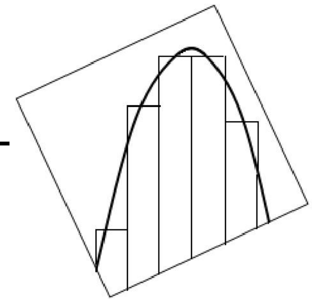
CHƯƠNG 1:

Các Mô hình Xác suất trong Kỹ thuật Điện và Kỹ thuật Máy tính



Ngày nay các kỹ sư thiết kế thường phải xây dựng những hệ thống làm việc trong các môi trường "hỗn loạn" (chaotic environment):

- Hệ thống máy tính lớn phải thoả mãn những yêu cầu khác nhau thay đổi bất thường của cộng đồng người mà hệ thống đó phục vụ.
- Mạng truyền thông phải luôn luôn sẵn sàng đối phó với những yêu cầu bất thường, hiếm gặp của khách hàng về "các kênh truyền tin (information pipeline)" đối với âm thanh và dữ liệu.
- Các hệ thống truyền thông phải cung cấp dịch vụ liên lạc thông suốt không gián đoạn và không gặp lỗi trên các kênh khó tránh khỏi nhiễu.
- Các hệ thống nhận dạng tiếng nói phải giải mã được tiếng nói từ đầu vào với độ tin cậy cao.

Các mô hình xác suất là một trong những công cụ cho phép kỹ sư thiết kế hiểu được hiện tượng hỗn độn, khó xác định quy luật nhằm xây dựng thành công các hệ thống chặt chẽ, đáng tin cậy và có hiệu quả về mặt kinh tế. Quyển sách này là phần mở đầu vào lý thuyết cơ sở của mô hình xác suất cũng như các kỹ thuật cơ bản được sử dụng trong việc phát triển những mô hình như vậy.

Chương này giới thiệu các mô hình xác suất và cho thấy sự khác biệt giữa chúng với các mô hình tất định – loại mô hình bao trùm trong kỹ thuật. Chúng tôi đề cập đến những tính chất quan trọng nhất của các khái niệm xác suất được phát triển, và vô vàn các ví dụ trong kỹ thuật điện và kỹ thuật máy tính, những lĩnh vực mà các mô hình xác suất đóng vai trò chủ đạo, cũng được trình bày. Phần 1.6 tóm tắt toàn bộ cuốn sách.

1.1 MÔ HÌNH TOÁN HỌC NHƯ LÀ CÔNG CỤ TRONG PHÂN TÍCH VÀ THIẾT KẾ

Việc thiết kế hoặc sửa đổi bất cứ một hệ thống phức tạp nào đều đòi hỏi việc chọn lựa từ nhiều phương án thay thế khả thi. Các lựa chọn đưa ra đều dựa trên các tiêu chí như như là chi phí, độ tin cậy và việc triển khai. Những phân tích định lượng các tiêu chí trên rất ít khi được tiến hành thông qua kiểm nghiệm trong thực tiễn hoặc thử nghiệm trên các mô hình lựa chọn. Trái lại, những quyết định lại được đưa ra căn cứ trên những ước lượng thu được nhờ sử dụng các mô hình của các phương án lựa chọn.

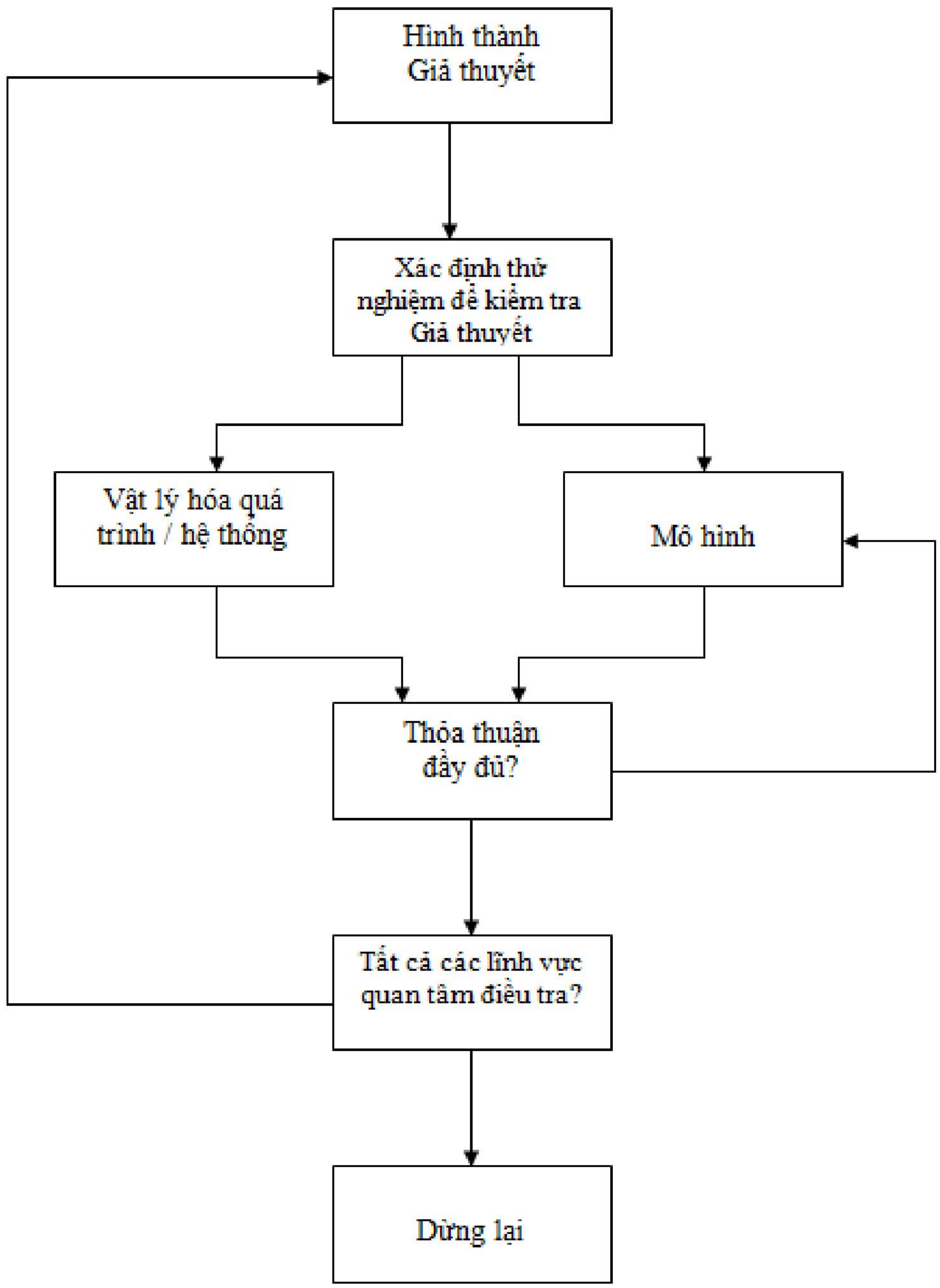
Mô hình (**model**) là biểu diễn gần đúng, tương đối chính xác của một hiện tượng vật lý. Mô hình cố gắng mô phỏng những thuộc tính quan sát được nhờ sử dụng tập hợp các luật mô tả đơn giản và dễ hiểu. Những luật đó được sử dụng để dự đoán kết quả các thí nghiệm liên quan đến các hiện tượng vật lý cho trước. Một ***mô hình hữu dụng*** (*useful* model) mô tả tất cả các mặt liên quan của hiện tượng xem xét. Do vậy, những mô hình đó có thể được sử dụng thay thế các thí nghiệm nhằm trả lời những câu hỏi liên quan tới thí nghiệm xem xét. Vì thế các mô hình cho phép kỹ sư tránh những chi phí cho thí nghiệm, tức là tiết kiệm lao động, trang thiết bị, và thời gian.

Mô hình toán học (**mathematical models**) được sử dụng khi hiện tượng quan sát được có những đại lượng thuộc tính là đo được. Một mô hình toán học là tập hợp các giả định về hoạt động của hệ thống hoặc quá trình vật lý. Các giả định đó nằm trong mẫu các quan hệ toán học chứa tham số quan trọng của hệ thống. Các điều kiện mà thí nghiệm đòi hỏi từ hệ thống được tiến hành xác định các giá trị cho trước trong các quan hệ toán học và việc giải quyết các mối quan hệ đó cho phép ta dự đoán các phép đo mà sẽ thu được nếu thí nghiệm được tiến hành.

Các mô hình toán học được các kỹ sư sử dụng một cách phổ biến trong các hệ thống chỉ dẫn thiết kế và các quyết định sử đổi. Trực giác và các kinh nghiệm thực tế, phương pháp thực nghiệm (rules of thumb) không phải bao giờ cũng đáng tin cậy trong những dự đoán việc sử dụng những hệ thống phức tạp và mới lạ cũng như trong việc thử nghiệm là không thể trong giai đoạn đầu của thiết kế hệ thống. Hơn nữa, chi phí cho những thử nghiệm đắt đỏ trong hệ thống hiện hành thường cho thấy là không thể chấp nhận được. Tính có giá trị của mô hình tương xứng hợp của các thành phần hệ thống phức tạp kết hợp với những tri thức về tương tác giữa chúng cho phép các khoa học gia và các kỹ sư phát triển toàn bộ mô hình toán học cho hệ thống. Do đó có thể đưa ra các câu trả lời cho việc sử dụng hệ thống phức tạp một cách nhanh chóng và không đắt đỏ. Thực vậy, các chương trình tìm lời giải cho các mô hình toán học tạo nền tảng cho các phân tích thiết kế hệ thống sử dụng máy tính.

Để trở nên hữu dụng, một mô hình phải thoả mãn các điều kiện thực tiễn của trường hợp quan sát. Từ đó quá trình phát triển và kiểm định mô hình nhất định phải bao hàm chuỗi các thử nghiệm và sửa đổi mô hình biểu diễn trên hình 1.1. Mỗi thí nghiệm xem xét những khía cạnh thực sự của hiện tượng qua nghiên cứu và đòi hỏi những quan sát và đo đạc phải được tiến hành trong những điều kiện xác định. Mô hình được sử dụng để dự báo hệ quả của thí nghiệm và những dự đoán đó được so sánh với những mẫu quan sát thu được khi tiến hành thử nghiệm trong thực tiễn. Nếu như phát sinh sự sai khác nghiêm trọng, mô hình được sửa đổi để làm rõ hiện tượng đó. Quá trình mô hình hoá kéo dài cho đến khi nhà nghiên cứu thấy thoả mãn khi những phương án giải quyết tất cả trường hợp của hiện tượng đều có thể được dự đoán với độ chính xác mong muốn. Cần phải nhấn mạnh rằng quyết khi nào cho dừng quá trình mô phỏng phụ thuộc vào các *mục đích nhất định* của nhà nghiên cứu. Do đó, một mô hình tương xứng với một lĩnh vực này có thể được chứng minh là hoàn toàn không thích hợp với trong các lĩnh vực khác.

HÌNH 1.1: Quá trình làm mẫu



Những dự đoán của mô hình toán học có thể được coi như giả thuyết cho đến khi nó được kiểm định qua đối chiếu với những đo đạc trong thực nghiệm. Trong thiết kế hệ thống, trường hợp sau có thể coi là dẫn vào ngõ cụt: không thể kiểm định mô hình bằng thực nghiệm được vì hệ thống đó trong thực tế không tồn tại. Các mô phỏng trên máy tính đóng vai trò quan trọng trong trường hợp này nhờ đưa ra những phương tiện thay thế để dự đoán tịnh chất của hệ thống và do đó các phương tiện xác minh những dự báo đó được xây dựng từ các mô hình toán học. Một mô phỏng bằng máy tính bao gồm chương trình mô phỏng hoặc bắt chước hoạt động của hệ thống. Kết hợp chặt chẽ với hệ thống là những hướng dẫn cho phép đo các tham số hoạt động liên quan. Một cách tổng quát, các mô hình trên máy có khả năng biểu diễn các hệ thống đó chi tiết hơn là các mô hình toán học. Tuy vậy chúng tỏ ra kém linh hoạt hơn và thường xuyên đòi hỏi nhiều thời gian tính toán hơn là các mô hình toán học.

Trong hai phần tiếp sau đây chúng ta sẽ đề cập đến hai dạng cơ bản của mô hình toán học: mô hình quyết định và mô hình xác suất.

1.2 MÔ HÌNH QUYẾT ĐỊNH

Trong mô hình quyết định các điều kiện mà trong đó thí nghiệm được tiến hành xác định chính xác kết quả của thí nghiệm. Trong mô hình quyết định toán học, nghiệm của hệ các phương trình toán học xác định rõ kết quả chính xác của thí nghiệm. Lý thuyết mạch là ví dụ về mô hình quyết định toán học.

Lý thuyết mạch mô phỏng sự liên kết trong sơ đồ của các thiết bị điện tử thông qua bảng mạch lí tưởng chứa các thành phần rời rạc đặc điểm về thế hiệu và dòng lí tưởng. Lý thuyết mạch giả định rằng tương tác giữa các thành phần lí tưởng đó hoàn toàn tuân theo các định luật về thế hiệu và dòng điện của Kirchhoff. Lấy ví dụ định luật Ohm phát biểu rằng quan hệ đặc thù giữa hiệu điện thế và dòng điện của điện trở là $I = V/R$. Hiệu điện thế và dòng điện trong một mạch điện bất kỳ chứa nguồn và các điện trở đều có thể xác định được bằng cách giải hệ các phương trình tuyến tính thu được nhờ áp dụng các định luật Kirchhoff và Ohm.

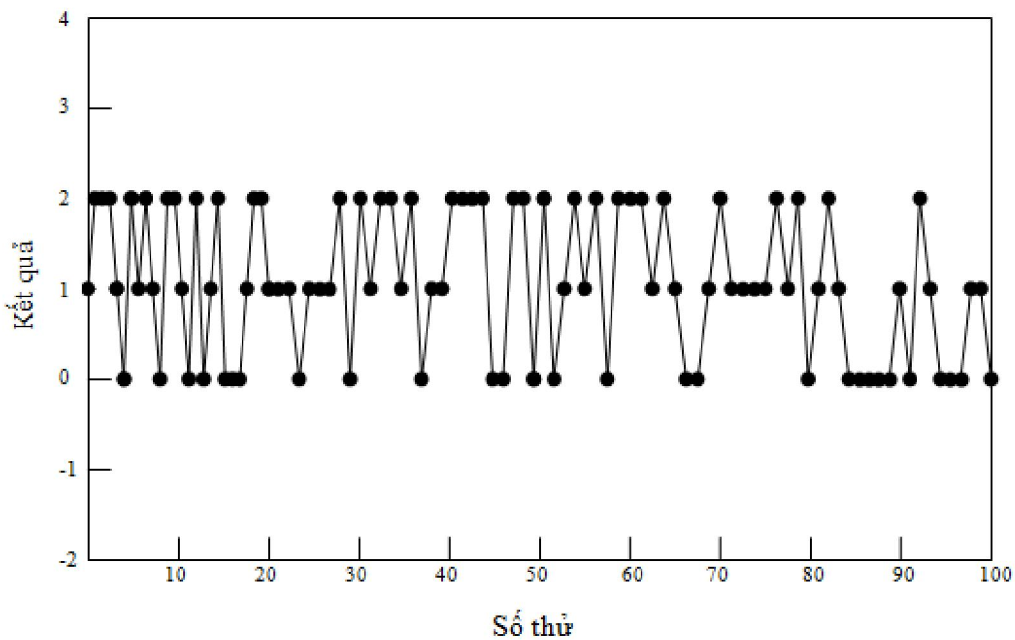
Nếu thí nghiệm đòi hỏi đo lường trên tập các hiệu điện thế lặp lại một số lần trong cùng điều kiện. Lý thuyết mạch dự báo rằng các kết quả thu được luôn không đổi. Trong thí nghiệm sẽ có sự thay đổi về kết quả đo đạc do sai số đo lường và các yếu tố không kiểm soát được.

1.3 MÔ HÌNH XÁC SUẤT

Nhiều hệ thống xem xét bao gồm các hiện tượng với các biểu hiện thay đổi một các ngẫu nhiên và không lường trước được. Chúng ta định nghĩa ***thí nghiệm ngẫu nhiên*** là thí nghiệm trong đó kết quả biến đổi theo cách không thể dự báo được khi thí nghiệm đó được lặp lại trong những điều kiện hệt trước. Các mô hình quyết định không thoả đáng với các thí nghiệm ngẫu nhiên do chúng dự báo một kết quả không đổi cho mỗi lần lặp thí nghiệm. Trong phần này chúng tôi giới thiệu các mô hình xác suất đã được phát triển cho các thí nghiệm ngẫu nhiên

Như là một ví dụ về thí nghiệm ngẫu nhiên, giả sử rằng quả bóng được chọn từ một cái bình chứa ba quả bóng giống nhau được đánh số 0, 1 và 2. Cái bình ban đầu được lắc để tạo sự ngẫu nhiên cho vị trí của các viên bi và sau đó chọn một viên. Số của viên bi được ghi lại và trả lại viên bi vào bình. Kết quả của thí nghiệm là số trong tập S = {0, 1, 2}. Chúng ta gọi tập S của tất cả các kết quả là ***không gian mẫu***. Hình 1.2 thể hiện kết quả của 100 lần lặp (phép thử) của chương trình máy tính mô phỏng thí nghiệm nhặt bi. Hoàn toàn hiển nhiên rằng kết quả kết quả của thí nghiệm này không thể tiên đoán một cách duy nhất một cách chính xác được.

HÌNH 1.2: Kết quả thử nghiệm URN



Tính đồng xác suất

Để trở nên hữu dụng, một mô hình phải cho phép chúng ta đưa ra những dự đoán về hoạt động trong tương lai của một hệ thống, và để cho trở nên dự báo được, một hiện tượng phải thể hiện tính đều đặn trong hoạt động của nó. Nhiều mô hình xác suất trong kỹ thuật dựa trên thực tế là trung bình thu được trong chuỗi dài các phép lặp (phép thử) của thí nghiệm ngẫu nhiên sinh ra một giá trị một cách nhất quán. Tính chất đó được gọi là ***tính đồng xác suất***.

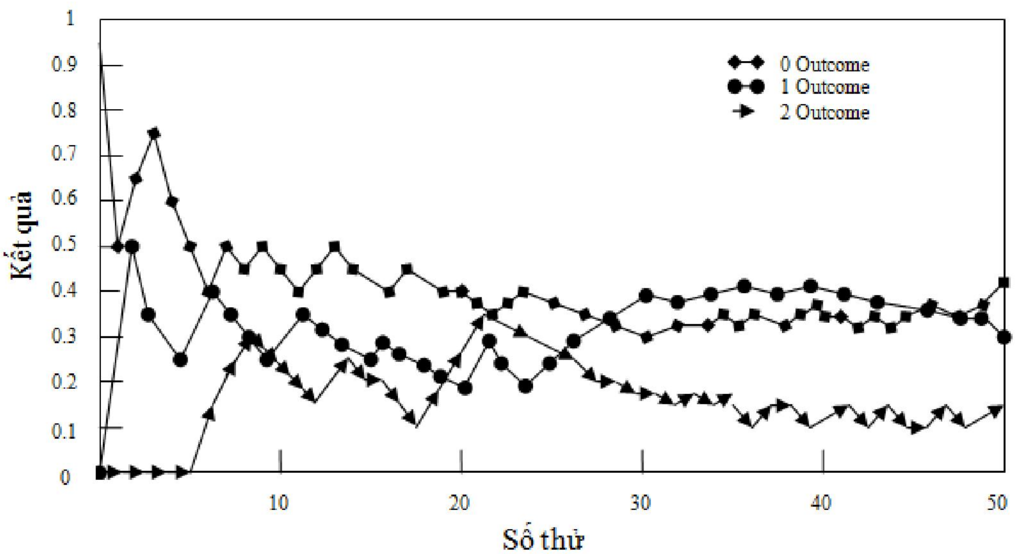
Giả định rằng thí nghiệm thí nghiệm với bình ở trên được lặp n lần dưới những điều kiện như nhau. Đặt $N_0(n)$, $N_1(n)$ và $N_2(n)$ là số lần trong đó kết quả là 0, 1 và 2 tương ứng và đặt tần số tương đối của kết quả k được định nghĩa bởi

$$f_k(n) = \frac{N_k(n)}{n} \qquad (1.1)$$

Trên cơ sở của tính đồng xác suất, chúng ta hiểu rằng $f_k(n)$ biến đổi càng ít hơn quanh một giá trị hằng khi n càng tăng, có nghĩa là

$$\lim_{n\to\infty} f_k(n) = p_k \qquad (1.2)$$

HÌNH 1.3: Kết quả của thử nghiệm URN



Hằng số p_k được gọi là ***xác suất*** của kết quả k. Phương trình (1.2) khẳng định rằng xác suất của kết quả là tỷ lệ của số lần xuất hiện trong chuỗi dài của các phép thử. Trong xuyên suốt cuốn sách chúng ta sẽ thấy rằng phương trình (1.2) đã mở ra con đường đi từ đo đạc các đại lượng vật lý cho đến các mô hình xác suất được đề cập trong cuốn sách. Hình 1.3 và 1.4 cho thấy các tần số tương đối của ba kết quả trong thí nghiệm nhặt bi nói trên khi số phép thử n tăng lên. Dễ thấy rằng tất cả các tần số tương đối đều hội tụ về 1/3. Điều đó không mâu thuẫn với "linh cảm" của chúng ta là cả ba kết quả là đồng xác suất.

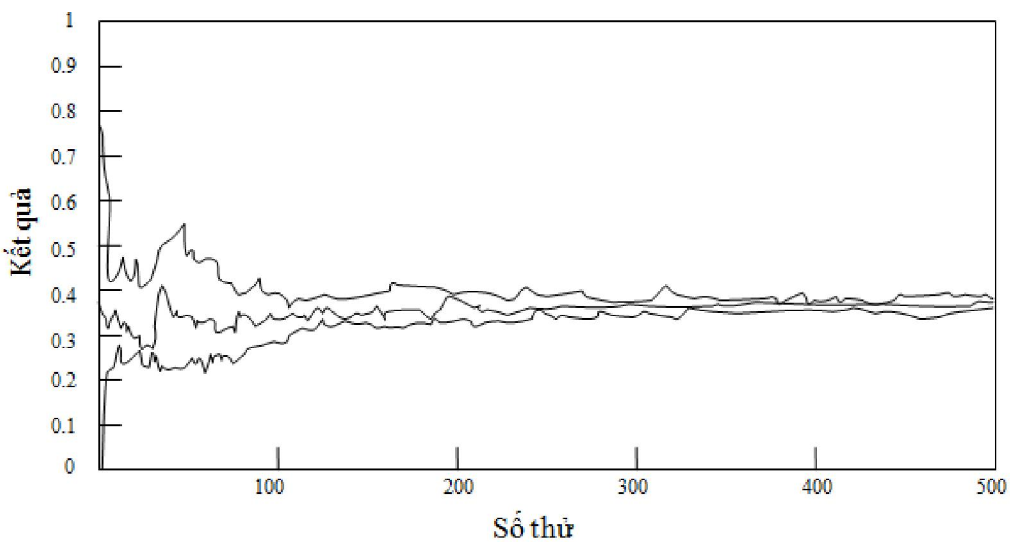
Giả sử rằng chúng ta thay thế thí nghiệm nhặt bi nói trên bằng cách cho thêm viên bi thứ tư giống chúng được đánh số 0. Xác suất của kết quả 0 bây giờ là 2/4 bởi vì hai trong số bốn viên bi trong bình được đánh số 0. Xác suất của kết quả 1 và 2 đều giảm xuống chỉ còn ¼. Điều này bộc lộ tính chất cơ bản của các mô hình xác suất, đó là: ***Các điều kiện mà trong đó một thí nghiệm ngẫu nhiên được tiến hành xác định xác suất của các kết quả của thí nghiệm***.

Các tính chất của tần số tương đối

Bây giờ chúng tôi sẽ giới thiệu vài tính chất của tần số tương đối. Giả sử rằng một thí nghiệm có K kết quả khả dĩ, có nghĩa là S = {1, 2,…, K}. Do số lần xuất hiện của bất kỳ kết quả nào trong n phép thử đều là số nằm giữa 0 và n, chúng ta có

$$0 \le N_k(n) \le n \qquad \text{với k = 1, 2,…, K}$$

HÌNH 1.4: Kết quả của thử nghiệm URN



và do đó khi chia các phương trình trên cho n, ta phát hiện ra rằng các tần số tương đối đều là số nằm giữa 0 và 1

$$0 \le f_k(n) \le 1 \qquad \text{với k = 1, 2,…, K} \qquad (1.3)$$

Tổng của số các lần xuất hiện của tất cả kết quả khả dĩ phải là n:

$$\sum_{k=1}^{K} N_k(n) = 1$$

Nếu chúng ta chia hai vế của phương trình trên cho n, chúng ta thấy rằng tổng của các tần số tương đối bằng 1:

$$\sum_{k=1}^{K} f_k(n) = 1 \qquad (1.4)$$

Đôi khi ta quan tâm đến sự xuất hiện của các sự kiện liên quan đến các kết quả của một thí nghiệm. Ví dụ xét sự kiện “xuất hiện một viên bi đánh số chẵn” trong thí nghiệm nói trên. Tần số tương đối của sự kiện đó sẽ là bao nhiêu? Sự kiện sẽ xuất hiện hễ khi số của viên bi là 0 hoặc 2. Số lần thí nghiệm trong đó kết quả là số chẵn do đó sẽ là $N_E(n) = N_0(n) + N_2(n)$. Tần số tương đối của sự kiện đó vì vậy là:

$$f_E(n) = \frac{N_E(n)}{n} = \frac{N_0(n) + N_2(n)}{n} = f_0(n) + f_2(n).$$

Ví dụ đó cho thấy rằng tần số tương đối của sự kiện là tổng các tần số tương đối của các kết quả liên quan. Một cách tổng quát hơn, đặt C là sự kiện "A hoặc B xuất hiện", trong đó A và B là hai sự kiện không thể xuất hiện đồng thời. Do đó số lần sự kiện C xuất hiện là $N_C(n) = N_A(n) + N_B(n)$ do đó

$$f_C(n) = f_A(n) + f_B(n) \tag{1.5}$$

Các phương trình (1.3), (1.4) và (1.5) là ba tính chất cơ bản của tần số tương đối mà từ đó ta sẽ thu được nhiều kết quả hữu ích.

Nghiên cứu lý thuyết xác suất thông qua các tiên đề

Phương trình (1.2) đưa ra giả thuyết là chúng ta định nghĩa xác suất của sự kiện thông qua tần số tương đối kéo dài theo thời gian của nó. Ta vấp phải những vấn đề khi sử dụng định nghĩa đó về xác suất với mục đích phát triển lý thuyết toán học về xác suất. Trước tiên là không rõ là khi nào và theo nghĩa nào về mặt toán học để giới hạn trong phương trình (1.2) tồn tại. Thứ hai, chúng ta không bao giờ có thể tiến hành một thí nghiệm với vô số phép thử, chính vì vậy chúng ta không bao giờ có thể biết được các xác suất p_k một cách chính xác. Cuối cùng, việc sử dụng tần số tương đối để định nghĩa xác suất sẽ làm hỏng khả năng ứng dụng của lý thuyết xác suất trong các trường hợp thí nghiệm không thể được lặp lại. Do đó nó làm cho ý nghĩa thực tiễn của việc phát triển lý thuyết toán học về xác suất không liên quan gì đến bất cứ một ứng dụng cá biệt nào hoặc bất cứ một ý niệm gì về những điều xác suất hàm ý. Mặt khác chúng tôi nhấn mạnh rằng một lý thuyết chính xác phải cho phép chúng ta sử dụng "linh cảm" và làm sáng tỏ xác suất như là tần số tương đối.

Để cho nhất quán với lời giải thích cho tần số tương đối, bất kỳ một định nghĩa nào về "xác suất của một sự kiện" phải thoả mãn các tính chất trong các phương trình (1.3) cho đến (1.5). Lý thuyết xác suất hiện đại bắt đầu với việc xây dựng tập các tiên đề để xác định rõ việc gán xác suất phải thoả mãn các tính chất đó. Lý thuyết đó giả định rằng: (1) cho trước một thí nghiệm ngẫu nhiên và S là tập của các kết quả cũng được chỉ rõ; (2) lớp các tập con của S, được gọi là các sự kiện, phải được xác định; (3) mỗi sự kiện A được gắn với một số P[A] sao cho các tiên đề sau được thoả mãn:

1. $0 \le P[a] \le 1$.

2. P[S] = 1.

3. Nếu A và B là các sự kiện xuất hiện không đồng thời thì

   P[A hoặc B] = P[A] + P[B].

Sự tương ứng giữa ba tiên đề với các tính chất của tần số tương đối được nêu trong các phương trình (1.3) đến (1.5) là hoàn toàn dễ thấy. Ba tiên đề này dẫn đến nhiều hệ quả có giá trị và rất mạnh. Thực vậy, chúng ta sẽ sử dụng toàn bộ phần còn lại của cuốn sách cho việc phát triển nhiều trong số những hệ quả đó.

Chú ý rằng lý thuyết xác suất không tự gắn nó với cách thu được các xác suất hoặc các xác suất đó có nghĩa gì. Bất kỳ việc gán xác suất với các sự kiện thoả mãn các tiên đề trên đều thoả đáng. Việc gán còn phụ thuộc vào người áp dụng lý thuyết, người thiết kế mô hình, để xác định việc gán xác suất phải làm như thế nào và việc làm sáng tỏ xác suất đều có nghĩa trong bất kỳ một ứng dụng nào.

Thiết kế mô hình xác suất

Chúng ta hãy cân nhắc làm sao xuất phát từ các bài toán thực tiễn hàng ngày có liên quan đến sự ngẫu nhiên đến mô hình xác suất cho bài toán đó. Lý thuyết đòi hỏi chúng ta phải xác định rõ các thành phần của các tiên đề trên. Điều này đòi hỏi:

(1) phải định nghĩa thí nghiệm ngẫu nhiên vốn có trong ứng dụng

(2) chỉ rõ tập S của các kết quả có thể và các sự kiện quan tâm

(3) chỉ rõ việc gán xác suất mà từ đó các xác suất của tất cả các sự kiện xem xét đều có thể tính toán trên máy được. Điều thách thức là phải phát triển một mô hình dễ nhất có thể được có thể giải thích được tất cả các khía cạnh liên quan của vấn đề trong hiện thực.

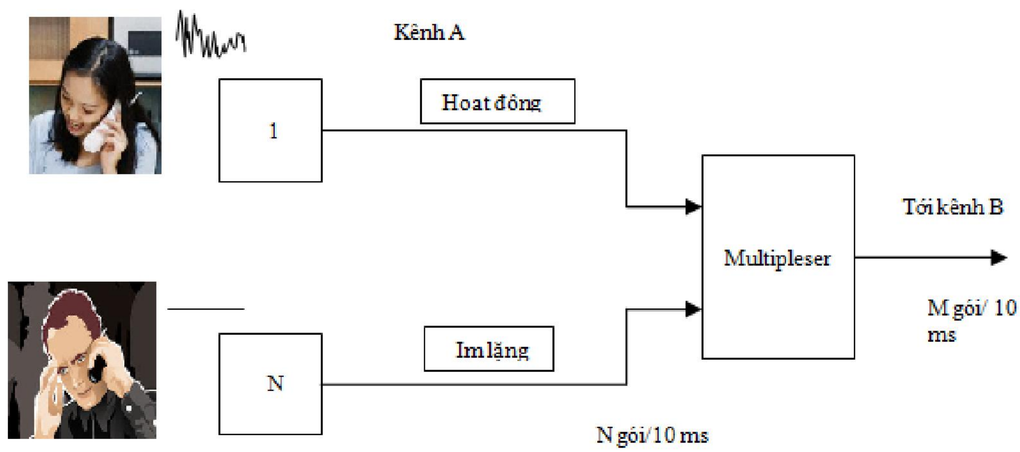
Như là một ví dụ, giả định rằng chúng ta thử nghiệm với cuộc trò chuyện qua điện thoại để xác định một người đang nói hay im lặng. Chúng ta biết rằng về trung bình một người sử dụng điện thoại bình thường chỉ dùng 1/3 thời gian cuộc gọi để nói; toàn bộ phần thời gian còn lại anh ta hoặc là nghe người còn lại hoặc ngừng giữa các từ và đoạn câu. Chúng ta có thể mô phỏng trường hợp này như là một thí nghiệm nhặt bi trong đó chúng ta nhặt bi từ một bình chứa hai bi trắng (im lặng) và một bi đen (đang nói). Chúng ta đang tiến hành đơn giản hoá rất lớn ở khâu này: không phải người dùng điện thoại nào cũng như nhau, không ngôn ngữ nào cũng cũng có hoạt động im lặng-nói như nhau, v.v… Tính có ích và sức mạnh của việc đơn giản hoá này trở nên rõ ràng khi chúng ta bắt đầu đặt câu hỏi mà sẽ được đặt ra khi thiết kế hệ thống đại loại như: Xác suất của sự kiện hơn 24 người dùng điện thoại trong số 48 người dùng độc lập cùng nói trong cùng một thời điểm là bao nhiêu? Câu hỏi này tương đương với: Xác suất của sự kiện hơn 24 viên bi trắng được chọn ra trong 48 lần lặp lại thí nghiệm nhặt bi một cách độc lập là bao nhiêu? Trước khi kết thúc chương 2 bạn đọc sẽ có thể trả lời câu hỏi sau và toàn bộ các bài toán hiện thực có thể quy về nó!

1.4 VÍ DỤ CỤ THỂ: HỆ THỐNG TRUYỀN GÓI ÂM THANH

Trong phần đầu của chương này, chúng tôi đã khẳng định rằng mô hình xác suất cho người thiết kế công cụ để xây dựng các hệ thống một cách thành công mà có tác dụng trong môi trường ngẫu nhiên, nhưng nó tuy thế mà phải hiệu quả, đáng tin cậy, và hiệu quả về kinh tế. Trong phần này chúng tôi giới thiệu một ví dụ cụ thể về một trong những hệ thống như vậy. Việc giới thiệu đó có mục đích là đánh thức tưởng tượng của bạn đọc. Nhiều kết quả thu được ở các bước có thể xuất hiện một cách không chính xác giờ sẽ được làm chính xác về sau trong cuốn sách.

Giả định rằng một hệ thống truyền tin phải tải 48 cuộc trò chuyện từ thành phố A đến thành phố B sử dụng các "gói" thông tin âm thanh. Lời của mỗi người trò chuyện được chuyển thành các hiệu điện thế có dạng hình sóng mà trước tiên sẽ được số hoá (có nghĩa là chuyển thành dãy các số nhị phân) và sau đó sẽ được gom lại trong các gói thông tin tương ứng với các đoạn 10 ms của câu chuyện. Địa chỉ nguồn và đích được thêm vào mỗi gói tin trước khi nó được chuyển (Xem hình 1.5).

HÌNH 1.5



Thiết kế đơn giản nhất cho hệ thống truyền tin này có thể tải 48 gói tin trong mỗi 10 ms theo mỗi hướng. Sơ đồ thiết kế như vậy là không hiệu quả, tuy vậy do ta biết rằng về trung bình có tới khoảng 2/3 của tất cả các gói chứa "im lặng" và do đó không có thông tin gì về lời nói. Nói cách khác, về trung bình trong số 48 người trò chuyện chỉ có 48/3 = 16 người đang nói trong mỗi gói với chu kỳ 10 ms. Từ đó chúng ta xem xét hệ thống khác mà chỉ tải M < 48 gói tin trong 10 ms.

Cứ 10 ms, hệ thống mới sẽ xác định những người trò chuyện nào đã "làm ra" các gói tin với lời nói. Gọi kết quả của thí nghiệm ngẫu nhiên này là A, số các gói tin có lời nói được sinh trong đoạn 10 ms. Đại lượng A nhận các giá trị từ 0 (tất cả các người trò chuyện đều im lặng) đến 48 (tất cả đều nói). Nếu A ≤ M, thì tất cả các gói đều được tải. Tuy vậy nếu A > M thì hệ thống sẽ không thể tải tất cả các gói chứa âm thanh vậy A – M gói tin có âm sẽ được chọn một cách ngẫu nhiên và bỏ bớt. Việc loại bỏ các gói tin có âm sẽ gây ra việc mất lời, vì vậy chúng ta sẽ muốn giữ lại phần nào đó của các gói tin bị loại đến mức mà người trò chuyện sẽ không thấy khó chịu.

Trước tiên xác định xác suất tương đối của A. Giả định rằng thí nghiệm nói trên được lặp n lần. Đặt A(j) là kết quả của lần thử thứ j. Đặt $N_k(n)$ là số các phép thử trong đó số các gói tin là k. Tần số tương đối của kết quả k trong n phép thử đầu tiên là $f_k(n) = N_k(n) / n$ trong đó chúng giả định rằng nó hội tụ về p_k :

$$\lim_{n\to\infty} f_k(n) = p_k \qquad 0 \le k \le 48 \qquad (1.6)$$

Tiếp theo xét tốc độ tạo ra các gói âm thanh. Số trung bình của các gói âm thanh được tạo ra mỗi khoảng thời gian 10 ms được cho bởi trung bình mẫu của số gói âm thanh:

$$\langle A\rangle_n = \frac{1}{n}\sum_{j=1}^{n} A(j) \qquad (1.7)$$

$$= \frac{1}{n}\sum_{k=0}^{48} kN_k(n) \qquad (1.8)$$

Biểu thức đầu cộng gói các số âm thanh được tạo ra trong n phép thử đầu tiên theo thứ tự quan sát được. Biểu thức thứ 2 tính xem co bao nhiêu quan sát có k gói âm thanh theo các giá trị có thể của k, và lấy tổng của k(1). Khi n tiến ra vô cùng. Tỷ số $N_k(n)/n$ trong biểu thức thứ 2 tiến tới p_k. Như vậy số trung bình các gói âm thanh được tạo ra mỗi 10 – ms tiến tới:

$$\langle A\rangle_n \to \sum_{k=0}^{48} kp_k \triangleq E[A] \qquad (1.9)$$

Biểu thức ở vế phải được định nghĩa là giá trị kỳ vọng của A trong phần 3.6. Chú ý rằng E[A] hoàn toàn được xác định bởi các xác suất p_k. Hệ thức (1.9) phát biểu rằng giá trị giới hạn của số trung bình của các gói âm thanh được tạo ra trong mỗi chu kỳ 10 – ms bằng E[A].

Tỷ số của số âm thanh bị loại bởi hệ thống trong n phép thử là:

---

(1) Giả sử bạn lấy từ túi của bạn như sau; 1 đồng 25 cent, 1 đồng 10 cent, 1 đồng 25 cent, 1 đồng 5 cent. Hệ thức (1.7) nói rằng tổng số của bạn là: 25 + 10 + 25 + 5 = 65 cent. Hệ thức (1.8) nói rằng tổng số của .... là : (1) 5 + (1/10 + (2/25) = 65 cent.

Số các gói âm thanh bị loại............

Từ (k – M) trong tử thức là số các gói bị loại khi số các gói âm thanh được tạo ra k > M. Nếu chúng ta chia tử thức và mẫu thức cho n và cho n tiến ra vô cùng, chúng ta một lần nữa nhận được một biểu thức của các xác suất p_k:

$$\frac{\sum_{k=M+1}^{48}(k-M)N_k(n)/n}{\sum_{k=0}^{48} kN_k(n)/n} \rightarrow \frac{\sum_{k=M+1}^{48}(k-M)p_k}{\sum_{k=0}^{48} kp_k} \quad (1.10)$$

Biểu thức ở vế phải là tỉ số giới hạn của số các gói âm thanh bị loại. Như vậy chúng ta đã chỉ ra rằng tất cả các độ đo mà chúng ta quan tâm trong bài toán này có thể tính được nếu ta biết các xác suất p_k của số gói âm thanh được tạo ra trong mỗi khoảng thời gian 10 – ms.

Nói chung cần phải thực hiện thí nghiệm một số lần để tính tần suất tương đối và dùng các tần suất này để ước lượng các xác suất p_k. Tuy nhiên trong bài tập thực tế này, nó chứng tỏ rằng điều này là không cần thiết do lý thuyết xác suất cho phép chúng ta tìm được xác suất p_k dựa vào xác suất của thí nghiệm đơn. Trong chương sau chúng ta nhận thấy rằng p_k được cho bởi phân phối nhị thức (công thức 2.32 chương 2).

Như vậy chúng ta nhận được biểu diễn chính xác cho tất cả các độ đo mà ta quan tâm trong hệ thống truyền tải âm thanh này.

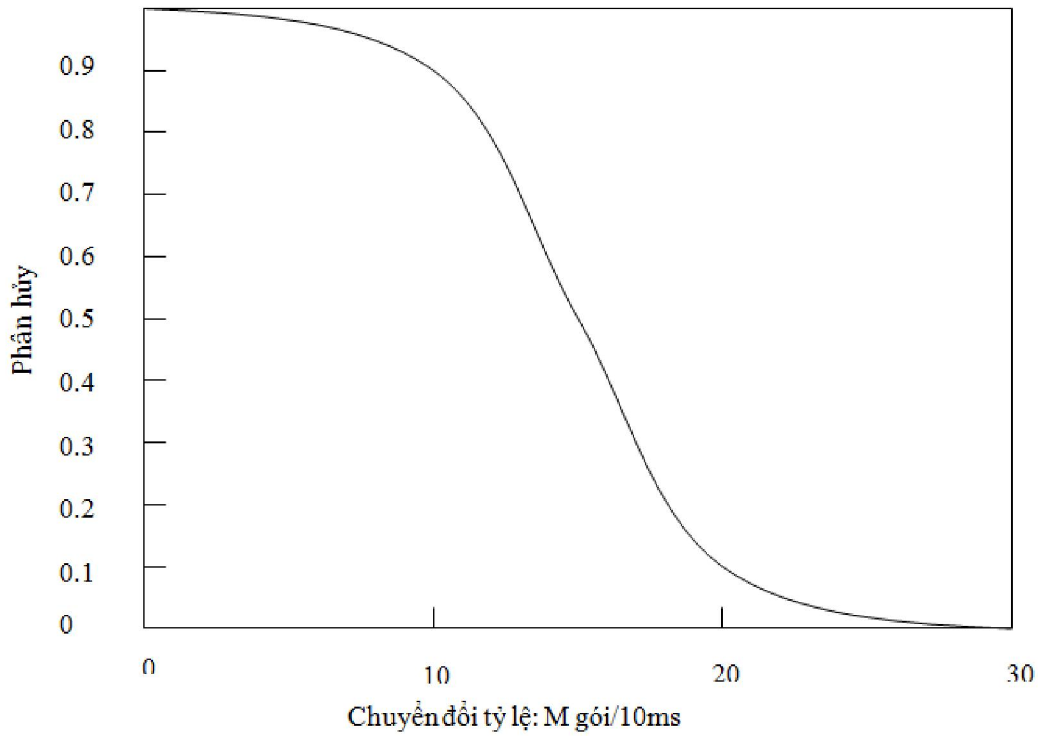
Hình 1.6 chỉ ra tỉ số giới hạn của các gói âm thanh bị loại trong hệ có 48 – bộ phát âm thanh với độ chính xác cao. Chúng ta có thể nhận thấy rằng các tỉ số này giảm khi M tăng. Nếu chúng ta giả thiết rằng các khách hàng sẽ cho phép mất 1% gói âm thanh, khi đó hình 1.6 chỉ ra rằng giá trị M cần thiết là 24. Như vậy chúng ta có thể đảm bảo việc truyền tải âm thanh cho tất cả 48 – bộ phát âm chỉ với 1 nửa tốc độ cần thiết mà vẫn đạt xấp xỉ tốt. Nếu chúng ta thực hiện với các cuộc điện đàm đường dài đắt đỏ, hệ thống mới có thể đưa đến một sự tiết kiệm có ý nghĩa.

Chúng ta tóm lược những điều chúng ta đã trình bày trong phần này. Chúng ta đã đưa ra một ví dụ mà ở đó sự hoạt động của hệ thống diễn ra một cách ngẫu nhiên, và ở đó các độ đo được phát biểu theo thuật ngữ các thành phần giới hạn. Chúng ta cũng chỉ ra rằng các độ đo giới hạn này dẫn tới các biểu thức chứa các xác suất của các kết cục khác nhau.Cuối cùng chúng ta đã chỉ ra rằng trong một số trường hợp lý thuyết xác suất cho phép tìm ra các xác suất này. Khi đó chúng ta có thể dự đoán được các trung bình giới hạn của các đại lượng khác nhau mà ta quan tâm và tiến hành việc thiết kế hệ thống.

HÌNH 1.6

Tỉ lệ của các gói âm thanh bị loại từ 48 – máy phát âm

với tốc độ truyền M, với M = 24, cho phép tỉ lệ mất âm là1%.



1.5 CÁC VÍ DỤ KHÁC

Trong phần này chúng ta sẽ trình bày thêm các ví dụ từ kỹ thuật điện và kỹ thuật máy tính, ở đó các mô hình xác suất được sử dụng để thiết kế các hệ thống làm việc trong mọi trường hợp ngẫu nhiên. Ở đay chúng ta muốn chỉ ra các xác suất và các trung bình giới hạn tiến đến một cách tự nhiên như thế nào tới các độ đo trong các hệ thống. Tuy nhiên chúng ta cũng lưu ý rằng, cuốn sách này chỉ trình bày các khái niệm cơ sở của lý thuyết xác suất mà không đi trình bày chi tiết các ứng dụng. Với các bạn đọc quan tâm, các tài liệu tham khảo cho việc đọc sâu hơn được giới thiệu ở phần cuối các chương.

Truyền thông trên kênh không tin cậy

Nhiều hệ thống truyền thông họat động theo cách sau. Mỗi T giây, máy phát nhận được tín hiệu vào nhị phân, ký hiệu là 0 và 1 và truyền đi tín hiệu tương ứng.

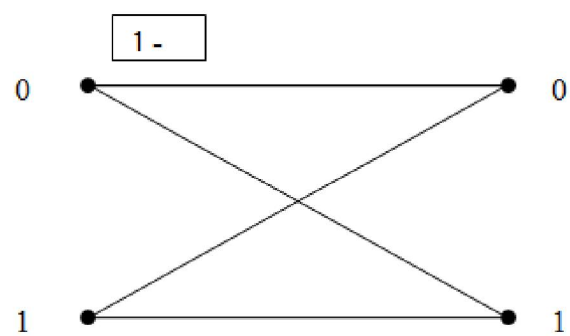
Vào thời điểm kết thúc T giây, máy thu sẽ đưa ra quyết định tín hiệu nào đã được đưa vào, dựa trên tín hiệu mà nó nhận được. Hầu hết các hệ truyền thông là không tin cậy do quyết định của máy thu không luôn trùng với tín hiệu vào của đường truyền. Hình 1.7(a) mô hình hóa các hệ truyền thông mà ở đó sai sót khi truyền xảy ra một cách ngẫu nhiên với xác suất ε. Như được chỉ ra trong hình vẽ, tín hiệu ra, khác với tín hiệu vào với xác suất là ε. Như vậy ε là tỉ lệ giới hạn của các bit được nhận sai bởi máy thu. Trong điều kiện mà ở đó tốc độ sai số này là không chấp nhận được, các kỹ thuật kiểm ta sai số được giới thiệu để giảm bớt sai số trong thông tin nhận được.

Một phương pháp giảm tốc độ sai số trong thông tin nhận được là dùng mà hiệu chỉnh sai số như được chỉ ra ở hình 1.7(b). Như một ví dụ đơn giản chúng ta xét mà lặp, ở đó mỗi bit thông tin được truyền lặp lại 3 lần:

$0 \rightarrow 0\ 0\ 0$

$1 \rightarrow 1\ 1\ 1$

Nếu chúng ta giả thiết rằng máy giải mã đưa ra quyết định về bit thông tin bởi cách lấy đa số trong 3 bit đầu ra nhận được bởi máy thu.

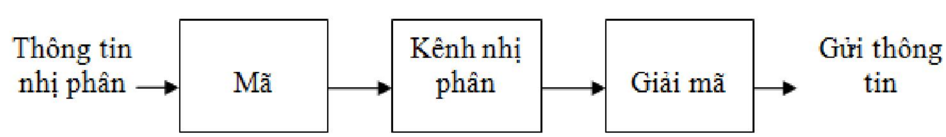


HÌNH 1.7 (a)

Mô hình một kênh truyền thông nhị phân

HÌNH 1.7(b)

Hệ kiểm tra sai số



Khi đó bộ giải mã sẽ đưa ra quyết định sai nếu 2 hoặc 3 bit nhận được bị sai. Trong ví dụ 2.37 chúng ta chứng tỏ rằng điều này xảy ra với xác suất $3\,\varepsilon^2 - 2\,\varepsilon^3$. Như vậy nếu tốc độ bit sai của kênh không mã là 10^{-3}, khi đó tốc độ bit sai của mã đơn giản ở trên là 3.10^{-6}, như vậy tốc độ bit sai được giảm đi 3 bậc. Tuy nhiên, sự cải thiện này nhận được với cái giá phải trả là : tốc độ truyền thông tin giảm xuống còn 1 bit mỗi 3 T giây. Bằng việc nghiên cứu tiếp những mã phức tạp hơn, nó cho phép giảm tốc độ sai số nhưng không làm giảm mạnh tốc độ truyền tin như trong ví dụ đơn giản này.

Xử lý các tín hiệu ngẫu nhiên

Kết cục của một thí nghiệm ngẫu nhiên không nhất thiết là một số đơn nhưng có thể là một hàm nhận giá trị nguyên của thời gian. Ví dụ kết cục của một thí nghiệm có thể là điện áp tương ứng với giọng nói hay điệu nhạc. Khi đó chúng ta sẽ quan tâm đến tính chất của tín hiệu và các phiên bản đã được xử lý của tín hiệu.

Như một ví dụ đặc biệt, giả sử chúng ta quan sát điện áp Y(t), kết quả của tổng 2 điện áp mà ta quan tâm S(t), (tín hiệu thông tin) với điện áp không mong muốn, N(t) ;(tín hiệu ồn). Ví dụ máy ghi âm từ, tín hiệu có thể là điện áp tương ứng với điệu nhạc và tiếng ồn, có thể là điện áp sẵn có của nó trong môi trường từ tính. Phép đo chất lượng của các hệ thống dạng này là tỉ lệ giữa tín hiệu và tiếng ồn (SNR) được xác định như là tỉ số giữa công suất trung bình của tín hiệu với công suất trung bình của tiếng ồn. Chất lượng của tín hiệu được quan sát được cải tiến khi, SNR tăng, do khi đo tiếng ồn, gây ra ảnh hưởng nhỏ hơn lên tín hiệu mong muốn.

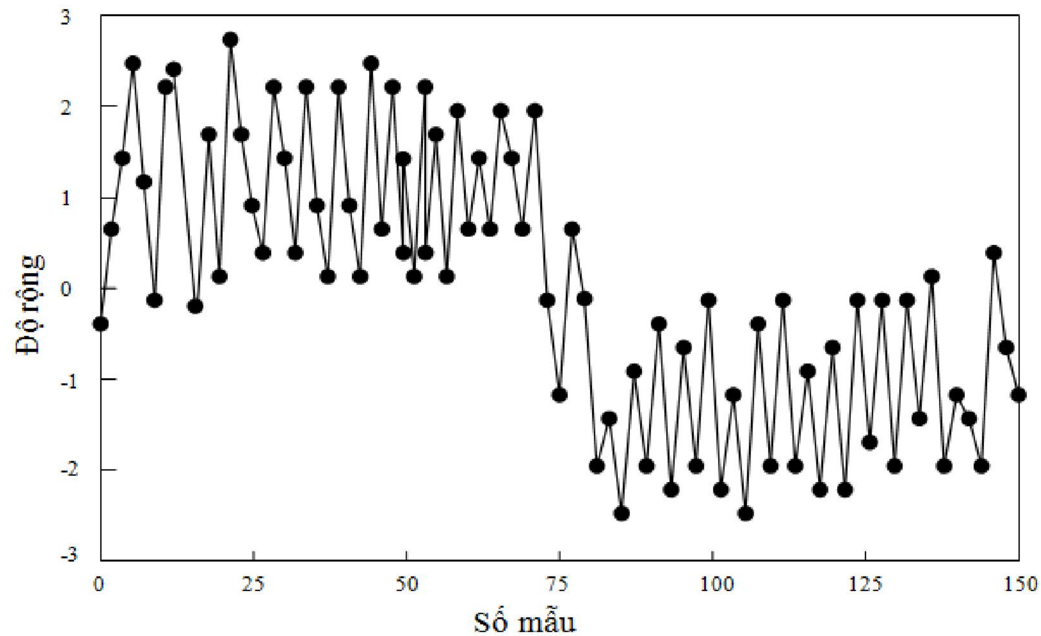
Các hệ thống có chung nguồn tài nguyên

Trong nhiều ứng dụng, các nguồn đắt đỏ như là kênh truyền thông và máy tính là đối tượng phục vụ vào những yêu cầu không ổn định và ngẫu nhiên. Những người dùng rải rác đưa ra những yêu cầu phục vụ trong những chu kỳ phục vụ ngắn giữa các chu kỳ nhàn rỗi dài. Những yêu cầu của người dùng có thể gặp được những nguồn đủ để đáp ứng tới mỗi người riêng lẻ. Tuy nhiên, cách tiếp cận này là cực kỳ lãng phí do các nguồn đáp ứng này không được sử dụng đến khi khách hàng không cần. Thách thức đối với người thiết kế là tạo ra hệ thống kinh tế và hiệu quả, mà ở đó yêu cầu của các khách hàng được đáp ứng thông qua sự chia sẻ năng động các nguồn tài nguyên .

Các hệ máy tính nhiều người dùng là một ví dụ của hệ thống có chung nguồn tài nguyên. Một hệ máy tính nguyên vẹn có thể coi như một nguồn tài nguyên riêng lẻ. Nếu chỉ một người dùng được phép sử dụng hệ thống, hệ luon phiên giữa chu kỳ, máy tính nhàn rỗi và đợi lệnh từ người dùng, với chu kỳ máy tính làm việc và người dùng chỉ đợi máy đáp ứng yêu cầu. Trong một ứng dụng điển hình, máy tính sử dụng được một lượng lớn thời gian nhàn rỗi.

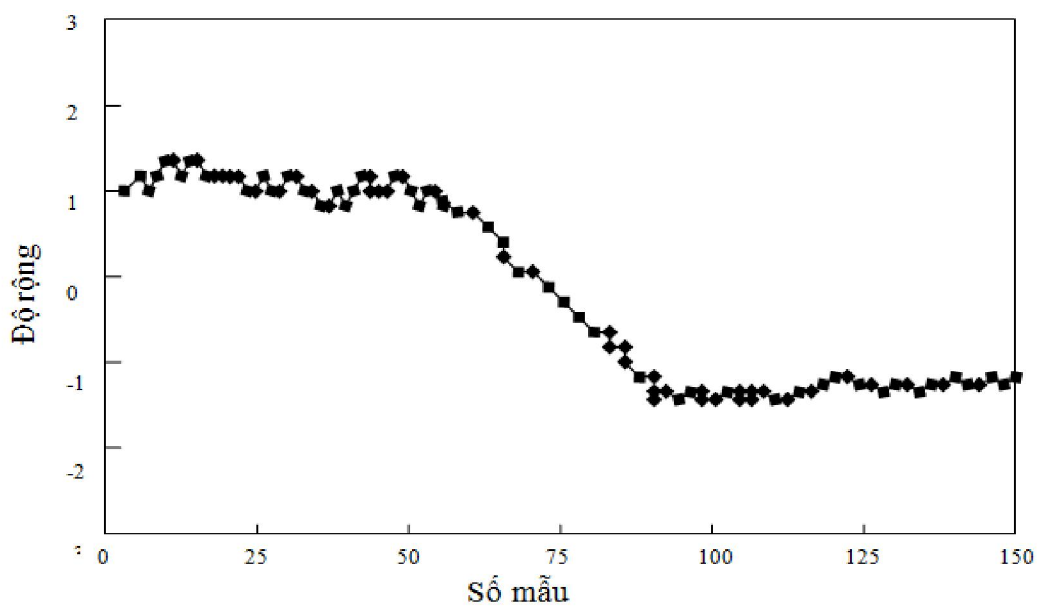
HÌNH 1.8(a)

Tín hiệu cộng tiếng ồn

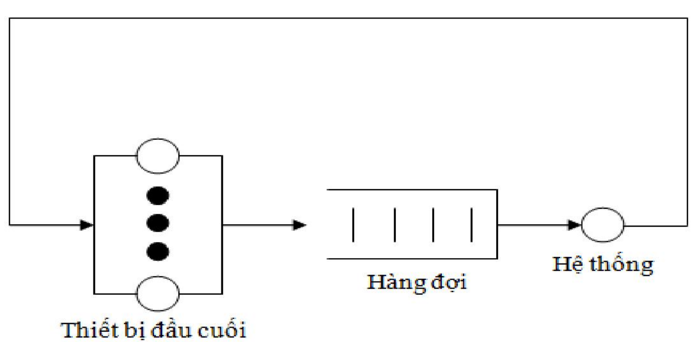


HÌNH 1.8(b)

Tín hiệu được lọc cộng tiếng ồn



HÌNH 1.9: Mô hình đơn giản cho một hệ máy tính nhiều người dùng

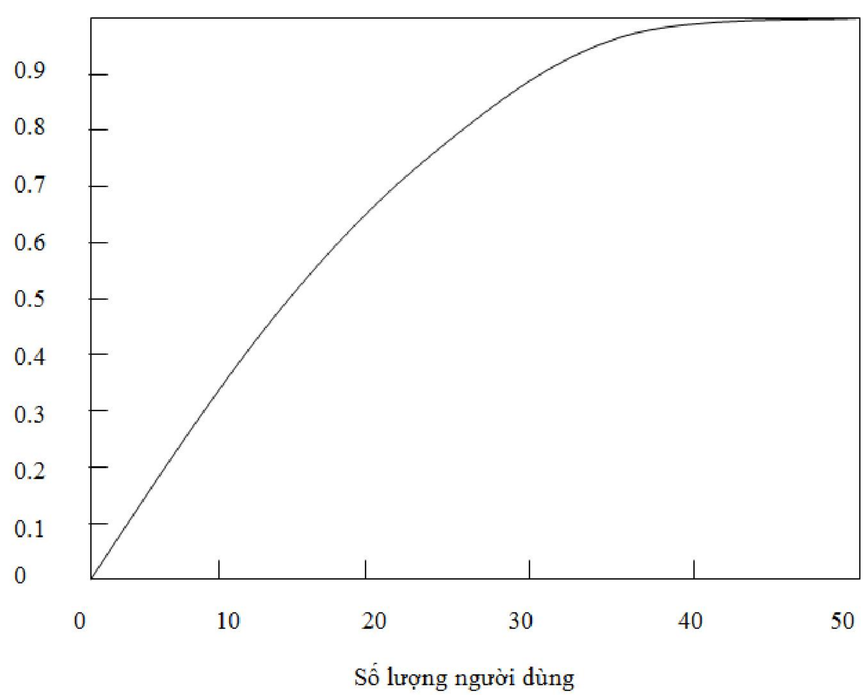


Để khắc phục tình trạng này, các máy tính thường được kết nối để dùng chung trong nhóm các người sử dụng. Hệ máy tính được thiết kế để điều hành số người dùng bằng việc đưa vào một hàng đợi mà ở đó các yêu cầu đợi được phục vụ từ máy tính như được chỉ ra trong hình 1.9. Tại một khoảng thời gian xác định bất kỳ, số các dùng đang trong quá trình chuẩn bị ra các mệnh lệnh và chờ sự đáp ứng từ hệ thống.

Độ đo ở đây, mà ta quan tâm là thời gian đáp ứng trung bình trội qua kể từ thời điểm người dùng đưa ra yêu cầu mọt thao tác đến thời điểm nhận được sự đáp yêu cầu và tốc độ trung bình để máy tính thực hiện xong một thao tác (khả năng thông qua). Các độ đo này có thể dự báo được khi dùng mô hình hàng đợi được thảo luận trong phần 9.5. Hình 1.10(a) và 1.10(b) chỉ ra thời gian đáp ứng trung bình và khả năng thông qua trung bình khi số người dùng trong hệ thống tăng lên. Kết quả nhận được đúng như mong đợi: Khi số người dùng tăng lên, thời gian mà hệ thống máy tính bận sẽ nhiều hơn, và do đó máy tính sẽ hoàn thành được nhiều công việc hơn mỗi giây, nhưng các công việc chờ được xử lý cũng tăng lên cùng với sự tăng lên của thời gian đáp ứang trung bình và sự không hài lòng của khách hàng cũng tăng lên.

HÌNH 10(a)

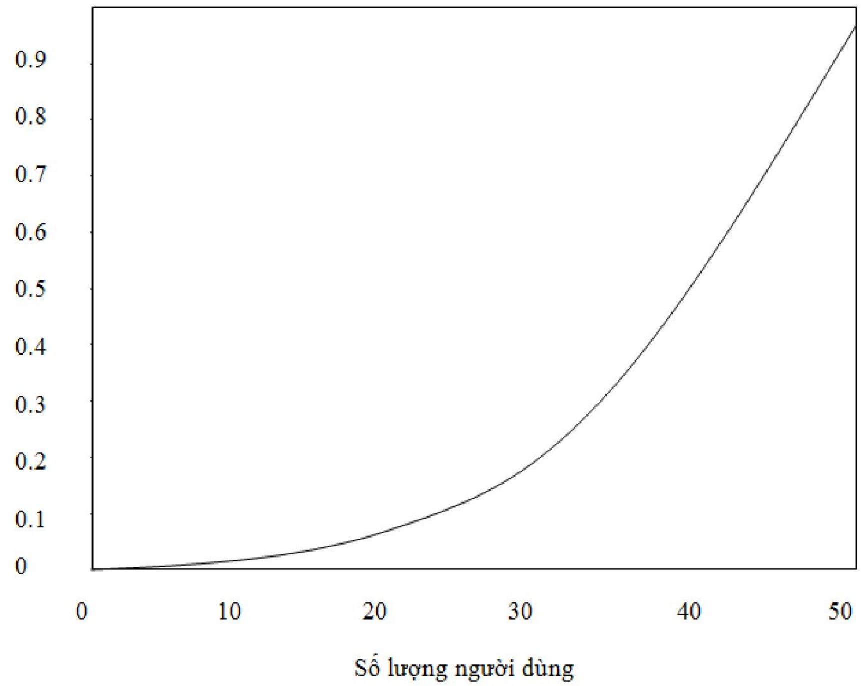
Hiệu suất thông qua của kênh nhiều người dùng



HÌNH 1.10(b)

Hệ số thời gian đáp ứng

của hệ nhiều người dùng



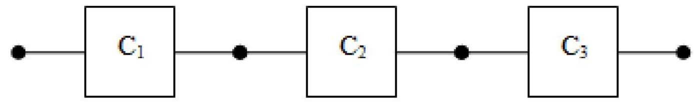
Độ tin cậy của hệ thống

Độ tin cậy là một khái niệm rất quan trọng trong thiết kế các hệ thống hiện đại. Ví dụ đầu tiên là hệ máy tính và các mạng truyền thông hỗ trợ việc chuyển khoản điện tử, giữa các ngân hàng. Vấn đề quan trọng là ở chỗ hệ thống vẫn tiếp tục làm việc ngay cả khi một phần hệ thống bị hỏng. Câu hỏi then chốt là, thiết một hệ tin cậy từ các thành phần không tin cậy như thế nào? Các mô hình xác suất cung cấp cho chúng ta các công cụ để trả lời câu hỏi này một cách định lượng

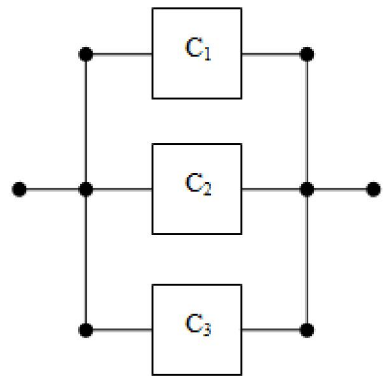
Sự hoạt động của một hệ thống yêu cầu sự hoạt của một số hoặc tất cả các thành phần của nó. Ví dụ hình 1.11(a) chỉ ra một hệ thống chỉ làm việc khi tất cả các thành phần của nó làm việc; hình 1.11(b) đưa ra một hệ thống làm việc khi còn ít nhất một thành phần của nó làm việc. Các hệ thống phức tạp hơn có thể nhận được như là tổ hợp của hai dạng thức cơ bản này.

Tất cả chúng ta đều biết từ kinh nghiệm là không thể dự báo chính xác khi nào một thành phần sẽ bị hỏng. Lý thuyết xác suất cho phép chúng ta tính được độ tin cậy, cũng như thời gian trung bình bị hỏng và xác suất để một thành phần còn làm việc sau khi đã hoạt động một khoảng thời gian nào đó. Hơn nữa chúng ta sẽ nhận thấy trong chương 2 và chương 3 là ... T

HÌNH 1.11
Các hệ với
n thành phần



(a) Hệ mắc nối tiếp



(b) Hệ mắc song song

Xác suất cho phép chúng ta xác định các giá trị trung bình và xác suất này của cả hệ thống dựa vào xác suất và giá trị trung bình của các thành phần của nó. Điều này cho phép chúng ta ước lượng được độ tin cậy của hệ thống, và chọn thiết kế hệ thống để đạt được độ tin cậy cần thiết.

1.6 TỔNG QUAN VỀ GIÁO TRÌNH

Trong chương này chúng ta đã thảo luận vai trò quan trọng của các mô hình xác suất trong việc thiết kế các hệ thống có sự tham gia của các yếu tố ngẫu nhiên. Nhiệm vụ của Giáo trình là giới thiệu cho sinh viên những khái niệm cơ sở của lý thuyết xác suất, những khái niệm cần thiết để hiểu được các mô hình xác suất trong kỹ thuật điện và kỹ thuật máy tính. Giáo trình không có ý định tập trung vào các ứng dụng, mà các ứng dụng thì có rất nhiều, mỗi ứng dụng đòi hỏi thảo luận chi tiết của riêng mình. Mặt khác chúng ta cố giữ các ví dụ liên quan tới các độc giả đã định hằng việc chọn lựa từ những lĩnh vực ứng dụng liên quan.

Mục tiêu khác của Giáo trình là trình bày một số kỹ thuật cơ sở cần thiết để phát triển các mô hình xác suất. Thảo luận ở chương này làm sáng tỏ một điều là các xác suất được sử dụng trong một mô hình cần phải được xác định thông qua thực nghiệm. Các kỹ thuật thống kê giúp chúng ta làm việc này, vì vậy chúng ta sẽ đưa vào một số thảo luận về kỹ thuật thống kê cơ sở nhưng bản chất chúng ta cũng sẽ đề cập đến sự có ích của các mô hình mô phỏng bằng máy tính khi xây dựng các mô hình xác suất. Hầu hết các chương đều có phần trình bày phương pháp máy tính hữu ích. Các phần này được đánh dấu bởi dấu hoa thị, và có thể bỏ qua.

Hơn nữa, sinh viên được khám phá các kỹ thuật này. Đây là một việc làm thú vị và chúng đưa đến một sự hiểu biết sâu sắc về bản chất của sự ngẫu nhiên.

Phần còn lại của cuốn sách được tổ chức như sau:

. Chương 2 trình bày các khái niệm cơ sở của lí thuyết xác suất. Chúng ta bắt đầu từ hệ tiên đề xác suất được phát biểu ở phần 1.3 và chúng ta sẽ thảo luận tính quan trọng của chúng. Một số mô hình xác suất cơ sở được giới thiệu trong chương 2.

. Nói chung lí thuyết xác suất không yêu cầu các kết cục của thí nghiệm ngẫu nhiên là một số. Khi đó các kết cục có thể là một sự vật (Ví dụ bóng đen hay trắng) hoặc hiện tượng (ví dụ hệ máy tính làm việc hay là không). Hơn nữa, chúng ta thường quan tâm đến những thí nghiệm mà ở đó kết cục là các số. Khái niệm biến ngẫu nhiên hướng đến các thí nghiệm này. Chương 3 và chương 4 thảo luận các thí nghiệm mà ở đó các kết cục là các số riêng lẻ hoặc các véc tơ số một cách tương ứng. Trong hai chương này chúng ta phát triển một số kỹ thuật đặc biệt hữu ích để giải các bài toán.

. Chương 5 trình bày các kết quả toán học (các định lí giới hạn) để trả lời câu hỏi điều gì sẽ xẩy ra khi các thí nghiệm độc lập lặp lại vô hạn lần. Các kết quả được trình bày sẽ chứng minh tính đúng đắn của việc dùng rộng rãi tần số tương đối để đưa ra khái niệm xác suất.

. Chương 6 giới thiệu khái niệm quá trình ngẫu nhiên, mà một thí nghiệm đơn giản là thí nghiệm mà ở đó kết cục là một hàm của thời gian.

. Chương 7 giới thiệu mật độ phổ công suất và ứng dụng của nó để phân tích và xử lí các tín hiệu ngẫu nhiên.

. Chương 8 thảo luận xích Markov, một dạng đặc biệt của quá trình ngẫu nhiên cho phép chúng ta mô hình hóa dãy các thí nghiệm không độc lập.

. Trình bày nhập môn lí thuyết hàng đợi và các ứng dụng khác nhau.

TÓM TẮT

. Các mô hình toán học liên quan đến các tham số và các biến quan trọng khi sử dụng các quan hệ toán học. Các mô hình này cho phép các nhà thiết kế hệ thống bằng việc sử dụng các phương trình khi các thí nghiệm là không khả thi hoặc quá đắt đỏ. **.** Các mô hình mô phỏng bằng máy tính là một sự lựa chọn thay thế khi dự báo hiệu suất hệ thống. Chúng có thể được sử dụng để công nhận các mô hình toán học. **.** Trong các mô hình tất định các điều	quyết định xác suất của các kết cục có thể. Nhiệm của các phương trình trong các mô hình xác suất là các xác suất của các kết cục và các biến cố cũng như các dạng khác nhau của các giá trị trung bình. **.** Các xác suất và các giá trị trung bình liên quan đến thí nghiệm ngẫu nhiên có thể tìm được một cách thực nghiệm bằng cách tính các tần suất và trung bình mẫu liên quan khi lặp lại thí nghiệm một số lớn lần. **.** Độ đo trong nhiều hệ thống ứng

kiện thực hiện thí nghiệm quyết định kết cục một cách chính xác. . Trong các mô hình xác suất các điều kiện thực hiện thí nghiệm ngẫu nhiên	dụng thực tiễn là tần suất tương đối và trung bình giới hạn. Các mô hình xác suất được sử dụng để thiết kế các hệ thống này.

DANH SÁCH CÁC THUẬT NGỮ QUAN TRỌNG

Mô hình tất định	Mô hình xác suất	Trung bình mẫu
Biến cố (sự kiện)	Thí nghiệm ngẫu nhiên	Không gian mẫu
Giá trị kỳ vọng	Tần suất tương đối	Sự chính quy thống kê
Xác suất		

TÀI LIỆU THAM KHẢO

1. M. E. Van Valkenburg, Network Anlysis, Prentice Hall, Englewood Cliffs, N. J., 1974.
2. L. Breiman, Probabylity and Stochastic Processes: With a View Toward Applications, Hougton Mifflin, Boston 1969.
3. P. L. Meyer, Introductory Probability and Statiscal Applications, Addison–Wesley, Reading, Mass., 1970.
4. W. B. Davenport, Probability and Random Processes: An Introduction for Applied Scientists and Engineers, McGraw–Hill, New York, 1970.
5. A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw–Hill, New York, 1965.
6. A. B. Clarke and R. L. Disney, Probability and Random processes: A. First Course with Applications, Wiley, New York, 1985.
7. C. W. Helstrom, Probability and Stochastic processes for

9. W. Feller, An Introduction to Probability Theory and Its Applications, Wiley, New York, 1968.
10. G. C. Clark and J. B. cain, Error–Correction Coding for Digital Communications, plenum Press, New York, 1981.
11. S. Lin and R. Costello, Error Cotrol Coding: Fundamentals and Applications, Prentice–all, Englewod Cliffs, N. J., 1983.
12. S. Haykin, Communication Systems, Wiley, New York, 1983.
13. N. Jayant and P. Noll, Digital Coding of Waveforms, Prentice–Hall, Englewod Cliffs, N. J., 1984.
14. A. V. Oppenheimand and R. W. Schafer, Digital Signal Processing, Prentice–Hall, Englewod Cliffs, N. J., 1975.
15. M.Schawrtz, Telecommuni–cation Networks: Protocols, Modeling, and Analysis, Addison–Wesley, Reading, Mass., 1987.
16. D. Bertsekas and R. G. Gallager, Data, Neworks,

Engineers, Macmillan, New York, 1984. 8. H. Cramer, mathematical Methods of Statistics, Princeton University Press, Princeton, N. J., 1946.	Prentice–Hall, Englewood Cliffs, N. J., 1987. 17. J. F. Hayes, Modeling and Analysis of Communications Netwoorks, Plenum Press, New York, 1984.

BÀI TẬP

1. Một thí nghiệm ngẫu nhiên là việc chọn hai quả bóng kế tiếp từ 1 hộp chứa hai quả bóng đen và một quả bóng trắng.
    a. Hãy biểu diễn không gian mẫu của thí nghiệm này.
    b. Giả sử rằng thí nghiệm được thay đổi như sau, quả bóng đầu tiên được hoàn lại vào hộp ngay sau khi được lấy. Khi đó không gian mẫu như thế nào?
    c. Tần suất của kết cục (trắng, trắng) bằng bao nhiêu nếu số lần lặp lại thí nghiệm là đủ lớn trong phần a? trong phần b?
    d. Kết quả của lần lấy thứ 2 phụ thuộc như thế nào vào kết cục của lần lấy thứ nhất trong cả hai thí nghiệm
2. Hãy giải thích các thí nghiệm sau tương đương như thế nào với thí nghiệm ngẫu nhiên hộp?
    a. Tung 1 đồng tiền cân đối.
    b. Tung một cặp xúc xắc cân đối.
    c. Lấy từ bộ bài 52 quân khác nhau.
3. Thí nghiệm ngẫu nhiên có không gian mẫu S = {a, b, c} và các xác suất p_a = 1/7, p_b = 2/7 và p_c = 4/7. Hãy mô tả một thí nghiệm ngẫu nhiên hộp có thể được dùng để mô phỏng thí nghiệm ngẫu nhiên này.
4. Với những điều kiện nào các thí nghiệm sau có thể được mô hình hóa bởi việc tung một đồng xu:
    a. Kết quả của việc kiểm tra một thiết bị điện tử trong một mạch gồm nhiều linh kiện
    b. Quan trắc tín hiệu ra nhị phân (1 hoặc 0) được tạo bởi một nguồn thông tin điện tử (ví dụ như máy tính hoặc máy fax):
5. Giả sử A là sự kiện gồm nhiều kết cục của một thí nghiệm ngẫu nhiên, và giả sử B là sự kiện ”xảy ra khi A không xảy ra”. Chứng tỏ rằng $f_B(n) = 1 – f_A(n)$.
6. Giả sử A, B và C là các sự kiện không đồng thời xảy ra đôi 1 hoặc bộ ba, và giả sử D là sự kiện ”A hoặc B hoặc C xảy ra”. Chứng tỏ rằng:
$$f_D(n) = f_A(n) + f_B(n) + f_C(n).$$
7. Giá trị trung bình mẫu của các dãy kết cục số X(1), X(2), ..., X(n) của dãy các thí nghiệm ngẫu nhiên được định nghĩa bởi:

$$\langle X\rangle_n = \frac{1}{n}\sum_{j=1}^{n} X(j)$$

Chứng tỏ rằng trung bình mẫu thỏa mãn công thức quy nạp sau:

$$\langle X\rangle_n = \langle N\rangle_{n-1} + \frac{X(n)-\langle X\rangle_{n-1}}{n} ;\ \langle X\rangle_0 = 0.$$

8. Giá trị trung bình bình phương mẫu của các kết cục số X(1), X(2),...,X(n) của dãy n thí nghiệm ngẫu nhiên được xác định bởi:

$$\langle X^2\rangle_n = \frac{1}{n}\sum_{j=1}^{n} X^2(j)$$

   a. bạn có thể hy vọng gì về sựm hội tụ của biến này tới một số khi n đủ lớn?
   b. Hãy tìm công thức quy nạp cho $\langle X^2\rangle_n$ tương tự như công thức trong bài tập 7.

9. Phương sai mẫu được định nghĩa như là giá trị trung bình bình phương của đọ lệch của mẫu khỏi trung bình mẫu:

$$\langle V^2\rangle_n = \frac{1}{n}\sum_{j=1}^{n}\left\{X(j)-\langle X\rangle_n\right\}^2 .$$

Chú ý rằng $\langle X\rangle_n$ cũng phụ thuộc vào các giá trị mẫu. Theo tập quán thay n trong mẫu số bởi n–1 bởi lý do kỹ thuật mà chúng ta sẽ xét ở chương 5. Từ giờ chúng ta sẽ sử dụng định nghĩa trên.

   a. Chứng tỏ rằng phương sai mẫu cũng có thể tính theo công thức:
   $\langle V^2\rangle_n = \langle X^2\rangle_n - \langle X\rangle_n^2$
   b. Chứng tỏ rằng phương sai mẫu thỏa mãn công thức truy hồi sau:
   $\langle V^2\rangle_n = \left(1-\frac{1}{n}\right)\langle V^2\rangle_{n-1} + \frac{1}{n}\left(X(n)-\langle X\rangle_{n-1}\right)^2$, với $\langle V^2\rangle_0 = 0$

10. Các số liệu sau nhận được bởi việc lấy mẫu từ một vôn kế:
    7, 3, –9, 4, 7, –2, –8, 4, 3, 4, –5, 5, 4, 1, –6, 3, –7, 1, –9, 0.

a. Hãy tìm tần suất tương đối của hiệu điện thế là dương.
b. Hãy tìm trung bình mẫu và phương sai mẫu.
c. Hãy tìm hàm phân phối thực nghiệm được xác định theo công thức:

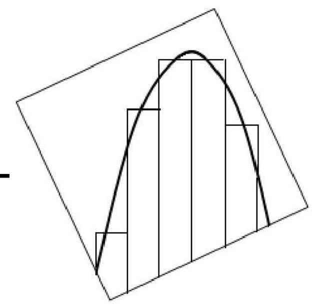
$$F(x) = \frac{So\ cac\ ket\ cuc\ nho\ hon\ hoac\ bang\ x}{Tong\ so\ cac\ ket\ cuc}; -\infty < x < \infty$$

11. Xét tập dữ liệu sau về thời gian làm việc (tính theo mi li giây) của một hệ máy tính: 14, 3, 11, 4, 12, 10, 2, 3, 7, 8, 14, 1, 5, 16, 14, 10, 2, 3, 20, 12.
    a. Hãy tìm trung bình mẫu và phương sai mẫu.
    b. Hãy tìm tần suất tương đối của sự kiện ”khoảng thời gian làm việc lớn hơn 10 ms”.
    c. Hãy vẽ hàm phân phối thực nghiệm được định nghĩa như bài toán trên.

12. Một người thợ có trách nhiệm sửa n máy trong một cửa hàng. Hãy giải thích hình 1.9 có thể được sử dụng để mô hình hóa sự thực hiện công việc của người thợ như thế nào?
13. Giả sử n thiết bị được lắp đặt theo mạng nối tiếp và song song như được chỉ ra trong hình 1.11. hãy giải thích tại sao mạng song song làm việc lâu hơn mạng nối tiếp.

CHƯƠNG 2:

Những Khái niệm Cơ sở của Lý thuyết Xác suất



Chương này trình bày những khái niệm cơ sở của lý thuyết xác suất. Trong phần còn lại của cuốn sách, chúng ta sẽ phát triển hơn nữa hoặc trình bày kỹ hơn nữa những khái niệm cơ sở được đưa ra ở đây. Bạn đọc sẽ được chuẩn bị tốt để tiếp thu những phần còn lại của cuốn sách, nếu có kiến thức đầy đủ những khái niệm cơ sở ở đây sau khi đọc xong chương này.

Sau đây những khái niệm cơ sở sẽ được trình bày. Trước hết, lý thuyết tập hợp đã được dùng để mô tả không gian mẫu và các biến cố của thí nghiệm ngẫu nhiên. Thứ hai, các tiên đề của xác suất chỉ rõ các tính cách xác suất của các biến cố. Thứ ba, khái niệm xác suất có điều kiện cho phép chúng ta xác định thông tin một phần về kết cục của một phép thử ngẫu nhiên ảnh hưởng như thế nào đến xác suất của biến cố. Xác suất có điều kiện cũng cho phép chúng ta đưa ra khái niệm độc lập của các biến cố và thí nghiệm. Cuối cùng, chúng ta xét dãy các thí nghiệm ngẫu nhiên được tạo thành bởi một dãy các thí nghiệm ngẫu nhiên đơn giản. Chúng ta cũng chỉ ra xác suất của các biến cố trong các thí nghiệm này có thể được xác định như thế nào từ các xác suất của các thí nghiệm đơn giản. Ngoài ra, cuốn sách cũng chỉ ra rằng các thí nghiệm ngẫu nhiên phức tạp có thể phân tích thành các thí nghiệm đơn giản.

2.1 MÔ TẢ CÁC THÍ NGHIỆM NGẪU NHIÊN

Thí nghiệm ngẫu nhiên là thí nghiệm mà ta không thể nói chắc kết cục của nó khi thí nghiệm được lặp đi lặp lại trong cùng một điều kiện. *Thí nghiệm ngẫu nhiên là một cách tiến hành thí nghiệm và tập hợp gồm một hay nhiều phép đo hoặc quan trắc.*

VÍ DỤ 2.1 *Thí nghiệm E_1*: Lấy 1 viên bi từ hộp gồm những viên bi được đánh số từ 1 đến 50. Ghi số của viên bi lại.

Thí nghiệm E_2: Lấy một viên bi từ hộp được đánh số từ 1 đến 4. Giả sử rằng những viên bi 1 và 2 có màu đen, viên bi 3 và 4 có màu trắng. Ghi lại số và màu của viên bi được lấy.

Thí nghiệm E_3: Tung đồng xu 3 lần và ghi dãy mặt ngửa và mặt sấp xuất hiện.

Thí nghiệm E_4: Tung đồng xu 3 lần và ghi số lần xuất hiện mặt ngửa.

Thí nghiệm E_5: Đếm số gói âm thanh im lặng từ nhóm n điện thoại trong chu kỳ thứ 10.

Thí nghiệm E_6: Một block thông tin được phát lặp lại trên kênh có nhiễu cho đến khi một block không sai được gửi tới nơi nhận. Đếm số lần phát block thông tin cần thiết.

Thí nghiệm E_7: Chọn một số một cách ngẫu nhiên giữa 0 và 1.

Thí nghiệm E_8: Đo thời gian giữa hai lần điện báo đến tại một trung tâm điện báo.

Thí nghiệm E_9: Đo thời gian sống của một chip máy tính ở trong một điều kiện môi trường nào đó.

Thí nghiệm E_{10}: Xác định giá trị của một vôn kế tại thời điểm t_1.

Thí nghiệm E_{11}: Xác định giá trị của một vôn kế tại thời điểm t_1 và t_2.

Thí nghiệm E_{12}: Chọn hai số một cách ngẫu nhiên giữa 0 và 1.

Thí nghiệm E_{13}: Lấy một số X một cách ngẫu nhiên giữa 0 và 1, sau đó lấy 1 số Y một cách ngẫu nhiên giữa 0 và X.

Thí nghiệm E_{14}: Một hệ thống được lắp ráp tại thời điểm $t = 0$. Với $t \geq 0$, $X(t) = 1$ cho đến khi các thành phần còn làm việc, và $X(t) = 0$ sau khi thiết bị không làm việc.

Xác định thí nghiệm ngẫu nhiên cần phải phát biểu rõ ràng, thực hiện phép đo hay là quan trắc. Ví dụ, thí nghiệm ngẫu nhiên có thể là trên cùng một sản phẩm nhưng khác nhau các quan trắc, như được minh họa ở thí nghiệm E_3 và thí nghiệm E_4.

Một thí nghiệm ngẫu nhiên có thể bao gồm nhiều hơn một phép đo hoặc quan trắc, như là các thí nghiệm E_2, E_3, E_{11}, E_{12} và E_{13}. Một thí nghiệm thậm chí bao gồm tập continum các phép đo như được chỉ trong thí nghiệm E_{14}.

Các thí nghiệm E_3, E_4, E_5, E_6, E_{12} và E_{13} là ví dụ về các thí nghiệm liên tiếp mà nó được quan sát như là dãy các thí nghiệm đơn giản. Bạn có thể xác định được các thí nghiệm đơn giản trong các thí nghiệm này hay không? Chú ý rằng, trong thí nghiệm E_{13}, thí nghiệm nhỏ thứ 2 phụ thuộc vào kết cục của thí nghiệm nhỏ thứ nhất.

Không gian Mẫu

Do các thí nghiệm ngẫu nhiên không thể gồm các kết cục giống nhau, cho nên cần thiết phải xác định tập các kết cục có thể. Chúng ta định nghĩa một **kết cục** hay **một điểm mẫu** của một thí nghiệm ngẫu nhiên như là một kết cục không thể phân chia thành các kết cục khác. Khi chúng ta thực hiện thí nghiệm ngẫu nhiên, thì 1 và chỉ 1 kết cục xảy ra. Như vậy các kết cục là xung khắc nhau theo nghĩa là chúng không thể xảy ra đồng thời. **Không gian mẫu** S của một thí nghiệm ngẫu nhiên được xác định như là tập tất cả các kết cục có thể.

Chúng ta sẽ ký hiệu một kết cục của một thí nghiệm bởi ζ , ở đây ζ là một phần tử hay một điểm của S. Mỗi lần thực hiện một thí nghiệm ngẫu nhiên có thể coi như là một phép chọn ngẫu nhiên một điểm (kết cục) riêng lẻ từ S.

Không gian mẫu có thể được mô tả hoàn toàn bởi việc sử dụng các ký hiệu tập hợp. Nó có thể mô tả bởi các bảng vẽ, các sơ đồ, các khoảng của đường thẳng thực, hoặc là các miền của mặt phẳng.

VÍ DỤ 2.2 Các không gian mẫu tương ứng với các thí nghiệm trong ví dụ 2.1 được đưa ra với việc dùng các ký hiệu tập hợp sau:

$S_1 = \{1, 2, \ldots, 50\}$

$S_2 = \{(1, b), (2, b), (3, w), (4, w)\}$

$S_3 = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

$S_4 = \{0, 1, 2, 3\}$

$S_5 = \{0, 1, 2, 3, \ldots, N\}$

$S_6 = \{1, 2, 3,\ldots\}$

$S_7 = \{x : 0 \le x \le 1\} = [0, 1]$ Xem Hình 2.1(a).

$S_8 = \{t : t \ge 0\} = [0, \infty)$

$S_9 = \{t : t \ge 0\} = [0, \infty)$ Xem Hình 2.1(b).

$S_{10} = \{v : -\infty < v < \infty \} = (\infty, \infty)$

$S_{11} = \{(v_1, v_2) : -\infty < v_1 < \infty$ và $-\infty < v_2 < \infty\}$

$S_{12} = \{(x, y) : 0 \le x \le 1$ và $0 \le y \le 1\}$ Xem Hình 2.1(c).

$S_{13} = \{(x, y) : 0 \le y \le x \le 1\}$ Xem Hình 2.1(d).

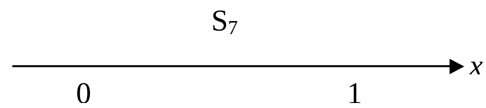
S_{14} = Tập hợp các hàm X(t) mà nó thoả mãn X(t) = 1 với $0 \le t < t_0$ và X(t) = 0 với $t \ge t_0$, ở đây $t_0 > 0$ là thời điểm tại đó chi tiết máy không làm việc.

Các thí nghiệm ngẫu nhiên có cùng một kết quả thí nghiệm có thể có các không gian mẫu khác nhau như đã được chỉ ra trong các thí nghiệm E_3 và E_4. Như vậy mục đích của thí nghiệm ảnh hưởng đến việc chọn không gian mẫu.

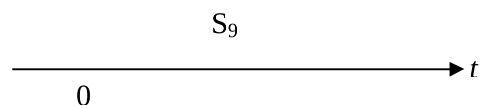
Có ba khả năng có thể xảy ra về số lượng các kết cục trong không gian mẫu. Một không gian mẫu có thể hứa hẹn vô hạn đếm được hoặc vô hạn không đếm được. Chúng ta sẽ gọi S là không gian mẫu đếm được, nghĩa là các kết cục của nó có thể được tương ứng 1 – 1 với các số nguyên dương. Chúng ta sẽ gọi S là không gian mẫu liên tục nếu S là không đếm được. Các thí nghiệm E_1, E_2, E_3, E_4 và E_5 có không gian mẫu rời rạc hữu hạn. Thí nghiệm E_6 có không gian mẫu vô hạn đếm được. Các thí nghiệm E_7 và E_{13} có không gian mẫu liên tục.

HÌNH 2.1

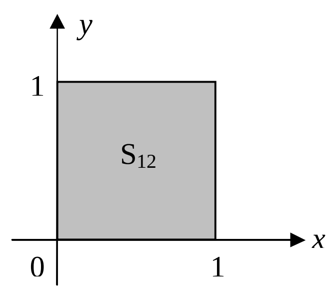
Không gian mẫu của các Thí nghiệm E_7, E_9, E_{12}, và E_{13}.



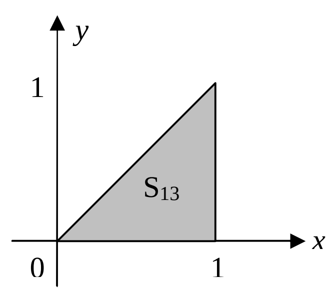
(a) Không gian mẫu cho Thí nghiệm E_7.



(b) Không gian mẫu cho Thí nghiệm E_9.



(c) Không gian mẫu cho Thí nghiệm E_{12}.



(d) Không gian mẫu cho Thí nghiệm E_{13}.

Do một kết cục của một thí nghiệm ngẫu nhiên có thể gồm 1 hoặc hơn 1 phép đo hoặc quan trắc, không gian mẫu S có thể nhiều chiều. Ví dụ, các kết cục trong các thí nghiệm E_2, E_{11}, E_{12} và E_{13} là hai chiều và các kết cục trong thí nghiệm E_3 là 3 chiều. Trong một số trường hợp, không gian mẫu có thể được viết như là tích Descartes của các tập hợp khác [1]. Ví dụ, $S_{11} = R \times R$, ở đây R là tập các số thực, và $S_3 = S \times S \times S$, ở đây S = {H; T}.

Để thuận tiện ta lấy cả không gian mẫu bao gồm cả những kết cục không thể. Ví dụ, trong ví dụ E_9 để thuận tiện ta xác định không gian mẫu như là đường thẳng thực dương, mặc dù rằng thiết bị không thể có thời gian sống vô hạn.

Các Biến cố

Chúng ta thường không quan tâm đến sự xảy ra của từng kết cục riêng lẻ, mà quan tâm đến sự xảy ra của một số các kết cục (Ví dụ khi kết cục thoả mãn những điều kiện đã biết nào đó). Ví dụ trong thí nghiệm E_{10} là phép đo của một vôn kế, chúng ta phải quan tâm đến biến cố "vôn kế chỉ số âm". Các điều kiện mà ta quan tâm xác định tập hợp con của không gian mẫu, được gọi là tập hợp các điểm ζ của S thoả mãn các điều kiện đã cho. Ví dụ, "vôn kế chỉ số âm" tương ứng với tập hợp $\{\zeta : -\infty < \zeta < 0\}$. Biến cố này xảy ra nếu và chỉ nếu kết cục của thí nghiệm ζ phụ thuộc vào tập con này. Với lý do đó chúng ta định nghĩa một **biến cố** như là một tập con của S.

Có hai biến cố được quan tâm đặc biệt đó là **biến cố chắc chắn** S, là biến cố bao gồm tất cả các kết cục và do đó luôn luôn xảy ra, và **biến cố không thể** hoặc **biến cố không**, $\varnothing$, mà nó không chứa một kết cục nào cả và do đó chúng ta định nghĩa một biến cố như là một tập con của S.

VÍ DỤ 2.3 Trong các thí nghiệm sau A_k là biến cố tương ứng với thí nghiệm E_k trong Ví dụ 2.1.

E_1: "Quả bóng có số chẵn được chọn", $A_1 = \{2, 4, \ldots, 50\}$.

E_2: "Quả bóng có màu trắng và số chẵn được chọn", $A_2 = \{(4, w)\}$.

E_3: "Ba lần tung cho kết cục như nhau", $A_3 = \{HHH; TTT\}$.

E_4: "Số lần xuất hiện mặt ngửa bằng số lần xuất hiện mặt sấp", $A_4 = \varnothing$.

E_5: "Không có một gói âm nào tạo ra", $A_5 = \{0\}$.

E_6: "Ít hơn 10 phép chuyển đã được thực hiện", $A_6 = \{1, 2, \ldots, 9\}$.

E_7: "Số được lấy ra là một số không âm", $A_7 = S_7$.

E_8 : "Khoảng thời gian giữa hai lần điện báo đến nhỏ hơn to"

$$A_8 = \{t: 0 \le t < t_0\} = [0, t_0).$$

E_9: "Chip làm việc nhiều hơn 1000 giờ nhưng nhỏ hơn 1500 giờ"

$$A_9 = \{t: 1000 < t < 1500\} = (1000, 1500).$$

(1) Tích Descartes của các tập hợp A và B là tập các cặp có thứ tự (a, b), với thành phần thứ nhất thuộc A và thành phần thứ hai thuộc B.

E_{10}: "Giá trị tuyệt đối của vôn kế là nhỏ hơn 1 vôn"

$$A_{10} = \{v: -1 < v < 1\} = (-1, 1).$$

E_{11}: "Hai vôn kế có dấu ngược nhau", $A_{11} = \{(v_1, v_2): (v_1 < 0$ và v_2

> 0) hoặc ($v_1 > 0$ và $v_2 < 0$)}.

E_{12}: "Hai số khác nhau một lượng ít hơn 1/10", $A_{12} = \{(x, y) : (x, y)$ thuộc S_{12} và $| x - y | < 1/10\}$.

E_{13}: "Hai số khác nhau một lượng ít hơn 1/10", $A_{13} = \{(x, y) : (x, y)$ thuộc S_{13} và $| x - y | < 1/10\}$.

E_{14}: "Hệ làm việc tại thời điểm t_1'', A_{14} = tập con của S_{14} mà ở đó $X(t_1) = 1$.

Một biến cố có thể chỉ gồm một kết cục riêng lẻ, như các biến cố A_2 và A_5. Một biến cố từ không gian mẫu rời rạc chỉ gồm một kết cục riêng lẻ được gọi là **biến cố sơ cấp**. Các biến cố A_2 và A_5 là các biến cố sơ cấp. Một biến cố cũng có thể là cả không gian mẫu như biến cố A_7. Biến cố không, $\varnothing$, xảy ra khi không có kết cục nào thoả mãn các điều kiện của biến cố này, như biến cố A_4.

Các Phép toán Tập hợp

Chúng ta cũng có thể kết hợp các biến cố bằng việc dùng **các phép toán tập hợp** để nhận được các biến cố khác. Chúng ta cũng có thể tạo ra các biến cố phức tạp như là tổ hợp các biến cố đơn giản.

Phép hợp của hai biến cố A và B được ký hiệu bởi $A \cup B$ và được xác định như là tập hợp các kết cục thuộc vào A hoặc thuộc vào B, hoặc thuộc cả 2. Biến cố $A \cup B$ xảy ra nên hoặc A hoặc B xảy ra hoặc cả A và B xảy ra.

Phép lấy giao hai tập hợp A và B được ký hiệu bởi $A \cap B$ và được xác định như là tập hợp của các kết cục thuộc vào cả A và B. Biến cố $A \cap B$ xảy ra khi cả hai biến cố A và B xảy ra. Hai biến cố được gọi là **xung khắc** nếu giao của chúng là biến cố không, $A \cap B = \varnothing$. Các biến cố xung khắc không thể xảy ra đồng thời.

Phép lấy phần bù của một biến cố A được ký hiệu A^c và được xác định như là tập hợp của tất cả các kết cục không thuộc A. Biến cố A^c xảy ra khi biến cố A không xảy ra và ngược lại.

Các Hình 2.2(a), 2.2(b) và 2.2(c) chỉ ra các phép toán tập hợp cơ sở khi dùng sơ đồ Venn. Trong các sơ đồ này hình chữ nhật thể hiện không gian mẫu S, và các miền được gạch chéo biểu diễn các biến cố khác nhau. Hình 2.2(d) chỉ ra hai biến cố xung khắc.

Nếu một biến cố A là tập con của biến cố B, nghĩa là A = B, khi đó biến cố B sẽ xảy ra khi biến cố A xảy ra bởi vì tất cả các kết cục thuộc A cũng thuộc B (Xem Hình 2.2(d)). Bởi lý do này, chúng ta nói rằng, biến cố A **kéo theo** biến cố B. Rõ ràng rằng, từ các sơ đồ trong các hình 2.2(a) và 2.2(b) suy ra $A \cap B$ kéo theo cả B và mỗi biến cố A và B kéo theo $A \cup B$.

Các biến cố A và B là **bằng nhau**, A = B, nếu chúng gồm các kết cục như nhau.

Ba phép toán tập hợp trên có thể kết hợp để tạo thành các biến cố khác. Các tính chất sau đây của các phép toán tập hợp và tổ hợp của chúng là có ích:

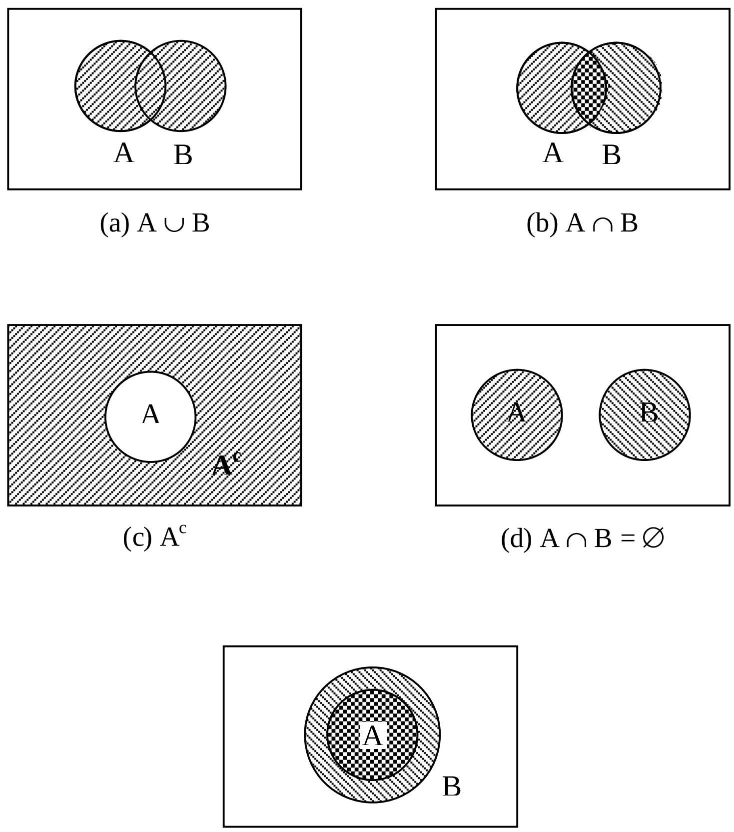
Tính chất giao hoán :

$A \cup B = B \cup A$ và $A \cap B = B \cap A$. (2.1)

Tính chất kết hợp :

$A \cup (B \cup C) = (A \cup B) \cup C$ và $A \cap (B \cap C) = (A \cap B) \cap C$ (2.2)

HÌNH 2.2
Các phép toán tập hợp và các quan hệ tập hợp



(a) $A \cup B$

(b) $A \cap B$

(c) A^c

(d) $A \cap B = \varnothing$

(e) $A \subset B$

Tính chất phân phối:

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ và

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. (2.3)

Qui tắc De Morgan:

$(A \cap B)^c = A^c \cup B^c$ và $(A \cup B)^c = A^c \cap B^c$ (2.4)

VÍ DỤ 2.4 Với thí nghiệm E_{10}, giả sử các biến cố A, B và C được xác định như sau:

$A = \{v : |v| > 10\}$, "Giá trị tuyệt đối của v lớn hơn 10 vôn,"

$B = \{v : v < -5\}$, "v nhỏ hơn –5 vôn," và

$C = \{v : v > 0\}$, "v là dương".

Các bạn có thể kiểm tra lại

$A \cup B = \{v : v < -5 \text{ hoặc } v > 10\}$,

$A \cap B = \{v : v < -10\}$,

$C^c = \{v : v \leq 0\}$,

$(A \cup B) \cap C = \{v : v > 10\}$,

$A \cup B \cup C = \varnothing$, và

$(A \cup B)^c = \{v : -5 \leq v \leq 10\}$.

VÍ DỤ 2.5 Hình 2.3 biểu diễn 3 hệ thống gồm 3 thành phần C_1, C_2 và C_3. Hình 2.3(a) là hệ nối tiếp mà ở đó hệ làm việc chỉ khi 3 thành phần cùng mắc cùng làm việc. Hình 2.3(b) là hệ mắc song song mà ở đó hệ làm việc cho đến khi còn ít nhất một trong ba thành phần làm việc. Hình 2.3(c) là hệ "hai-từ-ba" mà ở đó hệ làm việc cho đến khi có ít nhất hai trong ba thành phần làm việc. Giả sử A_k là biến cố "thành phần thứ *k* làm việc". Với mỗi cách bố trí của ba hệ thống có thể biểu diễn biến cố "hệ đang làm việc" theo thuật ngữ của các biến cố A_k.

Hệ mắc nối tiếp làm việc nếu và chỉ nếu tất cả các thành phần làm việc/ Như vậy biến cố D_a , "hệ a làm việc" được đưa ra bởi :

$$D_a = A_1 \cap A_2 \cap A_3.$$

Hệ mắc song song làm việc cho đến khi còn ít nhất một thành phần làm việc, nghĩa là, nếu thành phần 1 hoặc thành phần 2 hoặc thành phần 3 hoặc một tổ hợp bất kỳ của chúng còn làm việc. do vậy biến cố D_b, " Hệ mắc song song b làm việc", được đưa ra bởi

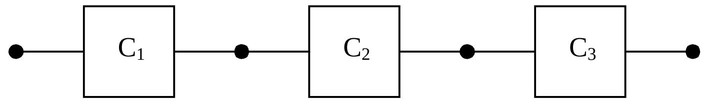
$$D_b = A_1 \cup A_2 \cup A_3.$$

Cuối cùng, hệ hai-từ-ba làm việc cho đến khi có không quá 1 thành phần không làm việc. Do vậy, biến cố D_c, " Hệ hai-từ-ba làm việc", được đưa ra bởi

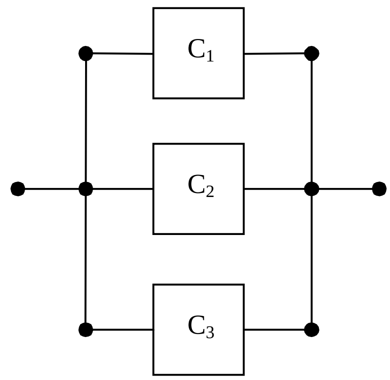
$$D_c = (A_1 \cap A_2 \cap A_3) \cup (A^c_1 \cap A_2 \cap A_3) \cup (A_1 \cap A^c_2 \cap A_3) \cup (A_1 \cap A_2 \cap A^c_3).$$

HÌNH 2.3

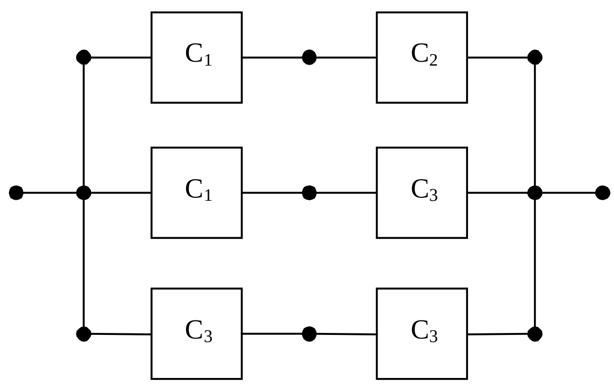
Dạng thức của một hệ ba thành phần. Khi một thành phần bị hỏng nó bị bỏ đi khỏi sơ đồ. Hệ thống làm việc khi có một đường đi từ bên phải sang bên trái.



(a) Hệ mắc nối tiếp



(b) Hệ mắc song song



(c) Hệ hai-từ-ba

Phép toán hợp và giao có thể được lặp lại với một số tuỳ ý các biến cố. Do vậy biến cố

$$\bigcup_{k=1}^{n} A_k = A_1 \cup A_2 \cup \ldots \cup A_n$$

xảy ra một hoặc hơn một biến cố từ các A_k xảy ra. Biến cố

$$\bigcap_{k=1}^{n} A_K = A_1 \cap A_2 \cap \ldots \cap A_n$$

xảy ra khi tất cả các biến cố $A_1, \ldots, A_n$ xảy ra. Các phép toán cũng có thể được sử dụng với dãy vô hạn đếm được các biến cố. Khi đó chúng ta cũng có các biến cố dạng

$$\bigcup_{k=1}^{\infty} A_k \quad \text{và} \quad \bigcap_{k=1}^{\infty} A_k .$$

2.2 CÁC TIÊN ĐỀ XÁC SUẤT

Các xác suất là các số được gán cho các biến cố để chỉ ra khả năng xảy ra các biến cố này khi thí nghiệm được thực hiện. **Luật xác suất** cho một thí nghiệm ngẫu nhiên là luật gán các xác suất cho các biến cố của thí nghiệm này. Do vậy một luật xác suất là một hàm mà nó gán một số với một tập hợp (biến cố). Trong phần 1.3 chúng ta đã tìm được một số tính chất của dãy tần suất mà một định nghĩa bất kỳ của xác suất có thể thoả mãn. Các tiên đề xác suất chỉ ra rằng các luật xác suất cần phải thoả mãn các tính chất này. Trong phần này, chúng ta phát triển một số các kết quả được suy ra từ các tiên đề này.

Giả sử E là một thí nghiệm ngẫu nhiên với không gian mẫu S. Một luật xác suất cho thí nghiệm E là một luật gán cho mỗi biến cố A một số P[A], được gọi là xác suất của A, mà nó thoả mãn các tiên đề sau:

Tiên đề I $0 \le P[A]$.

Tiên đề II $P[S] = 1$.

Tiên đề III Nếu $A \cap B = \varnothing$, thì $P[A \cup B] = P[A] + P[B]$.

Tiên đề III' Nếu $A_1, A_2, \ldots$ là dãy các biến cố thoả mãn

$A_i \cap A_j = \varnothing \ \forall\, i \neq j$, thì khi đó

$$P\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} P[A_k].$$

Các tiên đề I, II, và III đủ để giải quyết các thí nghiệm có không gian mẫu hữu hạn. Để sử dụng với các thí nghiệm có không gian mẫu vô hạn, tiên đề III cần phải được thay bởi tiên đề III'. Chú ý rằng, khi ta giả thiết $A_k = \varnothing$ với $k \ge 3$. do vậy chúng ta thực sự chỉ cần các tiên đề I, II, và III.Tuy vậy chúng ta sẽ thu được sự hiểu biết sâu sắc bởi sự bắt đầu với các tiên đề I, II, và III.

Các tiên đề cho phép chúng ta nhìn các biến cố như là các đối tượng có tính chất (tức là, xác suất của chúng) mà nó có thuộc tính đơn giản như là trọng lượng vật lý. Tiên đề I chỉ ra rằng, xác suất (trọng lượng) là không âm, tiên đề II chỉ ra rằng tổng các xác suất (trọng lượng) là một số cố định mà ta đặt là một đơn vị. Tiên đề III chỉ ra rằng xác suất của tổng hai biến cố không đồng thời bằng tổng của các xác suất thành phần.

Các tiên đề cung cấp cho chúng ta một tập hợp các qui luật bắt buộc mà một phép gán giá trị xác suất bất kỳ phải thoả mãn. Bây giờ chúng ta sẽ phát triển một số tính chất nảy sinh từ các tiên đề này mà nó có ích khi tính các xác suất.

Kết quả đầu tiên chỉ ra rằng, nếu chúng ta phân không gian mẫu thành hai biến cố xung khắc, A và A^c, thì xác suất của hai biến cố cộng lại bằng 1.

HỆ QUẢ 1. $P[A^c] = 1 - P[A]$

Chứng minh: Từ biến cố A và A^c xung khắc nhau $A \cap A^c = \emptyset$, chúng ta nhận được từ tiên đề III rằng

$$P[A \cup A^c] = P[A] + P[A^c].$$

Do $S = A \cup A^c$, tiên đề II,

$$1 = P[S] = P[A \cup A^c] = P[A] + P[A^c].$$

Hệ quả được chứng minh sau khi tìm ra $P[A^c]$.

Hệ quả tiếp theo chỉ ra rằng xác suất của một biến cố luôn luôn nhỏ hơn hoặc bằng 1. Hệ quả II kết hợp với tiên đề I cho chúng ta phép kiểm tra hết khi giải một bài toán: Nếu xác suất tìm ra của bạn là số âm hoặc lớn hơn 1, khi đó bạn đã mắc phải một sai lầm ở đâu đó.

HỆ QUẢ 2. $P[A] \le 1$.

Chứng minh: từ hệ quả 1,

$$P[A] = 1 - P[A^c] \le 1$$

do $P[A^c] \ge 0$.

Hệ quả 3 chỉ ra rằng, biến cố không thể có xác suất 0.

HỆ QUẢ 3. $P[\emptyset] = 0$

Chứng minh: Giả sử $A = S$ và $A^c = \emptyset$ trong hệ quả 1:

$$P[\emptyset] = 1 - P[S] = 0.$$

Hệ quả 4 cung cấp cho chúng ta một phương pháp chuẩn để tính xác suất của một biến cố hợp A. Phương pháp dẫn đến việc phân tích biến cố A thành hợp của các biến cố rời nhau $A_1, A_2, \ldots, A_n$. Xác suất của A bằng tổng các xác suất của các biến cố tính chất.

HỆ QUẢ 4. Nếu $A_1, A_2, \ldots, A_n$ là xung khắc với nhau từng đôi, khi đó :

$$P\left[\bigcup_{k=1}^{n} A_K\right] = \sum_{k=1}^{n} P[A_k] \qquad \text{với } n \ge 2.$$

Chứng minh : Chúng ta sử dụng phương pháp quy nạp toán học. Tiên đề III chỉ ra rằng, kết quả là đúng với $n = 2$. Tiếp theo chúng ta chỉ ra rằng nếu kết quả là đúng với n nào đó, thì nó cũng đúng với $n + 1$. Điều này, kết hợp với kết luận rằng kết quả là đúng với $n = 2$, chỉ ra rằng kết quả đúng với $n \ge 2$.

Giả sử rằng kết quả đúng với $n > 2$ nào đó, nghĩa là :

$$P\left[\bigcup_{k=1}^{n} A_K\right] = \sum_{k=1}^{n} P[A_k], \tag{2.5}$$

và xét từng trường hợp $n + 1$

$$P\left[\bigcup_{n=k}^{n+1} A_k\right] = P\left[\left\{\bigcup_{k=1}^{n} A_K\right\} \cup A_{n+1}\right] = P\left[\bigcup_{k=1}^{n} A_k\right] + P[A_{n+1}], \qquad (2.6)$$

ở đây chúng ta sử dụng tiên đề III với biểu thức thứ hai sau khi chú ý rằng hợp của các biến cố từ A_1 đến A_n là xung khắc với A_{n+1}. Khi đó chúng sử dụng tính chất phân phối

$$\left\{\bigcup_{k=1}^{n} A_k\right\} \cap A_{n+1} = \bigcup_{k=1}^{n} \{A_k \cap A_{n+1}\} = \bigcup_{k=1}^{n} \varnothing = \varnothing.$$

Thế đẳng thức (2.5) vào đẳng thức (2.6) cho trường hợp n + 1

$$P\left[\bigcup_{k=1}^{n+1} A_k\right] = \sum_{k=1}^{n+1} P[A_k].$$

Hệ quả 5 cho một biểu diễn của hợp hai biến cố không nhất thiết là phải xung khắc.

HỆ QUẢ 5. $P[A \cup B] = P[A] + P[B] - P[A \cap B]$

Chứng minh: Trước hết chúng ta phân tích $A \cup B$, A và B như là hợp của các biến cố rời nhau. từ sơ đồ Venn trong hình 2.4,

$P[A \cup B] = P[A \cap B^c] + P[B \cap A^c] + P[A \cap B]$

$P[A] = P[A \cap B^c] + P[A \cap B]$

$P[B] = P[B \cap A^c] + P[A \cap B]$

Bởi việc thế $P[A \cap B^c]$ và $P[B \cap A^c]$ từ hai đẳng thức sau vào đẳng thức đầu, chúng ta nhận được hệ quả.

Với việc nhìn vào sơ đồ Venn 2.4, bạn sẽ nhận thấy rằng tổng P[A] + P[B] tính xác suất của tập hợp $A \cap B$ hai lần. Biểu thức trong hệ quả 5 làm một sự hiệu chỉnh thích hợp.

Hệ quả 5 dễ dàng tổng quát hoá cho trường hợp 3 biến cố:

$$P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - P[B \cap C] + P[A \cap B \cap C], \qquad (2.7)$$

và tổng quát hoá cho trường hợp n biến cố được chỉ ra trong hệ quả 6.

HỆ QUẢ 6.

$$P\left[\bigcup_{k=1}^{n} A_k\right] = \sum_{j=1}^{n} P[A_j] - \sum_{j<k} P[A_j \cap A_k] + \ldots + (-1)^{n+1} P[A_1 \cap \ldots \cap A_n].$$

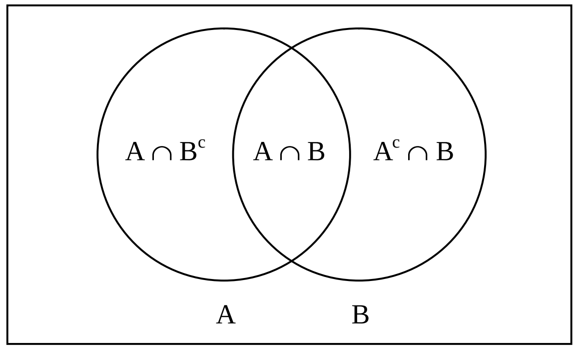
Chứng minh: Bằng phương pháp quy nạp (xem các Bài tập 18 và 19).

Do các xác suất là không âm, hệ quả 5 kéo theo rằng xác suất của hợp hai biến cố không lớn hơn bản tổng xác suất của các biến cố thành phần.

$$P[A \cup B] \leq P[A] + P[B]. \qquad (2.8)$$

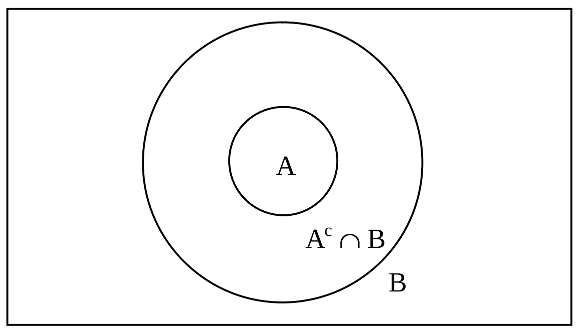
HÌNH 2.4

Phân tích $A \cup B$ thành 3 tập hợp rời nhau.



HÌNH 2.5

Nếu $A \subset B$, thì $P(A) \leq P(B)$.



Bất đẳng thức trên là trường hợp riêng của kết luận rằng, một tập hợp con của một tập hợp cần phải có xác suất nhỏ hơn. Kết quả này thường được dùng để nhận được cận trên của xác suất mà ta quan tâm. Trường hợp điển hình là khi ta cần tính xác suất của biến cố A, nhưng xác suất của biến cố này khó tính, khi đó ta tính xác suất của biến cố B chứa biến cố A như một tập hợp con.

HỆ QUẢ 7. Nếu $A \subset B$, thì $P[A] \leq P[B]$.

Chứng minh: Trong hình 2. 5, B là hợp của A và $A^c \cap B$, do vậy

$$P[B] = P[A] + P[A^c \cap B] \geq P[A],$$

do $P[A^c \cap B] \geq 0$.

Các tiên đề cùng với các hệ quả cung cấp cho chúng ta các quy tắc để tính xác suất của một biên cố đã cho dựa vào các biến cố khác. Tuy vậy, chúng ta vẫn cần **phép gán xác suất ban đầu** (**initial probability assignment**) cho tập các biến cố cơ bản nào đó, để dựa vào đó có thể tính xác suất của tất cả các biến cố khác. Bài toán này được giải quyết trong hai mục tiếp sau.

Không gian Mẫu Rời rạc

Trong phần này chúng ta sẽ chỉ ra rằng *luật xác suất của thí nghiệm với không gian mẫu đếm được có thể được xác định bởi việc đưa vào xác suất của các biến cố sơ cấp*. Trước hết, giả thiết rằng không gian mẫu là hữu hạn, S{a_1, a_2, …, a_n}. tất cả các biến cố sơ cấp khác nhau đều là xung khắc, do hệ quả 4, xác suất của biến cố bất kỳ B = $\{a_1^{'}, a_2^{'}, ..., a_m^{'}\}$ được cho bởi công thức

$$P[B] = P[\{a_1^{'}, a_2^{'}, ..., a_m^{'}\}]$$
$$= P[\{a_1^{'}\}] + P[\{a_2^{'}\}] + P[\{a_m^{'}\}]; \qquad (2.9)$$

nghĩa là xác suất của một biến cố bằng tổng của các xác suất của các kết cục thuộc vào biến cố. Nếu S là vô hạn đếm được, khi đó tiên đề III' chỉ ra rằng xác suất của biến cố D = { $b_1^{'}, b_2^{'}, ...$ } được cho bởi công thức

$$P[D] = P[\{b_1^{'}\}] + P[\{b_2^{'}\}] + ... \; . \qquad (2.10)$$

Hơn nữa, xác suất của một biến cố được xác định từ xác suất của các kết cục của nó. Do vậy, chúng ta kết luận rằng, luật xác suất của một thí nghiệm ngẫu nhiên với không gian mẫu rời rạc được xác định bởi xác suất của các biến cố sơ cấp.

Nếu không gian mẫu có *n* phương trình, S = {a_1, …, a_n}, phép gán xác suất đều nhau là trường hợp **các kết cục đồng khả năng** (**equally likely outcomes**). Khi đó xác suất của các biến cố sơ cấp bằng:

$$P[\{a_1\}] = P[\{a_2\}] = \mathrm{L} = P[\{a_n\}] = \frac{1}{n}. \qquad (2.11)$$

Xác suất của biến cố bất kỳ chứa k kết cục, $B = \{a_1^{'}, ..., a_k^{'}\}$, bằng

$$P[B] = P[\{a_1^{'}\}] + \mathrm{L} + P[\{a_k^{'}\}] = \frac{k}{n}. \qquad (2.12)$$

Như vậy, *nếu các kết cục đồng khả năng, khi đó xác suất của một biến cố bằng số các kết cục phụ thuộc vào biến cố chia cho tổng số các kết cục của không gian mẫu*. Phần 2.3 thảo luận các phương pháp tính thường được để tìm xác suất trong thí nghiệm có các kết cục đồng thời khả năng.

VÍ DỤ 2.6 Một hộp chứa 10 viên bi được đánh số 0, 1, …, 9. Phép thử ngẫu nhiên là phép lấy 1 viên bi từ hộp và ghi lại số viên bi. Hãy tìm xác suất của các biến cố sau:

A = "Số của viên bi là số lẻ",

B = "Số của viên bi là tích của 3",

C = "Số của viên bi được lấy nhỏ hơn 5",

và xác suất của các biến cố A ∪ B và A ∪ B ∪ C.

Không gian mẫu là S = {0, 1, …, 9}, và tập hợp của các kết cục tương ứng với các biến cố trên là :

A = {1, 3, 5, 7, 9}, B = {3, 6, 9}, và C = {0, 1, 2, 3, 4}.

Nếu chúng ta giả sử rằng, các kết cục đồng khả năng, khi đó

$$P[A] = P[\{1\}] + P[\{3\}] + P[\{5\}] + P[\{7\}] + P[\{9\}] = \frac{5}{10}.$$

$$P[B] = P[\{3\}] + P[\{6\}] + P[\{9\}] = \frac{3}{10}.$$

$$P[C] = P[\{1\}] + P[\{2\}] + P[\{3\}] + P[\{4\}] = \frac{5}{10}.$$

Từ Hệ quả 5,

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] = \frac{5}{10} + \frac{3}{10} - \frac{2}{10} = \frac{6}{10},$$

ở đây ta đã sử dụng thực tế là $A \cap B = \{3, 9\}$, do vậy $P[A \cap B] = 2/10$.

Từ Hệ quả 6,

$$\begin{aligned} P[A \cup B \cup C] &= P[A] + P[B] + P[C] - P[A \cap B] \\ &\quad - P[A \cap C] - P[B \cap C] + P[A \cap B \cap C] \\ &= \frac{5}{10} + \frac{3}{10} + \frac{5}{10} - \frac{2}{10} - \frac{2}{10} - \frac{1}{10} + \frac{1}{10} \\ &= \frac{9}{10}. \end{aligned}$$

Bạn có thể kiểm tra lại các xác suất $P[A \cup B]$ và $[A \cup B \cup C]$ bằng cách tính ra số các kết cục của biến cố này.

Nhiều mô hình xác suất được phát minh cho không gian mẫu và các biến cố bằng cách thay đổi phép gán xác suất; trong trường hợp không gian mẫu hữu hạn chúng ta cần làm sao để tổng xác suất của tất cả các biến cố sơ cấp bằng 1. Tất nhiên trong tình huống cụ thể bất kỳ, phép gán xác suất có thể được chọn dưới sự ảnh hưởng của các quan trắc thí nghiệm tới giá trị xác suất. Ví dụ sau chỉ ra rằng, một tình huống có thể xuất hiện khi có hơn một phép gán xác suất "hợp lý" và khi kết quả thí nghiệm đòi hỏi lựa chọn một phép gán thích hợp.

VÍ DỤ 2.7 Giả sử rằng một đồng xu được tung 3 lần, nếu chúng ta quan sát dãy xuất hiện mặt ngửa và mặt sấp, thì có 8 kết cục có thể S_3 = {HHH, HHT, HTH, THH, TTH, THT, HTT, TTT}. Nếu chúng ta giả sử rằng các kết cục của S_3 là đồng khả năng, khi đó xác suất của mỗi một biến cố sơ cấp bằng 1/8. Phép gán xác suất này kéo theo xác suất nhận được 2 lần xuất hiện mặt ngửa trong 3 lần tung ra, do Hệ quả 3,

P["2 lần xuất hiện mặt ngửa trong 3 lần tung"] =

= P[{HHT, HTH, THH}]

$$= P[\{ HHT\}] + P[\{HTH\}] + P[\{THH\}] = \frac{3}{8}.$$

Bây giờ giả sử rằng chúng ta tung đồng xu 3 lần, nhưng chúng ta chỉ đếm số lần xuất hiện mặt ngửa trong 3 lần tung thay vì quan sát dãy mặt ngửa và mặt sấp. Khi đó không gian mẫu là $S_4 = \{0, 1, 2, 3\}$. Nếu chúng ta giả thiết các kết cục của S_4 là xác suất, thì mỗi một biến cố sơ cấp của S_4 có xác suất = 1/4 . phép gán xác suất thứ hai này dẫn đến xác suất nhận được 2 lần xuất hiện mặt sấp trong 3 lần tung là:

P["2 lần xuất hiện mặt ngửa trong 3 lần tung"] = P[{2}] = $\frac{1}{4}$.

Phép gán xác suất thứ nhất kéo theo xác suất xuất hiện mặt sấp 2 lần trong 3 lần tung bằng 3/8, và phép gán xác suất thứ 2 dẫn tới xác suất của biến cố này bằng 1/4. Như vậy hai phép gán là không nhất quán với nhau. Về mặt lý thuyết chúng ta quan tâm đến một phép gán khác là chấp nhận được. Điều đó vẫn tiếp diễn cho đến khi chúng ta chọn được một phép gán hợp lý hơn. Trong phần sau của chương chúng ta sẽ thấy rằng chỉ phép gán thứ nhất là phù hợp giả thiết rằng đồng xu là cân đối và các lần tung là độc lập. Phép gán này phù hợp với tần suất tương đối mà đã được quan sát trong thí nghiệm tung đồng xu thực tế.

Cuối cùng chúng ta xét ví dụ với không gian mẫu vô hạn đếm được.

VÍ DỤ 2.8 Một đồng xu cân đối được tung lặp lại cho đến khi lần đầu tiên xuất hiện mặt ngửa; kết cục của thí nghiệm là số lần tung cho đến khi mặt ngửa lần đầu tiên xuất hiện. Tìm luật xác suất cho thí nghiệm này.

Có thể xảy ra số lần tung lớn tuỳ ý cho đến khi xuất hiện mặt ngửa, do vậy không gian mẫu là S = {1, 2, 3, …}. Giả sử rằng thí nghiệm được lặp lại *n* lần. Đặt N_j là số các phép thử mà ở lần tung xu thứ *j* mặt ngửa lần đầu tiên xuất hiện. Nếu *n* là số rất lớn chúng ta hy vọng N_1 xấp xỉ $n/2$ do đồng xu là cân đối. Điều này kéo theo kết quả ở lần tung thứ hai cần khoảng $n - N_1 \approx n/2$ lần, và một lần nữa ta hy vọng rằng khoảng một nửa của số này, nghĩa là $n/4$, sẽ có kết quả là mặt ngửa, và vân vân, như được chỉ ra trong Hình 2.6. Như vậy với *n* lớn, tần suất tương đối là:

$$f_j \approx \frac{N_j}{n}\left(\frac{1}{2}\right)^j \qquad j = 1, 2, \ldots .$$

Do đó chúng ta kết luận rằng luật xác suất hợp lý cho thí nghiệm này là

$$P[j \text{ lần tung xu để mặt ngửa lần đầu tiên xuất hiện}] = \left(\frac{1}{2}\right)^j$$

$$j = 1, 2, \ldots \qquad (2.13)$$

Chúng ta có thể kiểm tra lại rằng các xác suất này cộng lại bằng 1 bằng cách dùng cấp số nhân với công bội α = 1/2:

$$\sum_{j=1}^{\infty} \alpha^j = \left.\frac{\alpha}{1-\alpha}\right|_{\alpha=1/2} = 1.$$

Không gian Mẫu Liên tục

Các không gian mẫu liên tục xuất hiện trong các thí nhiệm mà các kết cục là các số, các biến cố mà ta quan tâm trong các thí nghiệm này là các khoảng của đường thẳng thực hoặc các miền hình chữ nhật trong mặt phẳng và các phép lấy phần bù, hợp và giao của các biến cố này, với các không gian mẫu liên tục là qui tắc gán các số cho các khoảng của đường thẳng thực hoặc các miền hình chữ nhật trong mặt phẳng.

VÍ DỤ 2.9 Xét thí nghiệm ngẫu nhiên "lấy ngẫu nhiên một số *x* nằm giữa 0 và 1." Không gian mẫu S cho thí nghiệm này là đoạn đơn vị [0, 1], không có số điểm vô hạn không đếm được. Nếu chúng ta giả sử rằng tất cả kết các cục của S được lấy đồng khả năng, thì chúng ta có thể dự đoán rằng xác suất để kết cục thuộc vào khoảng [0; 1/2] bằng với xác suất để kết cục thuộc khoảng [1/2; 1]. Chúng ta cũng có thể dự đoán rằng xác suất để xảy ra giá trị đúng bằng 1/2 có thể bằng 0 do có một số vô hạn không đếm được các biến cố đồng khả năng.

Xét luật xác suất sau: "Xác suất để kết cục thuộc vào một tập con của S là bằng độ dài của tập con đó," nghĩa là,

$$P[[a, b]] = (b - a) \qquad \text{với } 0 \le a \le b \le 1, \qquad (2.14)$$

ở đây chúng ta ký hiệu P[[*a*, *b*]] có nghĩa là xác suất của biến cố tương ứng với khoảng [*a*, *b*]. Rõ ràng rằng, Tiên đề I thoả mãn do $b \ge a \ge 0$. Tiên đề II được suy ra từ S = [*a*, *b*] với *a* = 0 và *b* = 1.

Bây giờ chúng ta chỉ ra rằng luật xác suất là phù hợp với dự đoán ở trên về các xác suất của các biến cố [0, 1/2], [1/2, 1], và {1/2}:

$$P[[0, 0.5]] = 0.5 - 0 = .5$$

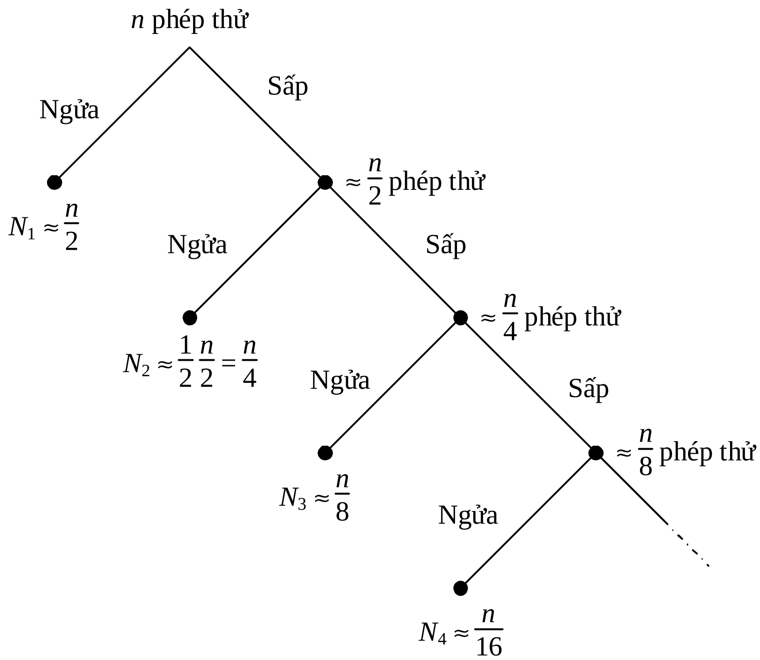
$$P[[0.5, 0]] = 1 - 0.5 = .5$$

Hơn nữa, nếu x_0 là điểm bất kỳ của S, khi đó $P[[x_0, x_0]] = 0$ do các điểm riêng biệt có độ dài bằng 0.

Bây giờ giả sử rằng, chúng ta quan tâm đến một biến cố là hợp của một vài khoảng ví dụ "kết cục cách tâm của đoạn thẳng đơn vị tối thiểu 0,3 ", nghĩa là, A = [0, 0.2] ∪ [0,8, 1]. Do hai khoảng là không giao nhau, chúng ta có kết quả từ Tiên đề III

$$P[A] = P[[0, 0.2]] + P[[0.8, 1]] = .4.$$

HÌNH 2.6 Trong n phép thử, mặt ngửa xuất hiện ở lần gieo đầu tiên trong xấp xỉ $n/2$ lần gieo, ở lần gieo thứ hai trong xấp xỉ $n/4$ lần gieo, và vân vân.



Ví dụ tiếp theo chỉ ra rằng phép gán xác suất ban đầu mà xác định rõ xác suất của các khoảng nửa vô hạn cũng đủ để xác định xác suất của tất cả các biến cố mà ta quan tâm.

VÍ DỤ 2.10 Giả sử rằng thời gian sống của một chip nhớ máy tính được đo, và chúng ta tìm được rằng "Tỉ lệ thời gian sống của chip vượt qua t là hàm mũ âm với tốc độ α ". Hãy tìm một luật xác suất xấp xỉ.

Lấy không gian mẫu trong thí nghiệm này là $S = (0, \infty)$. Nếu chúng ta làm sáng tỏ khái niệm đưa ra ở trên như là "xác suất để thời gian sống của chip vượt qua t là hàm mũ âm với tốc độ α", khi đó chúng ta nhận được phép gán xác suất sau cho các biến cố có dạng (t, ∞):

$$P[(t, \infty)] = e^{-\alpha t} \qquad \text{với } t > 0. \tag{2.15}$$

ở đây $\alpha > 0$. Chú ý rằng giá trị hàm mũ này là một số nằm giữa 0 và 1 với $t > 0$, bởi vậy Tiên đề I được thoả mãn. Tiên đề II được thoả mãn do:

$$P[S] = P[(0, \infty)] = 1.$$

Còn xác suất để thời gian sống nằm trong khoảng $(r, s]$ được tìm bởi chú ý trong Hình 2.7 rằng $(r, s] \cup (s, C) = (r, \infty)$, bởi vì Tiên đề III,

$$P[(r, \infty)] = P[(r, s]] + P[(s, \infty)].$$

Chuyển vế đẳng thức trên chúng ta nhận được

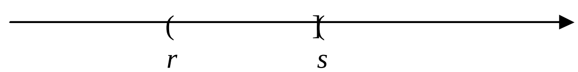
$$P[(r, s]] = P[(r, \infty)] - P[(s, \infty)] = e^{-\alpha r} - e^{-\alpha s}.$$

Như vậy chúng ta nhận được xác suất của các khoảng trong S.

Trong cả hai Ví dụ 2.9 và 2.10 đều có kết cục lấy xác suất bằng 0. Bạn có thể hỏi: Liệu một kết cục (hoặc biến cố) có xác suất 0, không có nghĩa là nó không xảy ra hay không? Và bạn cũng có thể hỏi: Trong một không gian mẫu có thể có bao nhiêu kết cục có xác suất 0? Chúng ta có thể giải thích nghịch lý này bằng việc dùng tần suất tương đối để giải thích xác suất. Một biến cố xảy ra chỉ một lần trong số vô hạn lần các phép thử sẽ có tần suất tương đối bằng 0. Do thực tế này một biến cố hoặc kết cục có tần suất tương đối bằng 0 không thể suy ra được rằng nó không thể xảy ra, nhưng có thể nói rằng nó rất hiếm khi xảy ra . Trong trường hợp không gian mẫu liên tục, tập tất cả các kết cục có thể nhiều đến mức tất cả các kết cục hiếm xảy ra có tần suất tương đối đúng bằng 0.

Chúng ta kết thúc phần này bằng một ví dụ mà ở đó các kết cục là các miền của mặt phẳng.

HÌNH 2.7
$(t,\infty)=(t,s]\cup(s,\infty)$

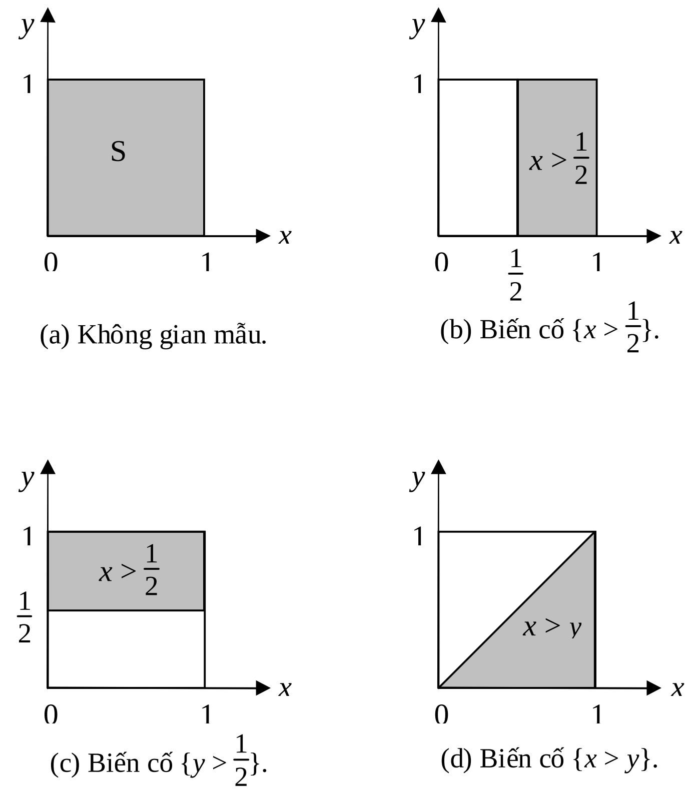


VÍ DỤ 2.11 Xét thí nghiệm E_{12}, mà ở đó chúng ta lấy một cách ngẫu nhiên 2 số x và y nằm giữa 0 và 1. Khi đó không gian mẫu là hình vuông đơn vị được chỉ ra ở Hình 2.8(a). Nếu chúng ta giả sử rằng tất cả các cặp số trong hình vuông đơn vị được lấy đồng khả năng, khi đó là hợp lý với phép gán xác suất mà ở đó xác suất của miền R bất kỳ của hình vuông đơn vị bằng diện tích của nó. Bây giờ ta tìm xác suất của các biến cố: A = $\{x > 0.5\}$, B = $\{y > 0.5\}$, và C = $\{x > y\}$.

Các Hình từ 2.8(b) tới 2.8(d) biểu diễn các miền tương ứng với các biến cố A, B và C. Rõ ràng mỗi một miền này có diện tích bằng 1/2 . Do đó:

$$P[A] = \frac{1}{2}, \qquad P[B] = \frac{1}{2}, \qquad P[C] = \frac{1}{2}.$$

HÌNH 2.8
Một không gian mẫu hai chiều và ba biến cố



(a) Không gian mẫu.

(b) Biến cố $\{x > \frac{1}{2}\}$.

(c) Biến cố $\{y > \frac{1}{2}\}$.

(d) Biến cố $\{x > y\}$.

Chúng ta lặp lại như thế nào trong phần sau từ việc thiết lập bài toán đến mô hình xác suất của nó. Phép đặt bài toán xác định ẩn hay hiện một thí nghiệm ngẫu nhiên, mà nó chỉ rõ kết quả thí nghiệm và tập hợp các phép đo và quan trắc. các phép đo và quan trắc quyết định tập tất cả các kết cục có thể và do đó không gian mẫu S.

Một phép gán xác suất ban đầu chỉ rõ xác suất của các biến cố có thể cần phải được xác định tiếp ngay sau đó. Phép gán xác suất này cần phải thoả mãn các tiên đề xác suất. Nếu S là rời rạc, khi đó nó đủ để xác định xác suất của một biến cố sơ cấp. Nếu S là liên tục, nó đủ để xác định xác suất của các khoảng của đường thẳng thực hoặc các miền của mặt phẳng. Xác suất của các biến cố khác mà ta quan tâm có thể được xác định từ phép gán xác suất ban đầu và các tiên đề xác suất và các hệ quả của nó. Nhiều phép gán xác suất có thể, song cần phải chọn phép gán xác suất cần phải liên quan đến các quan trắc thí nghiệm và/ hoặc thí nghiệm trước đây.

*2.3 TÍNH XÁC SUẤT BẰNG PHƯƠNG PHÁP ĐẾM[2]

Trong nhiều thí nghiệm với không gian mẫu hữu hạn, các kết cục có thể giả thiết đồng xác suất. Xác suất của một biến cố là tỉ số của số các kết cục thuộc vào biến cố mà ta quan tâm với tổng số các kết cục của không gian mẫu (Đẳng thức 2.12). Phép tính xác suất qui về việc đếm số các kết cục thuộc vào biến cố. Trong phần này chúng ta phát triển một vài công thức (tổ hợp đếm hay dừng).

Giả sử rằng một bài kiểm tra trắc nghiệm có k câu hỏi và câu thứ i sinh viên cần phải lựa chọn 1 trong n_i câu trả lời có thể. Hỏi rằng, tổng số tất cả các cách trả lời bài kiểm tra này bằng bao nhiêu? Câu trả lời của câu hỏi thứ i có thể được coi như thành phần thứ i của một véctơ k-chiều, do đó câu hỏi trên tương đương với: có bao nhiêu bộ-k sắp thứ tự khác nhau $(x_1, \ldots, x_k)$ trong đó x_i là một phần tử của một tập hợp n_i phần tử khác nhau?

Chúng ta xét trường hợp $k = 2$. Nếu như chúng ta sắp xếp tất cả các phép chọn có thể của x_1 và x_2 dọc theo các cạnh của bảng được chỉ ra trong Hình 2.9, chúng ta nhận thấy rằng có n_1n_2 cặp sắp thứ tự khác. Với bộ 3 chúng ta có thể sắp xếp n_1n_2 cặp (x_1, x_2) dọc theo cạnh thẳng đứng của bảng và n_3 cách chọn x_3 dọc theo cạnh nằm ngang. Rõ ràng rằng, số các bộ 3 có thể là $n_1n_2n_3$.

Một cách tổng quát, *số của các bộ k sắp thứ tự khác nhau* $(x_1, \ldots, x_k)$ *với thành phần* x_i *từ tập có* n_i *phần tử khác là*

$$\text{số các bộ } k \text{ sắp thứ tự khác nhau} = n_1n_2\ldots n_k. \qquad (2.16)$$

(2). Phần này và tất cả các phần được đánh dấu hoa thị có thể bị lược qua mà không mất tính liên tục.

Nhiều bài toán đếm có thể đặt như bài toán lấy mẫu mà ở đó chúng ta lấy "những viên bi" từ "những chiếc hộp" hoặc "những đối tượng" từ "những đám đông." Bây giờ chúng ta sẽ dùng Đẳng thức (2.16) để phát triển bài toán tổ hợp cho dạng đã biết của bài toán lấy mẫu.

HÌNH 2.9

Nếu có n_1 cách chọn khác nhau của x_1 và n_2, khi đó n_1n_2 cách chọn cặp thứ tự khác nhau (x_1, x_2).

$$
\begin{array}{ccc|cccc|}
 & & & & & x_1 & \\
 & & & a_1 & a_2 & \ldots & a_{n_1} \\
\hline
 & & b_1 & (a_1,b_1) & (a_2,b_1) & \ldots & (a_{n_1},b_1) \\
 & & b_2 & (a_1,b_2) & (a_2,b_2) & \ldots & (a_{n_1},b_2) \\
 x_2 & & \vdots & \vdots & & \mathrm{O} & \vdots \\
 & & b_{n_2} & (a_1,b_{n_2}) & (a_2,b_{n_2}) & \ldots & (a_{n_1},b_{n_2}) \\
\hline
\end{array}
$$

Phép Lấy Mẫu Có Hoàn lại và Có Thứ tự

Giả sử rằng chúng ta chọn k đối tượng từ tập hợp A có n đối tượng khác nhau theo cách có hoàn lại, nghĩa là, sau khi lấy một đối tượng và ghi lại các đặc điểm của nó vào một danh sách theo thứ tự, đối tượng được trả lại tập hợp trước khi chọn lần tiếp theo. Chúng ta coi A như là một "quần thể." Kết quả thí nghiệm là một bộ-k có thứ tự

$(x_1, \ldots, x_k)$

với $x_i \in$ A và $i = 1, \ldots, k$. Đẳng thức (2.16) với $n_1 = n_2 = \ldots = n_k = n$ dẫn đến rằng:

$$\text{số của các bộ-}k\text{ có thứ tự khác nhau} = n^k. \qquad (2.17)$$

VÍ DỤ 2.12 Một hộp có 5 viên bi đánh số từ 1 đến 5. Giả sử ta lấy 2 viên bi từ hộp có thay thế. Có bao nhiêu bộ đôi có thứ tự khác nhau? Xác suất để bộ đôi có hai số như nhau bằng bao nhiêu?

Đẳng thức (2.17) chỉ ra rằng số các bộ đôi có thứ tự là $5^2 = 25$. Hình 2.10(a) chỉ ra 25 cặp có thể. Năm trong 25 kết cục có hai lần lấy được số như nhau. Nếu chúng ta giả thiết tất cả các cặp đồng xác suất, thì xác suất để hai lần lấy được cùng một số là 5/25 = 0.2.

Phép Lấy Mẫu Có Hoàn lại và Có Thứ tự

Giả sử chúng ta chọn k đối tượng theo thứ tự mà không hoàn lại từ một quần thể A gồm n đối tượng khác nhau. Rõ ràng rằng $k \le n$. Số các kết cục có thể có trong lần rút đầu tiên là $n_1 = n$; số kết cục có thể trong lần lấy thứ 2 là $n_2 = n - 1$, gồm tất cả n đối tượng trừ một đối tượng được chọn trong lần lấy thứ nhất; và vân vân tới $n_k = n - (k - 1)$ trong lần lấy cuối cùng. Khi đó đẳng thức (2.16) cho:

$$\text{số các bộ-}k\text{ sắp thứ tự khác nhau} = n(n-1)\ldots(n-k+1). \qquad (2.18)$$

HÌNH 2.10 Bản liệt kê các kết cục có thể của các dạng khác nhau từ phép lấy mẫu 2 viên bi từ hộp có 5 quả bong khác nhau

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2,5)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)

(a) Các cặp được sắp thứ tự từ phép lấy mẫu có hoàn lại

	(1, 2)	(1, 3)	(1, 4)	(1, 5)
(2, 1)		(2, 3)	(2, 4)	(2, 5)
(3, 1)	(3, 2)		(3, 4)	(3, 5)
(4, 1)	(4, 2)	(4, 3)		(4, 5)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	

(b) Các cặp được sắp thứ tự từ phép lấy mẫu không hoàn lại

	(1, 2)	(1, 3)	(1, 4)	(1, 5)
		(2, 3)	(2, 4)	(2, 5)
			(3, 4)	(3, 5)
				(4, 5)

(c) Các cặp từ phép lấy mẫu không hoàn lại và không tính đến thứ tự

VÍ DỤ 2.13

Một hộp có 5 viên bi được đánh số từ 1 đến 5. Giả sử chúng ta lấy 5 viên bi theo thứ tự mà không hoàn lại. Hỏi có thể có bao nhiêu cặp sắp thứ tự khác nhau? Hỏi xác suất để viên bi thứ nhất có số lớn hơn số của viên bi thứ hai bằng bao nhiêu?

Đẳng thức (2.18) chỉ ra rằng số của các bộ đôi sắp thứ tự là: 5(4) = 20. 20 bộ đôi sắp thứ tự có thể chỉ ra được ở trong Hình 2.10(b). 10 bộ đôi sắp thứ tự trong Hình 2.10(b) có số thứ tự thứ nhất lớn hơn số thứ hai, do đó xác suất của biến cố này là 10/20 = 1/2.

VÍ DỤ 2.14

Một hộp có 5 viên bi đánh số từ 1 đến 5. Giả sử chúng ta lấy 3 viên bi có hoàn lại. xác suất để cả 3 viên bi khác nhau bằng bao nhiêu?

Từ Đẳng thức (2.17) có 5^3 = 125 kết cục có thể, mà chúng ta sẽ giả thiết đồng xác suất. Số của các kết cục mà ở 3 lần lấy dược đưa ra bởi đẳng thức (2.18): 5(4)(3) = 60. do đó xác suất để 3 viên bi khác nhau là 60/125 = .48.

Các phép Hoán vị của n phần tử

Xét phép lấy mẫu không hoàn lại với $k = n$. Đây là phép rút từng đối tượng một từ hộp có n đối tượng cho đến hết. Do vậy số các thứ tự có thể (chỉnh hợp, hoán vị) của n đối tượng khác nhau bằng số các bộ-n sắp xếp thứ tự theo phép lấy mẫu không hoàn lại với $k = n$. Từ đẳng thức (2.18) chúng ta có:

$$\text{số các hoán vị của } n \text{ phần tử} = n(n-1)\ldots(2)(1) \triangleq n! \qquad (2.19)$$

Chúng ta gọi $n!$ là ***n* giai thừa** (**factorial**).

Chúng ta sẽ gặp $n!$ xuất hiện trong nhiều công thức tổ hợp. Với n lớn, **công thức Stirling** là rất có ích :

$$n! \sim \sqrt{2\pi}\, n^{n+1/2} e^{-n}, \qquad (2.20)$$

ở đây kí hiệu ~ chỉ ra tỉ số của 2 vế tiến đến 1 khi $n \to \infty$ (Feller, p. 52).

VÍ DỤ 2.15

Hãy tìm số các hoán vị của 3 đối tượng khác nhau {1, 2, 3}. Đẳng thức (2.19) cho 3! = 3(2)(1) = 6. Sáu hoán vị là :

123 312 231 132 213 321

VÍ DỤ 2.16

Giả sử rằng 12 viên bi được đặt một cách ngẫu nhiên vào 12 ô, có thể đặt nhiều hơn 1 bi vào một ô. Hỏi xác suất để tất cả các ô đều có bi bằng bao nhiêu?

Phép đặt mỗi bóng vào một ô có thể hình dung như phép chọn mỗi ô một số từ 1 đế 12. Đẳng thức (2.17) có 12^{12} phép đặt có thể 12 bi vào 12 ô. Để tất cả các ô đều có bi, viên bi thứ nhất đặt vào ô bất kỳ từ 12 ô, viên bi thứ 2 đặt vào ô bất kỳ từ 11 ô còn lại, và vân vân. Do đó số các phép đặt để tất cả các ô đều có bi bằng 12!. Nếu chúng ta giả thiết rằng tất cả 12^{12} phép đặt có thể là đồng xác suất, chúng ta tìm được xác suất để tất cả các ô đều có bóng là

$$\frac{12!}{12^{12}}=\left(\frac{12}{12}\right)\left(\frac{11}{12}\right)\ldots\left(\frac{1}{12}\right)=5.37(10^{-5}).$$

Câu trả lời này là đáng ngạc nhiên nếu chúng ta nếu chúng ta giải thích lại câu hỏi như sau đây. Giả sử rằng có 12 vụ rơi máy bay, xảy ra một cách ngẫu nhiên trong một năm, xác suất để có đúng một vụ rơi máy bay mỗi tháng bằng bao nhiêu? Kết quả trên chỉ ra rằng xác suất này là rất nhỏ. Như vậy một mô hình với giả định rằng sự rơi máy bay xảy ra một cách ngẫu nhiên theo thời gian không đủ để dự báo trước rằng chúng phân phối đều theo thời gian (Feller, p. 32).

Phép Lấy Mẫu Không Hoàn lại và Không tính Thứ tự

Giả sử chúng ta lấy *k* đối tượng từ tập hợp có *n* đối tượng khác nhau bằng cách lấy không hoàn lại và chúng ta ghi lại kết quả mà không chú ý tới thứ tự lấy ra. (Bạn có thể hình dung đặt mỗi đối tượng được lấy ra vào một bình khác nhau, như vậy chúng ta có đủ tập hợp *k* đối tượng mà không chú ý tới thứ tự chọn các đối tượng. Chúng ta gọi tập *k* đối tượng được lấy là một "tổ hợp chập *k*".

Từ đẳng thức (2.19), có *k*! bộ *k*! sắp thứ tự có thể như thế C_n^k ký hiệu số các tổ hợp chập *k* của *n* phần tử, khi đó $C_n^k k!$ là tổng số tất cả các chỉnh hợp không lặp chập *k* của *n* phần tử được cho bởi Đẳng thức (2.18). Do đó:

$$C_n^k k! = n(n-1)\ldots(n-k+1), \qquad (2.21)$$

và số các tổ hợp chập *k* khác nhau của *n* phương trình, $k \le n$, là :

$$C_n^k = \frac{n(n-1)\ldots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!} \triangleq \binom{n}{k}. \qquad (2.22)$$

Biểu thức $\binom{n}{k}$ được gọi là **hệ số của khai triển nhị thức Newton** và được đọc là "tổ hợp chập *k* của *n* phần tử" (hay "*n* chọn *k*").

Chú ý rằng việc chọn *k* đối tượng từ *n* đối tượng là tương đương với việc chọn (*n* – *k*) đối tượng từ *n* đối tượng. Khi đó chúng ta có (xem Bài tập 46):

$$\binom{n}{k} = \binom{n}{n-k}.$$

VÍ DỤ 2.17 Tìm số cách lấy 2 đối tượng từ A = {1, 2, 3, 4, 5} mà không chú ý tới thứ tự.

Đẳng thức (2.22) cho :

$$\binom{5}{2} = \frac{5!}{2!\,3!} = 10\,.$$

Hình 2.10(c) cho 10 cặp.

VÍ DỤ 2.18

Hãy tìm số các hoán vị khác nhau của k bi trắng và $n - k$ bi đen.

Bài toán này là tương đương với bài toán lấy mẫu sau: Đặt n vật được đánh số từ 1 đến n vào trong một hộp, mà ở đó mỗi vật thể hiện một ví trí sắp xếp của bi; lấy một tổ hợp chập k vật và đặt k quả bóng trắng vào các vị trí tương ứng. Mỗi một tổ hợp chập k cho một sự sắp xếp khác nhau của k quả bóng trắng và $n - k$ viên bi đen. Do vậy số các hoán vị khác nhau của k viên bi trắng và $n - k$ viên bi đen là C_n^k.

Như là một trường hợp riêng lấy $n = 4$ và $k = 2$. Số các tổ hợp chập 2 của 4 phần tử là:

$$\binom{4}{2} = \frac{4!}{2!2!} = \frac{4(3)}{2(1)} = 6\,.$$

6 hoán vị khác nhau của 2 bi trắng (số 0) và 2 bi đen (số 1) là

1100 0110 0011 1001 1010 0101.

VÍ DỤ 2.19
Quản lý Chất lượng

Một lô có 50 sản phẩm trong đó có 10 phế phẩm. Giả sử 10 sản phẩm được lấy ngẫu nhiên và được kiểm tra. Hỏi xác suất có đúng 5 sản phẩm được kiểm tra là phế phẩm bằng bao nhiêu?

Số cách lấy ra 10 sản phẩm từ lô có 50 sản phẩm bằng số tổ hợp chập 10 từ 50 đối tượng:

$$\binom{50}{10} = \frac{50!}{10!\,40!}.$$

Số các cách chọn 5 phế phẩm và 5 chính phẩm từ lô 50 sản phẩm là tính N_1N_2, ở đây N_1 là số cách chọn 5 phế phẩm từ 10 phế phẩm, và N_2 là cách chọn 5 chính phẩm từ 40 chính phẩm. Do đó xác suất để đúng 5 sản phẩm được kiểm tra là phế phẩm bằng:

$$\frac{\binom{10}{5}\binom{40}{5}}{\binom{50}{10}} = \frac{10!\,40!\,10!\,40!}{5!\,5!\,35!\,5!\,50!} = .016\,.$$

Ví dụ 2.18 chỉ ra rằng phép lấy mẫu không hoàn lại và không tính đến thứ tự là tương đương với phép phân chia tập gồm *n* đối tượng khác nhau thành 2 tập: B, gồm *k* đối tượng được lấy ra khỏi hộp, và B^c gồm $n - k$ đối tượng còn lại trong hộp. Giả sử chúng ta chia tập gồm *n* đối tượng thành F tập con B_1, B_2, …,B_F, ở đây B_j có k_j phần tử và $k_1 + k_2 + \ldots + k_F = n$.

Trong Bài tập 47 đã chỉ ra số các cách phân chia khác nhau là

$$\frac{n!}{k_1!k_2!\ldots k_3!}.$$

Đẳng thức (2.23) được gọi là **hệ số của khai triển đa thức**. Hệ số của khai triển nhị thức là trường hợp $F = 2$ của hệ số khai triển đa thức.

VÍ DỤ 2.20 Một con xúc xắc được tung 12 lần. Hỏi có bao nhiêu dãy khác nhau của sự xuất hiện các mặt của con xúc xắc (có số thuộc tập {1, 2, 3, 4, 5, 6}) có mỗi số xuất hiện đúng 2 lần? Xác suất xuất hiện các dãy như vậy bằng bao nhiêu?

Số các dãy khác nhau mà ở đó mỗi mặt con xúc xắc xuất hiện đúng 2 lần là bằng số cách phân chia tập hợp {1, 2, 3, … , 12} thành 6 tập con, mỗi tập có 2 phương trình bằng:

$$\frac{12!}{2!\,2!\,2!\,2!\,2!\,2!} = \frac{12!}{2^6} = 7{,}484{,}400$$

Từ đẳng thức (2.17) chúng ta có 6^{12} kết cục có thể trong 12 lần tung 1 con xúc xắc. Nếu chúng ta giả thiết rằng tất cả các kết cục này đồng xác suất khi đó xác suất nhận được dãy mà ở đó mỗi mặt xuất hiện đúng 2 lần là

$$\frac{12!/2^6}{6^{12}} = \frac{7{,}484{,}400}{2{,}176{,}782{,}336} \approx 3{,}4(10^{-3}).$$

Phép Lấy Mẫu Có Hoàn lại và Không tính đến Thứ tự

Giả sử chúng ta lấy *k* đối tượng từ tập *n* đối tượng khác nhau theo cách có hoàn lại và chúng ta ghi lại kết quả mà không chú ý tới thứ tự. Điều này có thể hoàn thành bằng cách điền vào một bảng có *n* cột, mỗi một cột dành cho một đối tượng. Mỗi lần một đối tượng được lấy ra, một dấu "x" được đặt vào cột tương ứng. Ví dụ, nếu chúng ta lấy 5 đối tượng từ 4 đối tượng khác nhau, một kết cục có thể như sau:

Đối tượng 1		*Đối tượng 2*		*Đối tượng 3*		*Đối tượng 4*
xx	/		/	x	/	xx

ở đây, ký hiệu gạch chéo (" / ") được dùng để phân ra các cột khác nhau. Chú ý rằng bảng này có thể tóm tắt thành dãy:

xx / / x / xx

ở đây n – 1 dấu "/" là đường giữa các cột, và không có gì ở giữa các gạch chéo "/" nếu đối tượng tương ứng không được chọn ra. Mỗi cách sắp xếp khác nhau của 5 "x" và 3 "/" dẫn tới tới một bảng khác. Nếu chúng ta đồng nhất "x" với "viên bi trắng" và "/" với "viên bi đen", khi đó bài toán này đã được nhận xét ở ví dụ 2.18 và số các cách sắp xếp khác nhau bằng $\binom{8}{3}$.

Trong trường hợp tổng quát, khi bảng có *k* ký hiệu "x" và *n* – 1 ký hiệu "/". Khi đó *số các cách lấy ra k phần tử từ tập n phần tử khác nhau theo cách không hoàn lại và không để ý đến thứ tự* bằng

$$\binom{n-1+k}{k} = \binom{n-1+k}{n-1}.$$

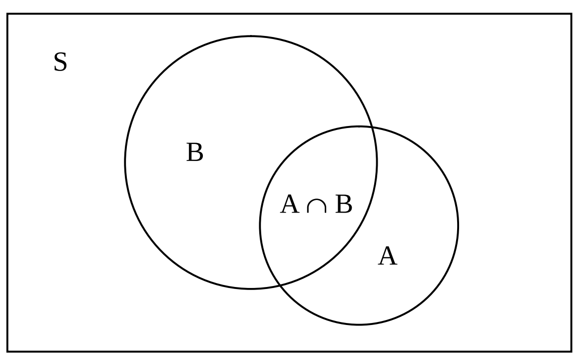
2.4 XÁC SUẤT CÓ ĐIỀU KIỆN

Hoàn toàn tự nhiên chúng ta xác định sự kiện liên quan giữa hai biến cố, A và B, khi thông tin về sự suất hiện của một biến cố, gọi là B, làm thay đổi xác suất xuất hiện của biến cố khác, gọi là A. Điều này dẫn đến việc tìm **xác suất có điều kiện**, P[A | B], của biến cố A khi biến cố B đã xảy ra.
Xác suất có điều kiện được xác định bởi:

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} \qquad \text{với } P[B] > 0. \qquad (2.24)$$

Thông tin về sự xảy ra của biến cố B dẫn đến kết cục của thí nghiệm nằm trong tập B. Do đó khi tính P[A | B] chúng ta có thể coi như thí nghiệm có không gian mẫu B như được chỉ ra trong hình vẽ 2.11. Biến cố A xảy ra trong không gian mẫu rút gọn nếu và chỉ nếu kết cục ζ nằm trong A ∩ B. Đẳng thức (2.24) chuẩn hoá xác suất của các biến cố xảy ra đồng thời với B. Khi đó, nếu chúng ta lấy A = B, đẳng thức (2.24) dẫn đến P[B | B] = 1, như được chờ đợi. Từ đó dễ dàng chỉ ra rằng P[A | B], với B cố định thoả mãn các tiên đề xác suất 9 xem Bài tập 49.)

HÌNH 2.11
Nếu đã biết B xảy ra, khi đó A xảy ra chỉ nếu A ∩ B xảy ra.



Nếu chúng ta coi xác suất như là tần suất tương đối, khi đó P[A | B] có thể tần suất tương đối của biến cố A ∩ B trong các thí nghiệm mà ở đó B xảy ra. Giả sử rằng thí nghiệm được tiến hành n lần, và giả sử rằng biến cố B xảy ra n_B lần, và biến cố A ∩ B xảy ra $n_{A\cap B}$ lần. Tần suất tương đối của biến cố mà ta quan tâm là:

$$\frac{n_{A\cap B}}{n_B}=\frac{n_{A\cap B}/n}{n_B/n}\to\frac{P[A\cap B]}{P[B]},$$

ở đây chúng ta giả thiết rằng P[B] > 0. Điều này phù hợp với đẳng thức (2.24).

VÍ DỤ 2.21 Một viên bi được lấy từ một hộp có hai bi đen, được đánh số 1 và 2, và hai bi trắng được đánh số 3 và 4. Số và màu của viên bi lấy ra được ghi lại, khi đó không gian mẫu là {(1, *b*), (2, *b*), (3, *w*),
(4, *w*)}. Giả sử rằng các kết cục là đồng khả năng, hãy tìm P[A | B] và P[A | C], ở đây A, B và C là các biến cố sau:

A = {(1, *b*), (2, *b*)}, "bi đen được lấy,"

B = {(2, *b*), (4, *w*)}, "bi có số chẵn được lấy," và

C = {(3, *w*), (4, *w*)}, "số của viên bi lớn hơn 2."

Do P[A ∩ B] = P[(2, *b*) và P[A ∩ C] = P[∅] = 0, Đẳng thức (2.21) dẫn đến:

$$P[A\,|\,B]=\frac{P[A\cap B]}{P[B]}=\frac{.25}{.5}=.5=P[A]$$

$$P[A\,|\,C]=\frac{P[A\cap C]}{P[C]}=\frac{0}{.5}\neq P[A].$$

Trong trường hợp thứ nhất, thông tin của B không ảnh hưởng tới xác suất của A. Trong trường hợp thứ hai, thông tin về B dẫn đến rằng A không xảy ra.

Nếu chúng ta nhân cả 2 vế của định nghĩa P[A | B] với P[B] chúng ta nhận được

$$P[A\cap B] = P[A\,|\,B]\,P[B]. \qquad (2.25a)$$

Bằng cách tương tự chúng ta cũng nhận được

$$P[A\cap B] = P[B\,|\,A]\,P[A]. \qquad (2.25b)$$

Trong ví dụ sau đây chúng ta chứng tỏ đẳng thức này có ích như thế nào trong việc tìm xác suất trong các dãy các thí nghiệm. Ví dụ cũng dẫn đến **sơ đồ hình cây** để dễ dàng tính xác suất.

VÍ DỤ 2.22 Một hộp có hai viên bi đen và 3 viên bi trắng. Hai viên bi được chọn một cách ngẫu nhiên từ hộp theo cách không hoàn lại và ghi lại dãy các màu của viên bi được lấy ra. Hãy tìm xác suất để cả hai viên bi đều màu đen.

Thí nghiệm này là dãy của 2 thí nghiệm nhỏ. Chúng ta có thể hình dung dưới dạng sơ đồ hình cây được chỉ ra trong hình 2.12 từ đỉnh cao nhất tới một từ các đỉnh sau đó. Chúng ta đi đến đỉnh 1 của cây nếu kết cục của lần lấy thứ nhất là viên bi đen; khi đó thí nghiệm nhỏ tiếp theo là việc lấy một viên bi từ hộp có một viên bi đen và ba viên bi trắng. Trên một cách khác, nếu kết cục của lần lấy thứ nhất là viên bi trắng khi đó chúng ta đi đến đỉnh 2 của cây và thí nghiệm thứ 2 là việc lấy một viên bi từ hộp có 2 viên bi trắng và hai viên bi đen. Do đó nếu chúng ta biết đỉnh đi đến sau bước thứ nhất, khi đó chúng ta có thể đưa ra được xác suất của kết cục trong thí nghiệm đơn tiếp theo.

Đặt B_1 và B_2 là các biến cố chứa kết cục lấy được viên bi đen lần thứ nhất và lần thứ 2 một cách tương ứng. Từ Đẳng thức (2.25b) chúng ta có

$$P[B_1 \cap B_2] = P[B_2 \mid B_1]P[B_1].$$

Theo sơ đồ thuật ngữ của sơ đồ hình cây trong hình 2.12, $P[B_1]$ là xác suất đi đến đỉnh 1 và $P[B_2 \mid B_1]$ là xác suất đi đến đỉnh trái nhất sau đó từ đỉnh 1. Ta có $P[B_1] = 2/5$ do lần lấy thứ nhất từ họp có 2 bóng đen và 3 bóng trắng; $P[B_2 \mid B_1] = 1/4$ do B_1 xảy ra, phép lấy thứ 2 từ hộp có 1 bi đen và 3 bi trắng.

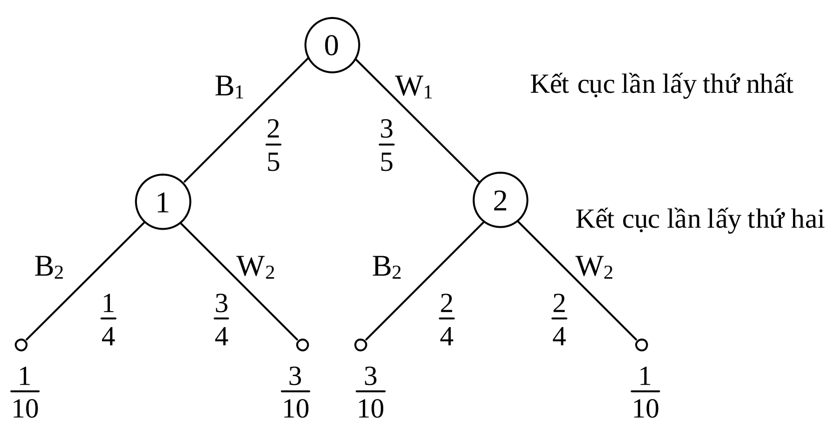
Như vậy

$$P[B_1 \cap B_2] = \frac{1}{4}\frac{2}{5} = \frac{1}{10}.$$

Một cách tổng quát, xác suất của dãy bất kỳ các màu nhận được bởi việc nhân các xác suất tương ứng với đỉnh phép chuyển từ đỉnh nọ tới đỉnh kia trong Hình 2.12.

HÌNH 2.12

Các phép chuyển từ đỉnh cao nhất, tới đỉnh tiếp theo tương ứng với các kết cục có thể khi lấy hai viên bi từ hộp mà không hoàn lại. Xác suất của phép chuyển là tích của các xác suất của các phép chuyển liên hợp.



VÍ DỤ 2.23

Hệ thống Thông

Rất nhiều hệ thống thông tin có thể được mô tả theo cách như tin Nhị thức

sau: Thứ nhất, người gửi dùng tín hiệu 0 hoặc 1 để truyền đi. Thứ hai, máy thu đưa ra quyết định tín hiệu nào được truyền đi, dựa trên cơ sở tín hiệu sẽ nhận được. Giả sử rằng người gửi tín hiệu 0 với xác suất 1– *p* và tín hiệu 1 với xác suất *p*, và giả sử rằng máy thu đưa ra quyết định một cách ngẫu nhiên với xác suất sai là ε. Với i = 0, 1 đặt A_i là biến cố "tín hiệu *i* được truyền đi," và B_i là biến cố "máy thu quyết định tín hiệu *i* đã được truyền đi." Hãy tìm các xác suất $P[A_i \cap B_j]$ với i = 0, 1 và j = 0, 1.

Sơ đồ hình cây cho thí nghiệm này được chỉ ra trong Hình 2.13. Khi đó chúng ta nhận được các xác suất cần thiết

$$P[A_0 \cap B_0] = (1 - p)\,(1 - \varepsilon),$$

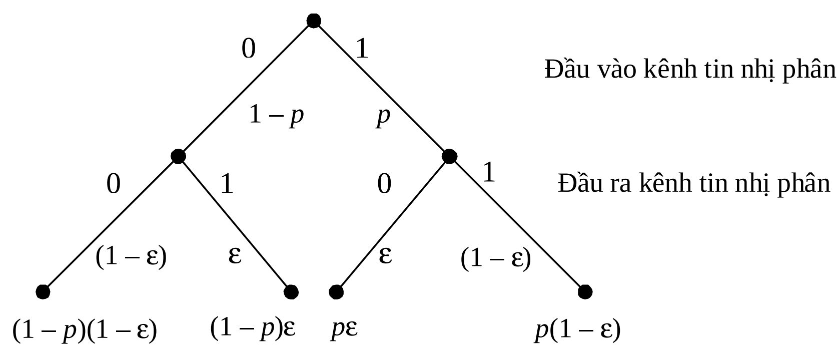
$$P[A_0 \cap B_1] = (1 - p)\,\varepsilon,$$

$$P[A_1 \cap B_0] = p\,\varepsilon, \quad \text{và}$$

$$P[A_1 \cap B_1] = p\,(1 - \varepsilon).$$

HÌNH 2.13

Xác suất của các cặp vào ra trong hệ truyền tin nhị phân.



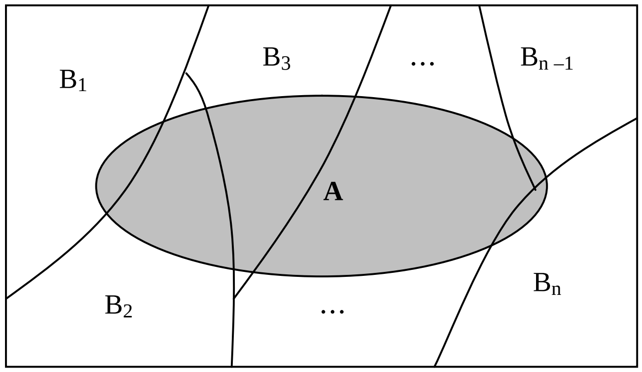
Giả sử $B_1, B_2, \ldots, B_n$ là các biến cố xung khắc từng đôi mà hợp của chúng bằng không gian mẫu như được chỉ ra trong Hình 2.14. Chúng ta coi các tập hợp này như là một **phân hoạch** (**partion**) của S. Biến cố A bất kỳ có thể biểu diễn như là hợp của các biến cố xung khắc từng đôi theo cách sau:

$$A = A \cap S = A \cap (B_1 \cup B_2 \cup \ldots \cup B_n)$$

$$= (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_n).$$

HÌNH 2.14

Một phân hoạch S thành n tập hợp rời nhau.



(xem Hình 2.14). Do Hệ quả 4, xác suất của A bằng:

$$P[A] = P[A \cap B_1] + P[A \cap B_2] + \cdots + P[A \cap B_n].$$

Bằng việc áp dụng Đẳng thức (2.25a) tới mỗi từ của vế phải chúng ta nhận được công thức **xác suất toàn phần** (**theorem on total probability**):

$$P[A] = P[A|B_1]\,P[B_1] + P[A|B_2]\,P[B_2] + \ldots P[A|B_n]\,P[B_n]. \quad (2.26)$$

Kết quả này thực sự có ích khi thí nghiệm có thể coi như một dãy gồm hai thí nghiệm đơn như được chỉ ra trong sơ đồ hình cây ở Hình 2.12.

VÍ DỤ 2.24

Trong thí nghiệm được thảo luận trong Ví dụ 2.22, hãy tìm xác suất của biến cố W_2 sao cho viên bi thứ 2 có màu trắng.

Các biến cố $B_1 = \{(b, b), (b, w)\}$ và $W_1 = \{(w, b), (w, w)\}$ là một phân hoạch của không gian mẫu, bằng việc ứng dụng đẳng thức (2.26) chúng ta có:

$$P[W_2] = P[W_2 \mid B_1]\,P[B_1] + P[W_2 \mid W_1]\,P[W_1]$$

$$= \frac{3}{4}\frac{2}{5} + \frac{1}{2}\frac{3}{5} = \frac{3}{5}.$$

Thật thú vị khi chúng ta chú ý rằng xác này là bằng xác suất lấy được viên bi trắng trong lần lấy đầu tiên. Kết quả này có được do chúng ta giả thiết rằng chúng ta không biết kết cục của lần lấy thứ nhất.

VÍ DỤ 2.25

Một quá trình sản xuất tạo ra một hỗn hợp gồm chip nhớ "tốt" và chip nhớ "xấu". Thời gian sống của chip tốt tuân theo luật phân phối mũ được đưa ra trong Ví dụ 2.10, với tốc độ hỏng α. Thời gian sống của chip xấu cũng tuân theo luật mũ, nhưng với tốc độ hỏng là 1000α. Giả sử rằng, tỉ lệ chip tốt là $1 - p$ và chip xấu là p. Hãy tìm xác suất để một chip được lấy một cách ngẫu nhiên vẫn làm việc sau t giây.

Lấy C là biến cố "chip vẫn làm việc sau t giây," và lấy

G là biến cố "chip được lấy là chip tốt" và B là biến cố "chip dược lấy là chip xấu." Do định lý xác suất toàn phần chúng ta có:

$$\begin{aligned} P[C] &= P[C \mid G]\, P[G] + P[C \mid B]\, P[B] \\ &= P[C \mid G]\,(1-p) + P[C \mid B]\, p \\ &= (1-p)e^{-\alpha t} + p e^{-1000\alpha t}, \end{aligned}$$

ở đây ta sử dụng giả thiết $P[C \mid G] = e^{-\alpha t}$ và $P[C \mid B] = e^{-1000\alpha t}$.

Quy tắc Bayes

Giả sử $B_1, B_2, \ldots, B_n$ là phân hoạch của không gian mẫu S. Giả sử rằng, biến cố A xảy ra, khi đó xác suất của biến cố B_j bằng bao nhiêu ? Do định nghĩa xác suất có điều kiện chúng ta có :

$$P[B_j \mid A] = \frac{P[A \cap B_j]}{P[A]} = \frac{P[A \mid B_j]P[B_j]}{\sum_{k=1}^{n} P[A \mid B_k]P[B_k]} \tag{2.27}$$

ở đây chúng ta sử dụng định lý xác suất toàn phần đối với P[A]. Đẳng thức (2.27) được gọi là **quy tắc Bayes**.

Quy tắc Bayes cũng thường được sử dụng trong tình thế sau chúng ta có một thí nghiệm ngẫu nhiên mà ở đó các biến cố mà ta quan tâm tạo ra một phân hoạch "xác suất tiên nghiệm" của các biến cố này, $P[B_j]$, là xác suất của các biến cố trước khi thí nghiệm được thực hiện. Bây giờ chúng ta giả thiết rằng thí nghiệm đã được thực hiện, và chúng ta biết rằng, biến cố A đã xảy ra; "Xác suất hậu nghiệm" là xác suất theo phân hoạch, $P[B_j \mid A]$, được tính theo thông tin đã có. Hai ví dụ sau đây minh họa cho tình thế này.

VÍ DỤ 2.26
Hệ thống Thông tin Nhị thức

Trong hệ truyền thông nhị thức trong Ví dụ 2.23, hãy tìm tín hiệu vào có xác suất lớn hơn với tín hiệu ra là 1. Giả sử rằng xác suất tiên nghiệm của 2 tín hiệu đầu vào 0 và 1 là đồng xác suất.

Gọi A_k là biến cố để tín hiệu vào là k, $k = 0, 1$ do A_0 và A_1 là một phân hoạch của không gian mẫu của cặp vào-ra. Đặt B_1 là biến cố "máy nhận tín hiệu ra là 1". Xác suất của B_1 bằng :

$$\begin{aligned} P[B_1] &= P[B_1 \mid A_0]\, P[A_0] + P[B_1 \mid A_1]\, P[A_1] \\ &= \varepsilon\left(\frac{1}{2}\right) + (1-\varepsilon)\left(\frac{1}{2}\right) = \frac{1}{2} \end{aligned}$$

Ứng dụng quy tắc Bayes, chúng ta nhận được xác suất hậu nghiệm

$$P[A_0 \mid B_1] = \frac{P[B_1 \mid A_0]P[A_0]}{P[B_1]} = \frac{\varepsilon/2}{1/2} = \varepsilon$$

$$P[A_1 \mid B_1] = \frac{P[B_1 \mid A_1]}{P[B_1]} = \frac{(1-\varepsilon)/2}{1/2} = (1-\varepsilon).$$

Do đó, nếu ε nhỏ hơn 1/2 , tín hiệu vào là 1 hợp lý hơn tín hiệu vào là 0 khi một tín hiệu nhận được ở đầu ra của kênh.

VÍ DỤ 2.27
Quản lý Chất lượng

Xét các chip nhớ được thảo luận ở Ví dụ 2.25. Gọi tỷ lệ chip xấu là *p* và tốc độ hỏng cần phải nhanh hơn chip tốt. Giả sử rằng, để "loại bỏ" chip xấu, mỗi chip được kiểm tra *t* giây trước khi xuất xưởng. Chip hỏng bị loại chip vẫn làm việc được đưa ra thị trường. Hãy tìm giá trị *t* để 99% chip được đưa ra thị trường là chip tốt.

Đặt *t* là biến cố "chip vẫn làm việc sau *t* giây", đặt G là biến cố "chip là tốt" và B biến cố "chip là xấu". Bài toán dẫn đến tìm giá trị *t* sao cho:

P[G | C] = .99.

Chúng ta tìm được P[G | C] bằng việc áp dụng quy tắc Bayes:

$$P[G \mid C] = \frac{P[C \mid G].P[G]}{P[C \mid G]P[G] + P[C \mid B]P[B]}$$

$$= \frac{(1-p)e^{-\alpha t}}{(1-p)e^{-\alpha t} + pe^{-\alpha 1000 t}}$$

$$= \frac{1}{1 + \dfrac{pe^{-\alpha 1000 t}}{(1-p)e^{-\alpha t}}} = .99$$

Đẳng thức trên có thể giải được theo *t*:

$$t = \frac{1}{999\alpha} 1n\left(\frac{99p}{1-p}\right).$$

Ví dụ, nếu 1/α = 20,000 h và *p* = .10, khi đó *t* = 48 h.

2.5 SỰ ĐỘC LẬP CỦA CÁC BIẾN CỐ

Nếu thông tin về sự xảy ra của biến cố B không ảnh hưởng đến xác suất của biến cố A nào đó, khi đó một cách hoàn toàn tự nhiên chúng ta nói rằng biến cố A là độc lập với biến cố B. Theo khái niệm xác suất tình thế này xảy ra khi:

$$P[A] = P[A \mid B] = \frac{P[A \cap B]}{P[B]}.$$

Đẳng thức trên có vấn đề khi mà vế phải không xác định với P[B] = 0.

Chúng ta sẽ định nghĩa 2 biến cố A và B độc lập nếu:

$$P[A \cap B] = P[A]\,P[B] \tag{2.28}$$

Đẳng thức (2.28) kéo theo cả 2 đẳng thức:

$$P[A \mid B] = P[A] \tag{2.29a}$$

và

$$P[B \mid A] = P[B] \tag{2.29b}$$

Chú ý rằng, đẳng thức (2.29a) kéo đẳng thức (2.28) khi $P[B] \neq 0$.

VÍ DỤ 2.28 Một viên bi được lấy từ hộp gồm hai viên bi đen, được đánh số 1 và 2, và hai viên bi trắng được đánh số 3 và 4. Đặt các biến cố A, B và C được xác định như sau:

$A = \{(1, b), (2, b)\}$, "viên bi đen được lấy"

$B = \{(2, b), (4, w)\}$, "viên bi có số chẵn được lấy"và

$C = \{(3, w), (4, w)\}$, "số trên viên bi lớn hơn 2".

Các biến cố A và B có độc lập? Các biến cố A và C có độc lập? Trước hết chúng ta xét biến cố A và B. Các xác suất theo yêu cầu của đẳng thức (2.28) là :

$$P[A] = P[B] = \frac{1}{2},$$

và

$$P[A \cap B] = P[\{(2, b)\}] = \frac{1}{4}.$$

Như vậy

$$P[A \cap B] = \frac{1}{4} = P[A]P[B],$$

và các biến cố A và B độc lập. Đẳng thức (2.29b) cho cái nhìn sáng tỏ hơn ý nghĩa của sự độc lập:

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{P[\{(2, b)\}]}{P[\{(2, b), (4, w)\}]} = \frac{1/4}{1/2} = \frac{1}{2}$$

$$P[A] = \frac{P[A]}{P[S]} = \frac{P[\{(1, b), (2, b)\}]}{P[\{(1, b), (2, b), (3, w), (4, w)\}]} = \frac{1/2}{1}.$$

Hai đẳng thức này dẫn đến $P[A] = P[A \mid B]$ do tỷ lệ của các kết cục trong S thuận lợi cho sự xảy ra A là bằng tỷ lệ của các kết cục trong B thuận lợi cho A. Như vậy, thông tin về sự xảy ra của B không ảnh hưởng đến xác suất xảy ra A.

Các biến cố A và C không độc lập do $P[A \cap C] = P[\varnothing] = 0$, vậy

$$P[A \mid C] = 0 \neq P[A] = .5.$$

Thực tế A và C là hai biến cố xung khắc do $A \cap C = \varnothing$, vậy sử

---

xảy ra của C có nghĩa rằng A không thể xảy ra.

Xét tổng quát, nếu hai biến cố có xác suất khác không và xung khắc, khi đó chúng không thể độc lập. Giả sử có hai biến cố độc lập *và* xung khắc, khi đó:

$$0 = P[A \cap B] = P[A]\, P[B],$$

mà điều này dẫn đến ít nhất một trong hai biến cố có xác suất bằng 0.

VÍ DỤ 2.29 Hai số x và y được lấy một cách ngẫu nhiên giữa 0 và 1. Đặt các biến cố A, B và C được xác định như sau:

$$A = \{x > 0.5\}, \qquad B = \{y > 0.5\}, \qquad \text{và } C = \{x > y\}.$$

Các biến cố A và B có độc lập hay không? Các biến cố A và C có độc lập hay không?

Hình 2.15 chỉ ra các miền của hình vuông đơn vị tương ứng với các biến cố trên. Sử dụng đẳng thức (2.29a), chúng ta có:

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{1/4}{1/2} = \frac{1}{2} = P[A],$$

như vậy các biến cố A và B độc lập. Hơn nữa chúng ta có tỷ lệ của các kết cục trong S thuận lợi với A là bằng tỷ lệ của các kết cục trong B thuận lợi với A.

Sử dụng đẳng thức (2.29b) chúng ta có:

$$P[A \mid C] = \frac{P[A \cap C]}{P[C]} = \frac{3/8}{1/2} = \frac{3}{4} \neq \frac{1}{2} = P[A]$$

như vậy các biến cố A và C là không độc lập. Thực vậy, từ Hình 2.15(b) chúng ta có thể nhận thấy rằng, thông tin về việc x lớn hơn y làm tăng xác suất để lớn hơn 0.5.

Điều kiện nào để ba biến cố A, B và C thoả mãn điều kiện độc lập? Trước hết chúng ta cần phải đôi 1 độc lập, nghĩa là,

$$P[A \cap B] = P[A]\, P[B], \qquad P[A \cap C] = P[A]\, P[C], \text{ và}$$

$$P[B \cap C] = P[B]\, P[C].$$

Hơn nữa thông tin về sự xảy ra đồng thời của hai biến cố bất kỳ, gọi là A và B , sẽ không ảnh hưởng đến xác suất của biến cố thứ 3, nghĩa là:

$$P[C \mid A \cap B] = P[C].$$

Để cho điều này xảy ra chúng ta cần phải có:

$$P[C \mid A \cap B] = \frac{P[A \cap B \cap C]}{P[A \cap B]} = P[C].$$

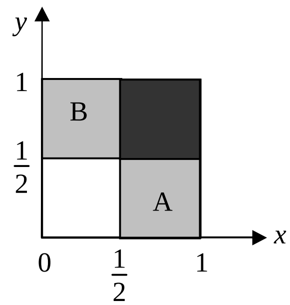
Đến lượt mình, điều này có nghĩa là chúng ta cần phải có

$$P[A \cap B \cap C] = P[A \cap B]\, P[C] = P[A]\, P[B]\, P[C],$$

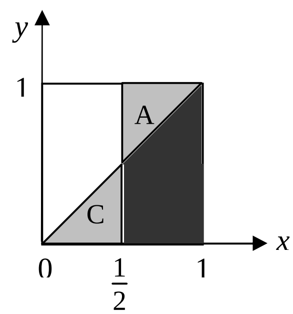
ở đây chúng ta đã sử dụng giả thiết A và B độc lập với nhau. Như thế chúng ta kết luận rằng *ba biến cố A, B và C độc lập nếu xác suất của tích hai biến cố bất kỳ và của ba biến cố là bằng tích của các biến cố thành phần.*

HÌNH 2.15

Ví dụ về sự độc lập và không độc lập của các biến cố



(a) Biến cố A và B độc lập.



(b) Biến cố A và C không độc lập.

Ví dụ sau đây chỉ ra rằng nếu ba biến cố đôi 1 độc lập, không nhất thiết kéo theo $P[A \cap B \cap C] = P[A]\, P[B]\, P[C]$.

VÍ DỤ 2.30 Xét thí nghiệm được thảo luận trong Ví dụ 2.29, ở đó hai số được lấy một cách ngẫu nhiên từ khoảng đơn vị. Đặt các biến cố B, D và F được xác định như sau:

$$B = \left\{ y > \frac{1}{2} \right\}, \qquad D = \left\{ x < \frac{1}{2} \right\}$$

$$F = \left\{ x < \frac{1}{2} \text{ và } y < \frac{1}{2} \right\} \cup \left\{ x > \frac{1}{2} \text{ và } y > \frac{1}{2} \right\}.$$

Ba biến cố được chỉ ra trong hình 2.16. Chúng ta dễ dàng kiểm tra rằng cặp bất kỳ của các biến cố này là độc lập:

$$P[B \cap D] = \frac{1}{4} = P[B]P[D]$$

$$P[B \cap F] = \frac{1}{4} = P[B]P[F], \text{ và}$$

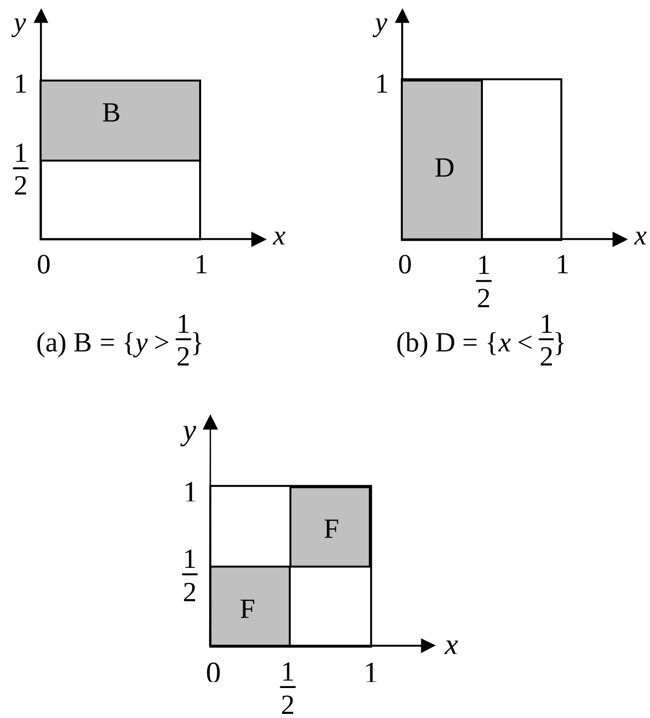
$$P[D \cap F] = \frac{1}{4} = P[D]P[F].$$

Tuy nhiên, ba biến cố là độc lập, do B ∩ D ∩ F = ∅, nên:

$$P[B \cap D \cap F] = P[\varnothing] = 0 \neq P[B]\, P[D]\, P[F] = \frac{1}{8}.$$

HÌNH 2.16

Các biến cố B, D và F là đôi một độc lập, nhưng bộ ba B, D và F là không độc lập.



(a) $B = \{y > \frac{1}{2}\}$

(b) $D = \{x < \frac{1}{2}\}$

(c) $F = \{x < \frac{1}{2} \text{ và } y > \frac{1}{2}\} \cup \{x > \frac{1}{2} \text{ và } y < \frac{1}{2}\}$.

Để tập n biến cố bất kỳ độc lập, xác suất có điều kiện của một biến cố nào đó không thay đổi khi xảy ra đồng thời tập con bất kỳ của các biến cố khác. Yêu cầu này một cách tự nhiên dẫn tới định nghĩa sau của sự độc lập. Các biến cố $A_1, A_2, \ldots, A_n$ *được gọi là* **độc lập** với $k = 2, \ldots, n$,

$$P[A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}] = P[A_{i_1}]P[A_{i_2}]\ldots P[A_{i_k}], \qquad (2.30)$$

ở đây $1 \le i_1 < i_2 < \ldots < i_k \le n$. Với tập gồm n biến cố chúng ta cần phải kiểm tra xác suất của $2^n - n - 1$ giao có thể với nhân tử trong vế phải.

Định nghĩa trên của tính độc lập có vẻ quá cồng kềnh do nó yêu cầu quá nhiều điều kiện cần phải kiểm tra. Tuy nhiên, ứng dụng phổ biến nhất của tính độc lập là khi chúng ta giả định rằng các biến cố của các thí nghiệm riêng lẻ là độc lập. Chúng ta quy cho các thí nghiệm như vậy là các **thí nghiệm độc lập**. Ví dụ, chúng ta thường giả định rằng, kết cục của phép tung đồng xu là các kết cục độc lập với tất cả các lần tung trước và sau đó.

VÍ DỤ 2.31

Giả sử rằng một đồng xu cân đối được tung ba lần và chúng ta quan trắc dãy kết quả của mặt sấp và mặt ngửa. Hãy tìm xác suất của các biến cố sơ cấp.

Không gian mẫu của thí nghiệm này là S = {HHH, HHT, HTH, THH, TTH, THT, HTT, TTT}. Giả thiết về tính cân đối của đồng xu có nghĩa là các kết cục của các lần tung riêng lẻ là đồng xác suất, nghĩa là, P[H] = P[T] = 1/2. Nếu chúng ta giả thiết thêm về tính độc lập của các lần tung đồng xu, khi đó:

$$P[\{HHH\}] = P[\{H\}]\,P[\{H\}]\,P[\{H\}] = \frac{1}{8},$$

$$P[\{HHT\}] = P[\{H\}]\,P[\{H\}]\,P[\{T\}] = \frac{1}{8},$$

$$P[\{HTH\}] = P[\{H\}]\,P[\{T\}]\,P[\{H\}] = \frac{1}{8},$$

$$P[\{THH\}] = P[\{T\}]\,P[\{H\}]\,P[\{H\}] = \frac{1}{8},$$

$$P[\{TTH\}] = P[\{T\}]\,P[\{T\}]\,P[\{H\}] = \frac{1}{8},$$

$$P[\{THT\}] = P[\{T\}]\,P[\{H\}]\,P[\{T\}] = \frac{1}{8},$$

$$P[\{HTT\}] = P[\{H\}]\,P[\{T\}]\,P[\{T\}] = \frac{1}{8}, \text{ và}$$

$$P[\{TTT\}] = P[\{T\}]\,P[\{T\}]\,P[\{T\}] = \frac{1}{8}.$$

VÍ DỤ 2.32
Độ tin cậy của Hệ thống

Một hệ gồm một bộ điều khiển và ba đơn vị ngoại vi. Hệ được gọi là sẵn sàng nếu bộ điều khiển và ít nhất hai đơn vị ngoại vi còn làm việc. Hãy tìm xác suất để hệ sẵn sàng giả thiết rằng tất cả các thành phần hỏng độc lập với nhau.

Định nghĩa các biến cố sau: A là biến cố "bộ điều khiển còn làm việc" và B_i là biến cố "đơn vị ngoại vi thứ *i* còn làm việc", ở đây i = 1, 2, 3. Biến cố F, "hai hoặc hơn hai đơn vị ngoại vi còn làm việc", xảy ra nếu cả ba đơn vị còn làm việc hoặc nếu có đúng hai đơn vị còn làm việc. Do vậy:

$$F = (B_1 \cap B_2 \cap B_3^c) \cup (B_1 \cap B_2^c \cap B_3)$$
$$\cup (B_1^c \cap B_2 \cap B_3) \cup (B_1 \cap B_2 \cap B_3)$$

Chú ý rằng các biến cố trong phép hợp ở trên là đôi một xung khắc. Do đó:

$$P[F] = P[B_1]\,P[B_2]\,P[B_3^c] + P[B_1]\,P[B_2^c]\,P[B_3]$$
$$+ P[B_1^c]\,P[B_2]\,P[B_3] + P[B_1]\,P[B_2]\,P[B_3]$$
$$= 3(1-a)^2a + (1-a)^3,$$

ở đây chúng ta giả thiết rằng mỗi đơn vị ngoại vi hỏng với xác suất a, do vậy $P[B_i] = 1 - a$ và $P[B_i^c] = a$.

Biến cố "hệ sẵn sàng" là khi $A \cap F$. Nếu chúng ta giả thiết rằng bộ điều khiển hỏng với xác suất p, khi đó :

$$
\begin{aligned}
P[\text{"Hệ sẵn sàng"}] = P[A \cap F] &= P[A]\, P[F] \\
&= (1 - p)\, P[F] \\
&= (1 - p)\, \{3(1 - a)^2 a + (1 - a)^3\}.
\end{aligned}
$$

Giả sử $a = 10\%$, khi đó tất cả ba đơn vị ngoại vi làm việc $(1 - a)^3 = 72.9\%$ thời gian và hai đơn vị ngoại vi làm việc và một hỏng $3(1 - a)^2 a = 24.3\%$ thời gian. Như vậy hai hoặc hơn hai đơn vị ngoại vi làm việc 97.2% thời gian. Giả sử rằng bộ điều khiển rất không ổn định với $p = 20\%$, khi đó hệ chỉ sẵn sàng 77.8% thời gian, phần lớn do bộ điều khiển bị hỏng.

Giả sử bộ điều khiển y hệt thứ hai được gắn vào hệ thống và là sẵn sàng nếu có ít nhất một bộ điều khiển đang làm việc và có ít nhất hai hoặc hơn hai đơn vị ngoại vi đang làm việc. Trong Bài tập 69, các bạn sẽ được hỏi để chỉ ra rằng ít nhất một bộ điều khiển đang làm việc với 96% thời gian, và là hệ thống làm việc với 93.3% thời gian. Điều này làm tăng thêm 16% thời gian làm việc của hệ so với hệ chỉ có một bộ điều khiển.

2.6 DÃY CÁC THÍ NGHIỆM LIÊN TIẾP

Nhiều thí nghiệm ngẫu nhiên có thể coi như các thí nghiệm liên tiếp, mà nó gồm một dãy các thí nghiệm đơn. Các thí nghiệm đơn này có thể độc lập, hoặc không. Trong phần này chúng ta sẽ thảo luận các phương pháp để nhận được các xác suất của các biến cố trong các thí nghiệm liên tiếp.

Dãy các Thí nghiệm Độc lập

Giả sử rằng một dãy thí nghiệm gồm các thí nghiệm $E_1, E_2, \ldots, E_n$. Khi đó một kết cục của thí nghiệm này là một bộ-*n* $S = (S_1, \ldots, S_n)$, ở đây S_k là kết cục của thí nghiệm đơn thứ *k*. Không gian mẫu của thí nghiệm liên tiếp được xác định như là tập các bộ-*n* được chỉ ra ở trên và được ký hiệu bởi tích Descartes của các không gian mẫu đơn $S_1 \times S_2 \times \ldots \times S_n$.

Chúng ta thường coi, do điều kiện vật lý, các thí nghiệm đơn là độc lập, theo nghĩa là các kết cục của thí nghiệm đơn bất kỳ không ảnh hưởng đến các kết cục của các thí nghiệm đơn khác. Đặt $A_1, A_2, \ldots, A_n$ là các biến cố sao cho A_k chỉ liên quan đến kết cục của thí nghiệm đơn thứ *k*. Nếu các thí nghiệm đơn là độc lập, khi đó hoàn toàn hợp lý khi giả thiết rằng biến cố ở trên $A_1, A_2, \ldots, A_n$ là độc lập. Do vậy

$$P[A_1 \cap A_2 \cap \ldots \cap A_n] = P[A_1]\, P[A_2] \ldots P[A_n]. \tag{2.31}$$

Biểu thức này cho phép chúng ta tính tất cả các xác suất của tất cả các biến cố của thí nghiệm liên tiếp.

VÍ DỤ 2.33 Giả sử rằng 10 số được lấy một cách ngẫu nhiên từ khoảng [0, 1]. Hãy tìm xác suất để 5 số đầu tiên nhỏ hơn 1/4 và 5 số sau lớn hơn 1/2. Gọi $x_1, x_2, \ldots, x_{10}$ là dãy 10 số, khi đó các biến cố mà ta quan tâm là :

$$A_k = \left\{x_k < \frac{1}{4}\right\} \quad \text{với } k = 1, \ldots, 5$$

$$A_k = \left\{x_k > \frac{1}{2}\right\} \quad \text{với } k = 6, \ldots, 10$$

Nếu chúng ta giả thiết rằng mỗi lần lấy là một số độc lập với các lần lấy khác, thì khi đó:

$$P[A_1 \cap A_2 \cap \ldots \cap A_{10}] = P[A_1]\, P[A_2] \ldots P[A_{10}]$$

$$= \left(\frac{1}{4}\right)^5 \left(\frac{1}{2}\right)^5.$$

Bây giờ chúng ta sẽ bắt đầu với một số mô hình quan trọng của các thí nghiệm bao gồm dãy các thí nghiệm đơn độc lập.

Luật Xác suất Nhị thức

Một **phép thử Bernoulli** bao gồm việc tiến hành thí nghiệm một lần và việc ghi lại biến cố A xảy ra hay không. Kết cục của phép thử Bernoulli được gọi là "thành công" nếu A xảy ra và "thất bại" nếu A không xảy ra. Trong phần này chúng ta sẽ quan tâm tới việc tìm xác suất của k lần thành công trong khi lặp lại *n* lần phép thử Bernoulli một cách độc lập.

Chúng ta có thể coi kết cục của phép thử Bernoulli đơn như kết cục của một lần tung đồng xu mà ở đó xác suất xuất hiện mặt ngửa (thành công) là $p = P[A]$. Xác suất của *k* lần thành công trong *n* phép thử Bernoulli tương đương với xác suất mặt ngửa xuất hiện *k* lần trong *n* lần tung đồng xu.

VÍ DỤ 2.34 Giả sử rằng một đồng xu được tung ba lần. Nếu chúng ta giả sử rằng *các lần tung là độc lập* và xác suất xuất hiện mặt ngửa là *p*, khi đó xác suất của dãy mặt ngửa và sấp là:

$$P[\{HHH\}] = P[\{H\}]\, P[\{H\}]\, P[\{H\}] = p^3,$$
$$P[\{HHT\}] = P[\{H\}]\, P[\{H\}]\, P[\{T\}] = p^2(1-p),$$
$$P[\{HTH\}] = P[\{H\}]\, P[\{T\}]\, P[\{H\}] = p^2(1-p),$$
$$P[\{THH\}] = P[\{T\}]\, P[\{H\}]\, P[\{H\}] = p^2(1-p),$$
$$P[\{TTH\}] = P[\{T\}]\, P[\{T\}]\, P[\{H\}] = p(1-p)^2,$$
$$P[\{THT\}] = P[\{T\}]\, P[\{H\}]\, P[\{T\}] = p(1-p)^2,$$
$$P[\{HTT\}] = P[\{H\}]\, P[\{T\}]\, P[\{T\}] = p(1-p)^2, \text{ và}$$
$$P[\{TTT\}] = P[\{T\}]\, P[\{T\}]\, P[\{T\}] = (1-p)^3.$$

ở đây chúng ta sử dụng việc là các phép tung là độc lập. Đặt k là số lần xuất hiện mặt ngửa trong ba lần tung, khi đó:

$$P[k = 0] = P[\{TTT\}] = (1 - p)^3,$$
$$P[k = 1] = P[\{TTH, THT, HTT\}] = 3p(1 - p)^2,$$
$$P[k = 2] = P[\{HHT, HTH, THH\}] = 3p^2(1 - p), \text{ và}$$
$$P[k = 3] = P[\{HHH\}] = p^3.$$

Kết quả trong Ví dụ 2.34 là trường hợp $n = 3$ của luật xác suất nhị thức.

ĐỊNH LÝ Giả sử k là số lần thành công trong n phép thử Bernoulli, khi đó xác suất của k được cho bởi **luật xác suất nhị thức**:

$$p_n(k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{với} \quad k = 0, \ldots, n, \qquad (2.32)$$

ở đây $p_n(k)$ là xác suất của k lần thành công trong n phép thử, và

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad (2.33)$$

là hệ số nhị thức.

Ký hiệu $n!$ trong Đẳng thức (2.33) được gọi là n giai thừa và được xác định bởi: $n! = n(n - 1) \ldots (2)(1)$. Với định nghĩa $0! = 1$.

Bây giờ chúng ta chứng minh định lý trên. Từ Ví dụ 2.34 chúng ta nhận thấy rằng mỗi một dãy với k thành công và $(n - k)$ thất bại có cùng xác suất bằng $p^k(1 - p)^{n-k}$. Đặt $N_n(k)$ là số các dãy khác nhau mà ở đó có k thành công và $(n - k)$ thất bại, khi đó :

$$p_n(k) = N_n(k)\, p^k (1 - p)^{n-k}. \qquad (2.34)$$

Biểu thức $N_n(k)$ là số cách lấy k điểm từ n điểm biểu thị cho thành công. Theo như chứng minh trong Ví dụ 2.18 ta sẽ có:

$$N_n(k) = \binom{n}{k} \qquad (2.25)$$

Định lý nhận được bởi việc thay Đẳng thức (2.35) vào Đẳng thức (2.34).

VÍ DỤ 2.35 Kiểm tra Đẳng thức (2.32) cho các xác suất tìm được trong Ví dụ 2.24.

Trong Ví dụ 2.34, đặt “kết quả tung được mặt ngửa” tương ứng với “thành công”, khi đó :

$$p_3(0) = \frac{3!}{0!\,3!} p^0 (1 - p)^3 = (1 - p)^3,$$

$$p_3(1) = \frac{3!}{1!\,2!}\, p^1(1-p)^2 = p(1-p)^2,$$

$$p_3(2) = \frac{3!}{2!\,1!}\, p^2(1-p)^1 = p^2(1-p), \text{ và}$$

$$p_3(3) = \frac{3!}{3!\,0!}\, p^3(1-p)^0 = p^3,$$

mà điều này phù hợp với kết quả trước đây.

Các bạn đã được làm quen với hệ số nhị thức trong giáo trình tính toán nhập môn khi **định lý nhị thức** được thảo luận :

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}. \tag{2.36}$$

Nếu chúng ta đặt $a = b = 1$, khi đó :

$$2^n = \sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} N_n(k),$$

mà điều này phù hợp với sự việc rằng có 2^n dãy có thể khác nhau của số lần thành công và thất bại trong n phép thử. Nếu chúng ta đặt $a = p$ và $b = 1 - p$ trong Đẳng thức (2.36), chúng ta khi đó nhận được :

$$1 = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} p_n(k),$$

mà điều này chứng tỏ rằng các xác suất của luật xác suất nhị thức có tổng bằng 1.

Giá trị của $n!$ tăng rất nhanh theo n, nên các bài toán số thường gặp phải những giá trị tương đối nhỏ của n nếu thử tính $p_n(k)$ trực tiếp theo Đẳng thức (2.32). Công thức truy hồi sau tránh khỏi ước lượng trực tiếp của $n!$ và do vậy mở rộng miền giá trị của n mà tại đó $p_n(k)$ có thể tính được trước khi gặp số khó tính :

$$p_n(k+1) = \frac{(n-k)p}{(k+1)(1-p)} p_n(k) \tag{2.37}$$

Trong phần sau của cuốn sách chúng ta sẽ trình bày hai công thức xấp xỉ xác suất nhị thức cho trường hợp n lớn.

VÍ DỤ 2.36 Đặt k là số các loa đang hoạt động (không im lặng) trong nhóm 8 loa không tương tác với nhau (tức là độc lập với nhau). Giả sử rằng mỗi loa hoạt động với xác suất 1/3. Hãy tìm xác suất để số loa hoạt động là lớn hơn 6.

Với $i = 1, \ldots, 8$ đặt A_i là biến cố "loa thứ i hoạt động". Số các loa hoạt động là số thành công trong 8 phép thử Bernoulli với $p = 1/3$. Do vậy xác suất để hơn 6 loa hoạt động là :

$$p[k=7]+p[k=8]=\binom{8}{7}\left(\frac{1}{3}\right)^7\left(\frac{2}{3}\right)+\binom{8}{8}\left(\frac{1}{3}\right)^8$$

$$= .00244 + .00015 = .00259.$$

VÍ DỤ 2.37
Mã Sửa Sai

Một hệ truyền thông thông tin nhị thức trên một kênh mà mỗi bit ngẫu nhiên bị sai với xác suất $\varepsilon = 10^{-3}$. Máy phát truyền mỗi bit thông tin 3 lần, và máy đọc quyết định lấy bit có số lần suất hiện nhiều hơn làm bit được truyền trên kênh. Hãy tìm xác suất để máy thu có được quyết định đúng đắn.

Máy thu có thể hiệu chỉnh một sai lầm riêng lẻ, như nó sẽ đưa ra quyết định sai lầm nếu kênh mắc phải 2 sai lầm trở lên. Nếu chúng ta coi mỗi phép chuyển như một phép thử Bernoulli mà ở đó mỗi "thành công" tương ứng với một lần sai, khi đó xác suất có hai lần sai trở lên trong ba phép thử Bernoulli là :

$$P[k \geq 2]=\binom{3}{2}(.001)^2(.999)+\binom{3}{3}(.001)^3 = 3(10^{-6})\,.$$

Luật Xác suất Đa thức

Luật xác suất nhị thức có thể được tổng quát hoá cho trường hợp mà ở đó chúng ta ghi được sự xảy ra của hơn một biến cố. Giả sử $B_1, B_2, \ldots B_M$ là một phân hoạch của không gian mẫu S của thí nghiệm ngẫu nhiên và đặt $P[B_j] = p_j$. Các biến cố là xung khắc, bởi vậy:

$$p_1 + p_2 + \ldots + p_M = 1.$$

Giả sử rằng n lần lặp lại thí nghiệm một cách độc lập được thực hiện. Đặt k_j là số lần xảy ra biến cố B_j, khi đó vectơ $(k_1, k_2, \ldots, k_M)$ mô tả số lần xuất hiện của mỗi biến cố B_j, xác suất của vectơ $(k_1, \ldots, k_M)$ thoả mãn **luật xác suất đa thức**:

$$P[(k_1, k_2, \ldots, k_M)] = \frac{n!}{k_1!\,k_2!\ldots k_M!}\,p_1^{k_1}\,p_2^{k_2}\ldots p_M^{k_M} \qquad (2.38)$$

ở đây $k_1 + k_2 + \ldots + k_M = n$. Luật xác suất nhị thức là trường hợp $M = 2$ của luật xác suất đa thức. Lời chứng minh cho luật xác suất đa thức hoàn toàn tương tự như cho luật xác suất nhị thức. Chúng ta chỉ cần chú ý rằng số các dãy khác nhau với $k_1, k_2, \ldots, k_M$ là số lần xuất hiện các biến cố $B_1, B_2, \ldots, B_M$ được cho bởi hệ số đa thức trong Đẳng thức (2.23).

VÍ DỤ 2.38

Một phi tiêu ném 9 lần vào bia gồm 3 miền. Mỗi lần ném có xác suất .2, .3, và .5 trúng vào miền 1, 2 và 3 tương ứng. Hãy tìm xác suất để phi tiêu ném trúng 3 lần vào mỗi miền.

Thí nghiệm này là sự lặp lại một cách độc lập 9 lần một thí nghiệm đơn có 3 kết cục có thể. Xác suất của số lần xảy ra mỗi kết cục được đưa ra bởi xác suất đa thức với các tham số n

= 9, và p_1 = .2, p_2 = .3, và p_3 = .5:

$$P[(3,3,3)] = \frac{9!}{3!\,3!\,3!}(.2)^3(.3)^3(.5)^3 = .04536\,.$$

VÍ DỤ 2.39 Giả sử rằng chúng ta lấy 10 số điện thoại một cách ngẫu nhiên từ một danh bạ điện thoại và ghi lại chữ số cuối cùng ở mỗi số điện thoại. Hỏi xác suất bằng bao nhiêu để chúng ta nhận được mỗi số nguyên từ 0 đến 9 chỉ 1 lần?

Xác suất của số lần xảy ra số nguyên được cho bởi nhị thức với các tham số M = 10, và n =10, và p_j = 1/10, nếu ta giả thiết rằng 10 số nguyên từ 0 đến 9 là đồng xác suất. Khi đó xác suất nhận được mỗi số nguyên 1 lần trong 10 lần lấy bằng:

$$\frac{10!}{1!\,1!\ldots 1!}(.1)^{10} = 3.6(10^{-4})\,.$$

Luật Xác suất Hình học

Xét một thí nghiệm liên tiếp mà ở đó chúng ta lặp lại các phép thử Bernoulli cho đến khi xảy ra thành công đầu tiên. Đặt kết cục của thí nghiệm này là *m*, là số các phép thử được thực hiện cho đến khi xuất hiện thành công đầu tiên. Không gian mẫu của thí nghiệm này là tập các số nguyên dương. Xác suất $p(m)$, để thực hiện *m* phép thử tìm được bằng sự lưu ý rằng, điều này chỉ có thể xảy ra nếu $m - 1$ phép thử đầu tiên kết quả thất bại và phép thử thứ *m* thành công[(3)]. Xác suất của biến cố này là :

$$\mathrm{p}(m) = \mathrm{P}[\mathrm{A}_1^c\mathrm{A}_2^c \ldots \mathrm{A}_{m-1}^c\mathrm{A}_m] = (1-p)^{m-1}p \qquad m = 1, 2, \ldots, \qquad (2.39)$$

ở đây A_i là biến cố "thành công ở phép thử thứ *i*". Phép gán xác suất được mô tả bởi đẳng thức (2.39) được gọi là luật xác suất hình học.

Các xác suất trong đẳng thức (2.39) có tổng bằng 1:

$$\sum_{m=1}^{\infty}\mathrm{p}(m) = p\sum_{m=1}^{\infty}q^{m-1} = p\frac{1}{1-q} = 1,$$

ở đây $q = 1 - p$, và chúng ta sử dụng công thức tính tổng của cấp số nhân. Xác suất để hơn K phép thử được thực hiện trước khi thành công có công thức đơn giản:

$$\begin{aligned}\mathrm{P}[\{m > K\}] &= p\sum_{m=K+1}^{\infty}q^{m-1} = pq^K\sum_{j=0}^{\infty}q^j \\ &= pq^K\frac{1}{1-q} \\ &= q^K \qquad\qquad (2.40)\end{aligned}$$

VÍ DỤ 2.40
Kiểm tra Lỗi bằng Phép Lặp

Máy tính A gửi một tin nhắn tới máy tính B trên kênh điện thoại không tin cậy. Tin nhắn được được mã hoá sao cho B có thể phát hiện ra một lỗi xảy ra trong quá trình truyền. Nếu B phát hiện ra một lỗi, nó sẽ yêu cầu A phát lại. Nếu xác suất xảy ra lỗi khi truyền tin nhắn là $q = .1$, xác suất tin nhắn phải truyền nhiều hơn 2 lần bằng bao nhiêu?

Mỗi phép truyền tin là một phép thử Bernoulli với xác suất thành công $p = 1 - q$. Các phép thử Bernoulli được lặp lại cho đến khi nhận được thành công đầu tiên (phép truyền không lỗi). Xác suất để có hơn hai lần truyền được thực hiện được cho bởi Đẳng thức (2.40):

$$P[m > 2] = q^2 = 10^{-2}.$$

(3) Xem Ví dụ 2.8 trong Phần 2.2 về cách giải thích bằng tần suất tương đối về việc luật xác suất hình học xảy ra như thế nào

Dãy các Thí nghiệm Độc lập

Trong phần này chúng ta xét dãy hoặc "xích" các thí nghiệm đơn mà ở đó kết cục của thí nghiệm đơn quyết định việc thực hiện tiếp. Trước hết chúng ta đưa ra một thí dụ đơn giản về một thí nghiệm như và chỉ ra sơ đồ như thế nào có thể được dùng để mô tả không gian mẫu.

VÍ DỤ 2.41

Thí nghiệm liên tiếp là phép lấy lặp lại một viên bi từ một trong hai hộp, ghi lại số trên viên bi và hoàn lại hộp của nó. Hộp 0 gồm một viên bi với số 1 và hai viên bi với số 0 và hộp 1 gồm 5 viên bi với số 1 và một viên bi với số 0. Hộp mà từ đó phép thử thứ nhất được thực hiện được chọn một cách ngẫu nhiên bởi kết quả tung đồng xu cân đối. Hộp 0 được dùng nếu kết cục là mặt ngửa và hộp 1 được dùng nếu kết cục là sấp. Sau đó hộp được dùng trong thí nghiệm đơn là hộp có số tương ứng với số trên viên bi được lấy trong thí nghiệm đơn trước đó.

Không gian mẫu của thí nghiệm này gồm các dãy 0 và 1. Mỗi dãy có thể tương ứng với một đường đi xuyên qua sơ đồ "lưới mắt cáo" được trình bày trong hình 2.17(a). Các đỉnh mang sơ đồ này ký hiệu các hộp được dùng ở thí nghiệm đơn thứ ra, và các nhân của các nhánh ký hiệu kết cục của thí nghiệm đơn. Như vậy đường đi 0011 tương ứng với dãy: phép tung đồng xu xuất hiện mặt ngửa bởi vậy lần lấy thứ nhất là từ hộp số 0; kết cục của lần lấy thứ nhất là 0, bởi vậy lần lấy bi thứ hai là từ hộp số 0; kết cục của lần lấy thứ hai là bi số 1, bởi vậy lần lấy bi thứ 3 là từ hộp số 1; và kết cục của lần lấy thứ 3 được bi số 1, bởi vậy lần lấy bi thứ tư là từ hộp số 1.

Bây giờ giả sử rằng chúng ta muốn tính xác suất của dãy riêng các kết cục, gọi là s_0, s_1, s_2. Ký hiệu xác suất này bởi $P[\{s_0\} \cap \{s_1\} \cap \{s_2\}]$. Đặt $A = \{s_2\}$ và $B = \{s_0\} \cap \{s_1\}$, khi đó $P[A \cap B] = P[A \mid B]\,P[B]$ chúng ta có

$$\begin{aligned} P[\{s_0\} \cap \{s_1\} \cap \{s_2\}] &= P[\{s_2\} \mid \{s_0\} \cap \{s_1\}]\,P[\{s_0\} \cap \{s_1\}] \\ &= P[\{s_2\} \mid \{s_0\} \cap \{s_1\}]\,P[\{s_1\} \mid \{s_0\}]\,P[\{s_0\}] \qquad (2.41) \end{aligned}$$

Bây giờ chú ý rằng trong ví dụ trên xác suất $P[\{s_n\} \mid \{s_0\} \cap \ldots \cap \{s_{n-1}\}]$ chỉ phụ thuộc vào $\{s_{n-1}\}$ do kết cục gần nhất quyết định thí nghiệm đơn nào sẽ được thực hiện:

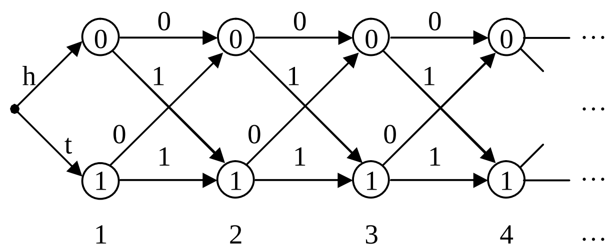
$$P[\{s_n\} \mid \{s_0\} \cap \ldots \cap \{s_{n-1}\}] = P[\{s_n\} \mid \{s_{n-1}\}] \qquad (2.42)$$

Từ đó với dãy mà ta quan tâm chúng ta có:

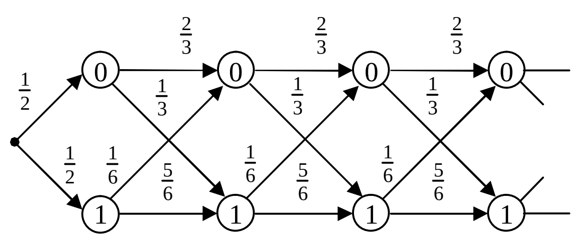
$$P[\{s_0\} \cap \{s_1\} \cap \{s_2\}] = P[\{s_2\} \mid \{s_1\}]\,P[\{s_1\} \mid \{s_0\}]\,P[\{s_0\}] \quad (2.43)$$

HÌNH 2.17

Sơ đồ hình mắt cáo của xích Markov



(a) Mỗi dãy kết cục tương ứng với một quỹ đạo xuyên qua sơ đồ hình mắt cáo này.



(b) Xác suất của dãy các kết cục là tích các xác suất dọc theo quỹ đạo kết nối.

Các thí nghiệm liên tiếp thoả mãn Đẳng thức (2.42) được gọi là **xích Markov** (**Markov chains**). Với các thí nghiệm này, xác suất của dãy $s_0, s_1, \ldots, s_n$ được cho bởi:

$$P[s_0, s_1, \ldots, s_n] = P[s_n \mid s_{n-1}]\,P[s_{n-1} \mid s_{n-2}] \ldots P[s_1 \mid s_0]\,P[s_0] \quad (2.44)$$

ở đây chúng ta có biểu thức đã được đơn giản bởi sự bỏ qua ký hiệu móc nối. Do vậy xác suất của dãy $s_0, \ldots, s_n$ được cho bởi tích xác suất của kết cục thứ nhất s_0 và các xác suất của tất cả các phép chuyển về sau, s_0 tới s_1, s_1 tới s_2, và vân vân. Chương 8 sẽ bàn về xích Markov.

VÍ DỤ 2.42 Hãy tìm xác suất của dãy 0011 với thí nghiệm hộp được giới thiệu trong Ví dụ 2.41.

Gọi lại hộp 0 gồm hai bi với nhãn 0 và một bi với nhãn 1, và hộp 1 gồm 5 bi với nhãn 1 và một bi với nhãn 0. Chúng ta hoàn toàn có thể tính được xác suất của dãy các kết cục theo nhãn các nhánh trong sơ đồ hình mắt cáo với xác suất tương ứng với phép chuyển như hình 2.17(b). Như vậy xác suất của dãy 0011 được cho bởi

$$P[0011] = P[1 \mid 1]\, P[1 \mid 0]\, P[0 \mid 0]\, P[0],$$

ở đây các xác suất chuyển được cho như sau:

$$P[1 \mid 0] = \frac{1}{3} \quad \text{và} \quad P[0 \mid 0] = \frac{2}{3}$$

$$P[1 \mid 1] = \frac{5}{6} \quad \text{và} \quad P[0 \mid 1] = \frac{1}{6},$$

và các xác suất ban đầu:

$$P(0) = \frac{1}{2} = P[1].$$

Nếu chúng ta thay thế các giá trị này vào biểu thức của

$$P[0011] = \left(\frac{5}{6}\right)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right) = \frac{5}{54}.$$

Thí nghiệm hai hộp trong các Ví dụ 2.41 và 2.42 là ví dụ đơn giản nhất của các mô hình xích Markov mà sẽ được thảo luận trong Chương 8. Thí nghiệm hai hộp thảo luận ở đây được dùng cho mô hình mà ở đó chỉ có hai kết cục, và các kết cục xảy ra đột nhiên. Ví dụ, mô hình hai hình đã được dùng cho mô hình xuất hiện đột nhiên của gói âm thanh được tạo ra bởi một loa riêng lẻ mà ở đó sự xuất hiện đột nhiên của gói âm được tách ra bởi chu kỳ yên lặng tương đối dài. Mô hình cũng được áp dụng cho dãy các chấm đen và trắng là kết quả từ việc scan các đường ảnh đen trắng.

*2.7 PHƯƠNG PHÁP TẠO RA SỰ NGẪU NHIÊN: SẢN XUẤT SỐ NGẪU NHIÊN

Phần này chúng ta sẽ giới thiệu phương pháp để sản xuất dãy số ngẫu nhiên bằng máy tính. Bất kỳ sự mô phỏng nào bằng máy tính một hệ có tính ngẫu nhiên nhất thiết phải có phương pháp tạo ra dãy số ngẫu nhiên. Các số ngẫu nhiên này cần phải thoả mãn các tính chất trung bình mà chúng mô phỏng. Trong phần này chúng ta tập trung vào bài toán tạo ra số ngẫu nhiên có “phân phối đều" trên khoảng [0, 1]. Trong chương sau chúng ta sẽ chỉ ra các số ngẫu nhiên này sẽ được dùng như thế nào để tạo ra các số tuân theo luật xác suất tuỳ ý.

Vấn đề đầu tiên chúng ta cần phải giải quyết khi tạo ra số ngẫu nhiên trong khoảng [0, 1] là việc có một số vô hạn không đếm được các điểm trong khoảng
[0, 1], nhưng sự tính toán chỉ giới hạn với các số đại diện chỉ với độ chính xác hữu hạn. Bởi vậy chúng ta cần phải bằng lòng với việc tạo ra các số đồng xác suất từ tập hữu hạn, gọi là $\{0, 1, \ldots, M-1\}$ hoặc $\{1, 2, \ldots, M\}$. Bằng việc chia các số này cho M, chúng ta nhận được các số trong khoảng đơn vị các số này có thể làm trù mật trong khoảng đơn vị bằng việc lấy M rất lớn.

Bước tiếp theo bao gồm việc tìm cơ chế sản sinh số ngẫu nhiên. Cách tiếp cận trực tiếp bao gồm cả việc thực hiện các phép thử ngẫu nhiên. Chúng ta có thể tạo ra các số nguyên từ 0 tới $2^m - 1$ bằng phép tung đồng xu cân đối m lần và thay dãy mặt ngửa và sấp bằng các số 0 và 1 để nhận được biểu diễn nhị thức của một số nguyên. Ví dụ khác gồm việc lấy một viên bi từ hộp có các viên bi được đánh số từ 1 tới M. Mô phỏng tính toán gồm việc tạo ra dãy các số ngẫu nhiên đủ dài. Nếu chúng ta sử dụng cơ chế trên để tạo các số ngẫu nhiên, chúng ta sẽ tiến hành thí nghiệm với số lần đủ lớn và chứa các kết cục vào bộ nhớ máy tính để bổ trợ thêm chương trình mô phỏng. Rõ ràng rằng sự tiếp cận này là vụng về và nhanh chóng tỏ ra không thực tế.

Cách tiếp cận thích hợp hơn để tạo ra các số ngẫu nhiên bằng máy tính là việc sử dụng các công thức truy hồi mà nó có thực hiện đầy đủ một cách dễ dàng và nhanh chóng. Chúng ta sẽ thảo luận **phương pháp thặng dư luỹ thừa** (**power residue method**), mà nó bao gồm công thức truy hồi sau:

$$Z_k = \alpha Z_{k-1} \bmod M, \qquad (2.45)$$

ở đây α là một số nguyên được chọn cẩn thận giữa 1 và M, và M là số nguyên tố p hoặc luỹ thừa nguyên của số nguyên tố p^m. Đẳng thức (2.45) gồm cả việc lấy tích α và Z_{k-1}, chia tích này cho M, lấy Z_k là phần dư của phép chia. Số nhận được nằm trong khoảng từ 0 tới $M - 1$.

VÍ DỤ 2.43 Hãy tìm các dãy được tính theo đẳng thức (2.45) với: $M = 11$, $\alpha = 7$, $Z_0 = 1$; $M = 11$, $\alpha = 3$, $Z_0 = 1$; $M = 2^2$, $\alpha = 2$, $Z_0 = 1$.

Với $M = 11$, $\alpha = 7$, và $Z_0 = 1$, chúng ta có

$$Z_1 = \text{phần dư của } \frac{(7\times 1)}{11} = 7,$$

$$Z_2 = \text{phần dư của } \frac{(7\times Z_1)}{11} = \text{phần dư của } \frac{49}{11} = 5,$$

và vân vân. Các bạn có thể kiểm tra dãy kết quả là:

1, 7, 5, 2, 3, 10, 4, 6, 9, 8, 1, 7, 5, 2, 3, 10, 4, 6, 9, 8, 1, ….

Chú ý rằng dãy chạy vòng trên qua tất cả các số nguyên nằm trong khoảng từ 1 tới 10, và sau đó dãy lặp lại vô hạn lần.

Với $M = 11$, $\alpha = 3$, và $Z_0 = 1$, dãy được tính theo đẳng thức (2.45) là:

1, 3, 9, 5, 4, 1, 3, 9, 5, 4, 1, ….

Dãy này không chạy vòng tròn qua tất cả các số nguyên trong khoảng 1 đến 10 trước khi bắt đầu quá trình lặp.

Với $M = 2^2 = 4$, $\alpha = 2$, và $Z_0 = 1$, dãy được tính theo đẳng thức (2.45) là:

1, 2, 0, 0, ….

Chú ý rằng nếu α chia đều M, khi đó dãy được tạo bởi đẳng thức (2.45) phần cuối gồm toàn số 0. Mặt khác dãy tuần hoàn có chu kỳ lớn nhất là $M - 1$. Để dãy có độ dài lớn nhất có thể α cần phải là "nghiệm nguyên thuỷ của M". Chúng ta không thảo luận nghiệm nguyên thuỷ được chọn thế nào. Chúng ta hướng độc giả quan tâm tới tài liệu tham khảo [6] ở cuối chương để đọc thêm.

Ví dụ 2.43 chỉ ra rằng các dãy được tạo bởi đẳng thức (2.45) là tuần hoàn và không thực sự ngẫu nhiên. Đẳng thức (2.43) dẫn đến rằng dãy đưa vào sẽ bắt đầu lặp lại khi có số xuất hiện lần thứ hai trong dãy. Vì lý do này mà các dãy được tạo bởi đẳng thức (2.45) được gọi là giả ngẫu nhiên.

Nếu M được lấy cực lớn, khi đó các số trong dãy không lặp lại trong suốt quá trình mô phỏng. Khi đó câu hỏi then chốt dãy số xuất hiện một cách ngẫu nhiên, có chăng và kiểm tra điều này như thế nào ? Một cách là tính các tần số tương đối. Nếu chúng ta kiểm tra dãy số chuẩn hoá bởi M, các số đơn có thể xuất hiện theo phương pháp đều trong khoảng đơn vị, các cặp số xuất hiện theo phương pháp đều trong hình vuông đơn vị, các bộ ba trong hình hộp đơn vị và vân vân. Chương 3 trình bày các tính chất thống kê để kết luận các tần suất tương đối được quan sát có phù hợp với phương pháp được mô tả hay không.

Các tính chất thống kê của các dãy được tạo bởi công thức truy hồi dạng như đẳng thức (2.45) được nghiên cứu bao quát. Chỉ dãy nhỏ có được sự đồng nhất với các tính chất thống kê tốt. Các tham số sau được giới thiệu cho phương pháp thặng dư luỹ thừa.

$$Z_i = 7^5 Z_{i-1} \bmod (2^{31} - 1). \qquad (2.46)$$

Phép sinh này với nhân tử $\alpha = 7^5 = 16{,}807$ tạo ra dãy có độ dài $M - 1 = 2^{31} - 2$ =

= 2,147,483,646. Phép chọn Z_0 được gọi là "hạt giống " của phép sinh số ngẫu nhiên, và nó quyết định điểm mà tại đó dãy bắt đầu.

Các bạn có thể không có vấn đề khi viết chương trình máy tính để thực hiện Đẳng thức (2.46). Phụ lục C chứa chương trình như vậy theo ngôn ngữ lập trình C. Thuật toán tốn ít thời gian nhất trong chương trình như vậy là phép chia cho M. Điều này dẫn đến một chú ý quan trọng rằng, khi mô hình hoán yêu cầu tạo ra số ngẫu nhiên có giá trị rất lớn. Phương pháp để thực hiện Đẳng thức (2.46) được gọi là **phép chia phỏng theo** (**simulated division**), được phát minh ra để đạt được nhanh hơn thực sự phép chia thông thường. Tài liệu tham khảo [6] và [7] giải thích phép chia mô phỏng và tài liệu tham khảo [7] trình bày chương trình viết theo ngôn ngữ FORTRAN để thực hiện phương pháp.

TÓM TẮT

- Mô hình xác suất được đồng nhất với không gian mẫu S, các biến cố được quan tâm và một phép gán xác suất ban đầu, một "luật xác suất" để từ đó có thể tính xác suất của tất cả các biến cố.
- Không gian mẫu S là tập tất cả các kết cục có thể. Nếu nó gồm một số hữu hạn hoặc đếm được các phương trình, S được gọi là rời rạc, ngược lại, gọi là liên tục.
- Các biến cố là các tập con của S thoả mãn các điều kiện mà ta quan tâm trong thí nghiệm cụ thể. Khi S rời rạc, các biến cố gồm hợp của các biến cố sơ cấp. Khi S là liên tục, các biến cố là hợp hoặc giao của các khoảng trên đường thẳng thực.
- Các tiên đề xác suất là tập các tính chất mà xác suất của các biến cố cần phải thoả mãn. Các hệ quả suy ra từ các tiên đề cung cấp các quy tắc để tính xác suất của các biến cố dựa vào xác suất của các biến cố khác.
- Phép gán xác suất ban đầu chỉ ra rằng xác suất của những biến cố chắc chắn cần phải được xác định riêng với từng mô hình. Nếu S là rời rạc, nó cần phải chỉ ra xác suất của các biến cố sơ cấp. Nếu S là liên tục thì nó phải chỉ ra xác suất của các khoảng hoặc là nửa khoảng hữu hạn.
- Các công thức tổ hợp được dùng để tính các xác suất trong các thí nghiệm mà ở đó có một số hữu hạn các biến cố đồng xác suất.
- Xác suất có điều kiện định lượng sự ảnh hưởng của thông tin riêng về kết cục của một thí nghiệm lên xác suất của một biến cố. Nó đặc biệt có ích trong các thí nghiệm liên tiếp, mà ở đó các kết cục của các thí nghiệm đơn tạo thành thông tin riêng
- Quy tắc Bayes chỉ ra xác suất hậu nghiệm của một biến cố dựa trên sự xảy ra của biến cố khác. Nó có thể được dùng để tổng hợp các quy tắc ra quyết định, mà từ đó thử xác định "nguyên nhân" có khả năng nhất dưới ánh sáng của kết quả quan trắc.
- Hai biến cố độc lập với nhau nếu thông tin về việc xảy ra của biến cố này không ảnh hưởng đến xác suất của biến cố kia. Hai thí nghiệm là độc lập nếu tất cả các biến cố của mỗi thí nghiệm độc lập với tất cả các biến cố của thí nghiệm kia. Khái niệm độc lập là rất có ích để tính xác suất trong các thí nghiệm mà nó gồm các thí nghiệm đơn là không tương tác.
- Rất nhiều thí nghiệm có thể coi như là dãy của các thí nghiệm độc lập. Trong chương này chúng ta đã trình bày luật xác suất nhị thức, đa thức và hình học như là các mô hình xuất hhiện trong điều kiện này.
- Một xích Markov là một dãy các thí nghiệm đơn mà ở đó các kết cục của thí nghiệm đơn xác định thí nghiệm đơn tiếp theo được xác định như thế nào. Xác suất của dãy các kết cục trong xích Markov được cho bởi tích của các xác suất của các thí nghiệm

thứ nhất và xác suất của tất cả các dẫn chứng.

- Sự mô phỏng máy tính dùng phương trình hồi quy để tạo ra dãy các số giả ngẫu nhiên.

DANH SÁCH CÁC THUÂT NGỮ QUAN TRONG

Quy tắc Bayes
Phép thử Bernoulli
Hệ số nhị thức
Định lý xác suất nhị thức
Xác suất có điều kiện
Không gian mẫu liên tục
Không gian mẫu rời rạc
Biến cố sơ cấp
Biến cố
Các biến cố độc lập
Các thí nghiệm độc lập
Phép gán xác suất ban đầu
Xích Markov
Phân hoạch
Luật xác suất
Không gian mẫu
Các phép toán tập hợp
Định lý xác suất toàn phần
Sơ đồ hình cây

CHÚ GIẢI TÀI LIÊU THAM KHẢO

Có nhiều sách nhập môn về xác suất và thống kê. Các tài liệu tham khảo từ [1] đến [5] là những cuốn sách yêu của tôi, chúng bắt đầu từ những vấn đề sơ khai, chọn những ví dụ trực quan, chỉ ra quy luật ở những nơi ngoài rắc rối bí hiểm và thích thú khi đọc chúng! Các tài liệu tham khảo [6] và [7] là sự mở đầu tuyệt vời các phương pháp mô phỏng các hệ ngẫu nhiên bằng máy tính.

1. Y. A. Rozanov, *Probability Theory: A Concise Course*, Dover Publications, New York, 1969.
2. P. L. Meyr, *Itroductory Proba-bility and Statistical, Applications*, Addison-Wesley, Reading, Mass., 1970.
3. K. L. Chung, *Elementary Probability Theory*, Springer-Verlag, New York, 1974.
4. A. B. Clarke and R. L. Disney, *Probability and Random Processes*, 2d ed., Wiley, New York, 1985.
5. L. Breiman, *Probability and Stochastic processes*, Houghton Mifflin, Boston, 1969.
6. H. Kobayashi, *Modeling and Analysis: An Introduction to System Performance Evanluation Methods*, Addison-Wesley, Reading, Mass., 1978.
7. A. M. Law and W. D. Kelton, *Simulation Modeling and Analisis*, McGraw-Hill, New York, 1982.
8. W. Feller, *An Introduction to Probability Theory and Its Applications,* 3d ed., Wiley, New York, 1968.

BÀI TÂP

PHẦN 2.1
Mô tả các Thí

1. Một quả xúc xắc được tung và số các chấm của mặt ngửa trên được đếm và ghi lại.

a. Không gian mẫu là gì?

b. Tập hợp A tương ứng với biến cố "số các chấm của mặt trên là số chẵn" là gì?

c. Hãy tìm tập hợp A^c và mô tả biến cố tương ứng bằng lời.

2. Một quân xúc xắc được tung hai lần và số các chấm của mặt trên được đếm và ghi lại theo thứ tự xảy ra.

   a. Tìm không gian mẫu.

   b. Tìm tập A tương ứng với biến cố "tổng số chấm xuất hiện là chẵn."

   c. Tìm tập hợp B tương ứng với biến cố "cả hai lần tung được số chấm chẵn.

   d. A kéo theo B hay B kéo theo A?

   e. Tìm $A \cap B^c$ và mô tả biến cố này bằng lời.

   f. Đặt C tương ứng với biến cố "số chấm xuất hiện ở cả hai lần tung là khác 1". Tìm $A \cap C$.

3. Hai con xúc xắc được tung và cả tổng số chấm của các mặt trên được đếm và ghi lại.

   a. Tìm không gian mẫu.

   b. Tìm tập A tương ứng với biến cố "tổng số chấm xuất hiện là chẵn".

   c. Biểu diễn mỗi biến cố sơ cấp trong thí nghiệm này như là hợp của các biến cố sơ cấp từ Bài tập 2.

4. Một con xúc xắc được tung và N_1 là số điểm của mặt trên được tính và ghi lại, N_2 là một số nguyên được lấy một cách ngẫu nhiên từ khoảng 1 tới N_1.

   a. Tìm không gian mẫu.

   b. Tìm tập các kết cục tương ứng với biến cố "quân xúc xắc xuất hiện mặt 4 chấm."

   c. Tìm tập các kết cục tương ứng với biến cố "$N_2 = 3$".

   d. Tìm tập các kết cục tương ứng với biến cố "$N_2 = 6$".

5. Một ngăn kéo bàn có 5 bút mực, ba chiếc trong đó bị khô.

   a. Các bút được lấy một cách ngẫu nhiên từng chiếc một cho đến khi lấy được bút tốt. Kết quả của dãy kiểm tra được ghi lại. Không gian mẫu là gì?

   b. Giả sử rằng số các bút, mà không phải dãy, đưa ra kiểm tra được ghi lại. Hãy mô tả không gian mẫu.

c. Giả sử rằng các bút được lấy từng chiếc một và được kiểm tra cho đến khi nhận được hai bút tốt, và kết quả của dãy được kiểm tra ghi lại. Không gian mẫu là gì?

d. Hãy mô tả không gian mẫu trong từng trường hợp c nếu chỉ số bút được kiểm tra được ghi lại.

6. Hai thành phần trong một hệ, C_1 và C_2, được kiểm tra và ghi lại một trong ba trạng thái có thể: *F*, làm việc; *R*, không làm việc nhưng có thể sửa chữa; và *K*, bị hỏng.

   a. Trong thí nghiệm này không gian mẫu là gì?

   b. Tập hợp tương ứng với biến cố "không có thành phần nào bị hỏng" là gì?

7. Ba viên bi được đánh số từ 1 đến 3 trong một hộp được lấy một cách ngẫu nhiên từng viên một cho đến khi hộp rỗng. Dãy số các viên bi được lấy được ghi lại.

   a. Tìm không gian mẫu.

   b. Tìm tập A_k tương ứng với các biến cố "viên bi số *k* được lấy vào lần thứ *k*", với $k = 1, 2, 3$.

   c. Tìm tập hợp $A_1 \cap A_2 \cap A_3$ và hãy mô tả biến cố bằng lời.

   d. Tìm tập hợp $A_1 \cup A_2 \cup A_3$ và hãy mô tả biến cố bằng lời.

   e. Tìm tập hợp $(A_1 \cup A_2 \cup A_3)^c$ và mô tả bằng lời.

8. Không gian mẫu của một thí nghiệm là đường thẳng thực. Đặt các biến cố A và B tương ứng với các tập con sau của đường thẳng thực: $A = (-\infty, r]$ và $B = (-\infty, s]$, ở đây $r \le s$. Hãy tìm biểu diễn của biến cố $C = (r, s]$ qua các biến cố A và B. Hãy chứng tỏ rằng A = $= B \cup C$ và $A \cap C = \varnothing$.

9. Trong thí nghiệm E_9 ở Ví dụ 2.3, thời gian sống của chip được đo. Đặt các biến cố A, B và C được xác định bởi: $A = (5, \infty)$, $B = (7, \infty)$, và $C = (0, 3]$. Hãy mô tả các biến cố này bằng lời. Tìm các biến cố $A \cap B$, $A \cap C$, và $A \cup B$ và mô tả các biến cố này bằng lời.

10. Dùng sơ đồ Venn để kiểm chứng sự đồng nhất của các tập hợp trong Đẳng thức (2.2) và Đẳng thức (2.4). Các bạn cần dùng các màu khác nhau hoặc các vạch khác nhau để đánh dấu các miền khác nhau cho rõ ràng.

11. Dùng sơ đồ Venn để chứng tỏ rằng:

    a. Nếu biến cố A kéo theo B, và B kéo theo C, thì A kéo theo C.

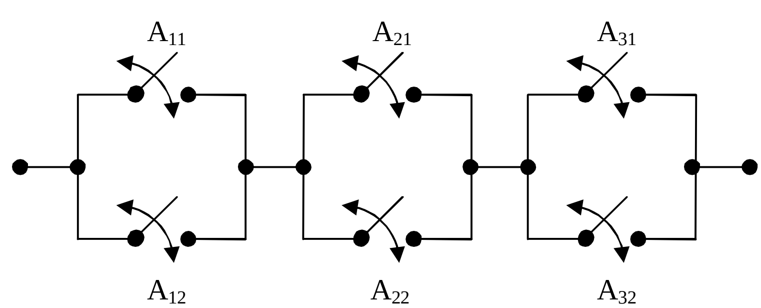
b. Nếu biến cố A kéo theo B, thì khi đó B^c kéo theo A^c.

12. Giả sử A và B là hai biến cố. Tìm biểu diễn của biến cố "đúng một biến cố trong hai biến cố A và B". Biểu diễn bằng sơ đồ Venn cho biến cố này.

13. Đặt A, B và C là các biến cố. Dùng sơ đồ Venn hãy tìm biểu diễn của các biến cố sau :
    a. Có đúng một biến cố trong ba biến cố xảy ra.
    b. Có đúng hai biến cố từ các biến cố này xảy ra.
    c. Có một hoặc hơn nữa từ các biến cố này xảy ra.
    d. Có hai hoặc hơn nữa từ các biến cố này xảy ra.
    e. Không có biến cố nào xảy ra.

14. Một hệ gồm ba con khoá giống hệt nhau. Hệ là "sẵn sàng" nếu trong mỗi hệ con có ít nhất một đơn vị làm việc.
    a. Đặt A_{jk} tương ứng là biến cố "đơn vị *k* trong hệ con *j* là làm việc", với $j = 1, 2, 3$ và $k = 1, 2$. Giải thích tại sao bài toán trên tương đương với bài toán có một sự kết nối trong mạng nối mạch được chỉ ra ở Hình P2.1.
    b. Hãy tìm mạng tương ứng với hệ "sẵn sàng" nếu cả hai đơn vị của hệ con thứ nhất làm việc và có ít nhất một đơn vị trong các hệ con còn lại làm việc.

HÌNH P2.1



15. Trong chu kỳ 24 giờ, một sinh viên thức dậy lúc T_1 và đi ngủ lúc T_2.
    a. Hãy tìm không gian mẫu và biểu diễn trên mặt phẳng tọa độ x - y nếu kết cục của thí nghiệm là cặp (T_1, T_2).
    b. Mô tả tập A và biểu diễn miền trên mặt phẳng tương ứng với biến cố "sinh viên dậy vào lúc 9 giờ".
    c. Hãy mô tả tập hợp B và biểu diễn miền trên mặt phẳng tương

ứng với biến cố "sinh viên ngủ nhiều hơn là thức".

d. Biểu diễn miền tương ứng với $A^c \cap B$ và diễn tả biến cố tương ứng bằng lời.

PHẦN 2.2
Các Tiên đề Xác suất

16. Một con xúc xắc được tung và số điểm của mặt hướng lên trên được ghi lại.

   a. Hãy tìm xác suất của các biến cố sơ cấp với giả thiết rằng tất cả các mặt của con xúc xắc là đồng xác suất hướng lên trên sau khi tung.

   b. Hãy tìm xác suất của các biến cố sơ cấp với giả thiết rằng mặt một chấm có xác suất hướng lên trên gấp đôi các mặt còn lại.

   c. Tìm xác suất để kết cục của phép tung là số chấm chẵn với các giả thiết như trong phần a và b của Bài tập này.

17. Con xúc xắc được tung hai lần và số các điểm của mặt hướng lên trên được ghi lại theo thứ tự xảy ra. Giả sử tất cả các kết cục đồng khả năng xảy ra, tìm xác suất của các biến cố sau :

   a. A_k : tổng của hai kết cục bằng k, với $k = 2, \ldots, 12$.

   b. B : các kết cục của hai lần tung là khác nhau.

18. Một thí nghiệm ngẫu nhiên có không gian mẫu S = {a, b, c}. Giả sử rằng P[{a, c}] = 5/8 và P[{b, c}] = 7/8. Hãy dùng các tiên đề xác suất để tìm xác suất của các biến cố sơ cấp.

19. Chứng tỏ rằng để có đúng một biến cố từ hai biến cố A và B xảy ra được cho bởi $P[A] + P[B] - 2P[A \cap B]$.

20. Giả sử các biến cố A và B có xác suất sau $P[A] = x$, $P[B] = y$ và $P[A \cap B] = z$. Dùng sơ đồ Venn để tìm $P[A^c \cup B^c]$, $P[A \cap B^c]$, $P[A^c \cup B]$, và $P[A^c \cap B^c]$.

21. Chứng minh rằng :

$$P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - P[B \cap C] + P[A \cap B \cap C].$$

22. Dùng các biến từ Bài tập 21 để chứng minh hệ quả 6 bằng phương pháp quy nạp.

23. Một đồng xu cân đối được tung 4 lần, đặt A_i là biến cố "kết quả lần tung thứ i là ngửa". hãy tìm xác suất của các biến cố sau: A_2, $A_1 \cap A_3$, $A_1 \cap A_2 \cap A_3 \cap A_4$, và $A_1 \cup A_2 \cup A_3 \cup A_4$.

24. Một đồng xu cân đối được tung cho đến khi mặt ngửa lần đầu tiên xuất hiện. Đặt k mà số lần tung ở đó mặt ngửa lần đầu tiên xuất hiện. Đặt A là biến cố "$k > 5$", và B là biến cố "$k > 10$". Hãy tìm

xác suất của các biến cố A, B, B^c, $A \cap B$, và $A \cup B$.

25. Hãy dùng Hệ quả 7 để chứng minh các hệ thức sau:

    a. $P[A \cup B \cup C] \le P[A] + P[B] + P[C]$.

    b. $P\left[\bigcup_{k=1}^{n} A_k\right] \le \sum_{k=1}^{n} P[A_k]$.

    Biểu thức thứ hai được gọi là **bao hợp** [**union bound**].

26. Một tài liệu gồm *n* ký tự được ghi vào trong máy tính. Xác suất để một tín hiệu đơn bị ghi sai là *p*. Dãy dùng bao hợp để nhận được cận trên của xác suất ghi sai ký tự bất kỳ trong tài liệu được ghi.

27. Một số *x* được lấy ngẫu nhiên trong khoảng [–1, 1]. Giả sử các biến cố $A = \{x < 0\}$, $B = \{|x - 0.5| < 1\}$, và $C = \{x > 0.75\}$.

    a. Tìm xác suất của các biến cố B, $A \cap B$ và $A \cap C$.

    b. Tìm xác suất của các biến cố $A \cup B$, $A \cup C$, và $A \cup B \cup C$, trước hết bằng cách tính trực tiếp các tập hợp và xác suất của chúng, và thứ hai, bằng cách dùng các Tiên đề và Hệ quả thích hợp.

28. Chọn số được lấy ngẫu nhiên trong khoảng [–1, 1]. Các số từ khoảng con [0, 1] xảy ra hai lần nhiều hơn từ khoảng [–1, 0).

    a. Tìm phép gán xác suất cho toàn khoảng [–1, 0), toàn khoảng [0,1], và từng phần trong mỗi khoảng trên.

    b. Lặp lại Bài tập 27 với phép gán xác suất này.

29. Thời gian sống của một thiết bị phù hợp với luật xác suất mũ được giới thiệu ở Ví dụ 2.10, với $\alpha = 1$. Đặt A là biến cố "thời gian sống lớn hơn 5", và B là biến cố "thời gian sống lớn hơn 10".

    a. Tìm xác suất của các biến cố $A \cap B$ và $A \cup B$.

    b. Hãy tìm xác suất của biến cố "thời gian sống lớn hơn 5 nhưng nhỏ hơn hoặc bằng 10".

30. Xét một thí nghiệm mà không gian mẫu là đường thẳng thực. Luật xác suất gán các xác suất cho các tập con có dạng $(-\infty, r]$.

    a. Chứng minh rằng phải có: $P[(-\infty, r]] \le P[(-\infty, s]]$ khi $r < s$.

    b. Hãy tìm biểu diễn của $P[(r, s]]$ qua $P[(-\infty, r]]$ và $P[(-\infty, s]]$.

31. Hai số được chọn ngẫu nhiên từ khoảng [0, 1]. Hãy tìm xác suất để chúng khác nhau một lượng lớn hơn 1/2.

PHẦN *2.3
Tính Xác suất

32. Một từ khoá được tạo bởi ba số từ tập hợp {0, 1, 2, …, 59}. Hãy tìm số các tập hợp có thể.

bằng các
Phương pháp
Tính

33. Một con xúc xắc sáu mặt được tung, một đồng xu được gieo, và một quân bài được lấy ngẫu nhiên từ bộ bài 52 quân. Hãy tìm số các kết cục có thể.
34. Có một sinh viên có bốn đôi giày khác nhau và không bao giờ mang cùng một đôi trong hai ngày liên tiếp. Có bao nhiêu cách anh ta có thể mang giày trong 5 ngày?
35. Có bao nhiêu số điện thoại gồm 7 chữ số có thể, nếu chữ số khác 0 và 1.
36. Sắp thứ tự các bánh Pizza hảo hạng các bạn có bốn sự lựa chọn từ 15 loại thượng hạng sẵn có. Hỏi có bao nhiêu tổ hợp có thể nếu các loại thượng hạng có thể lặp lại? Nếu chúng không thể lặp lại?
37. Hỏi có bao nhiêu cách có thể để 10 sinh viên ngồi vào 10 chiếc bàn? 12 chiếc bàn?
38. Một đứa bé kéo ba tập của cuốn bách khoa toàn thư từ giá sách và sau khi bị quở mắng, đã đặt chúng trở lại giá sách theo thứ tự ngẫu nhiên. Xác suất bằng bao nhiêu để các cuốn sách được đặt đúng thứ tự?
39. Một tập bài gồm 10 quân bài đỏ được đánh số từ 1 đến 10, và 10 quân bài đen được đánh số từ 1 đến 10. Hỏi có bao nhiêu cách để xếp 20 quân bài thành một hàng? Giả sử rằng chúng ta lấy các quân bài một cách ngẫu nhiên và đặt chúng thành hàng. Xác suất bằng bao nhiêu để các quân bài đỏ và đen xen kẽ nhau?
40. Một quầy thức ăn nhanh cung cấp hành tây, hương liệu, mù tạc, ketchup và ớt cay cho món "hot dog" của bạn. Có bao nhiêu tổ hợp có thể khi dùng một gia vị? Hai gia vị? Không, một vài hoặc tất cả các gia vị?
41. Một lô 100 sản phẩm có k phế phẩm. M sản phẩm được chọn ngẫu nhiên và kiểm tra. Xác suất bằng bao nhiêu để M sản phẩm đó có m phế phẩm?
42. Một khu rừng có N gấu trúc, trong đó có 10 con trước đây đã bị bắt và được gán nhãn. Giả sử rằng 20 đã bị bắt. Hãy tìm xác suất để trong đó có 5 con đã được gán nhãn. Ký hiệu xác suất này là $p(N)$. Hãy tìm giá trị N để xác suất này đạt giá trị lớn nhất. *Gợi ý:* so sánh tỉ số $p(N)/p(N-1)$ với 1.
43. Bạn thắng xổ số nếu bạn dự đoán đúng số của 6 quả bóng được lấy từ hộp gồm các quả bóng được đánh số từ 1 đến 49, với phép lấy không hoàn lại và không để ý đến thứ tự. Xác suất bằng bao nhiêu nếu bạn mua một vé?

44. Hỏi rằng có bao nhiêu hoán vị khác nhau của tập gồm 4 viên bi đỏ và 2 viên bi trắng, và 3 viên bi đen?

45. Hãy tìm xác suất để tổng các kết cục của ba lần tung một con xúc xắc là 7.

46. Chứng minh rằng

$$\binom{n}{k} = \binom{n}{n-k}.$$

47. Trong Bài tập này chúng ta nhận được hệ số đa thức. Giả sử chúng ta phân hoạch một tập gồm n phương trình khác nhau thành F tập con B_1, B_2, …, B_F cỡ k_1, …, k_F, một cách tương ứng, ở đây $k_i > 0$, và $k_1 + k_2 + \ldots + k_F = n$.

    a. Giả sử N_i là số các kết cục có thể khi tập con thứ i được chọn. Chứng minh rằng:

    $$N_1 = \binom{n}{k_1},\ N_2 = \binom{n-k_1}{k_2},\ \ldots,\ N_{\Im-1} = \binom{n-k_1-\mathrm{L}-k_{\Im-2}}{k_{\Im-1}}.$$

    b. Chứng minh rằng khi đó số các phân hoạch:

    $$N_1 N_2 \ldots N_{\Im-1} = \frac{n!}{k_1!k_2!\ldots k_\Im!}.$$

PHẦN 2.4
Xác suất có Điều kiện

48. a. Hãy tìm P[A | B] nếu A ∩ B = ∅, nếu A ⊂ B, và nếu B ⊂ A.

    b. Chứng minh rằng nếu P[A | B] > P[A], thì P[B | A] > P[B]. Hãy cho nhận xét.

49. Hãy chứng minh P[A | B] thỏa mãn các Tiên đề xác suất:

    i. 0 ≤ P[A | B] ≤ 1.

    ii. P[S | B] = 1.

    iii. Nếu A ∩ C = ∅, khi đó P[A ∪ C | B] = P[A | B] + P[C | B].

50. Chứng minh rằng P[A ∩ B ∩ C] = P[A | B ∩ C] P[B | C] P[C].

51. Một con xúc xắc được tung hai lần và số các điểm của mặt trên được tính và được ghi lại theo thứ tự xảy ra. Đặt biến cố A là biến cố "tổng các điểm là số chẵn", và đặt B là biến cố "cả hai lần tung có số điểm chẵn". Hãy tìm P[A | B] và P[B | A].

52. Số x được lấy ngẫu nhiên trong khoảng [–1, 1]. Đặt B là biến cố $\{|x - 1/2| < 1\}$ và đặt C là biến cố $\{x > 3/4\}$. Hãy tìm P[B | C] và P[C | B].

53. Trong mỗi lô 100 sản phẩm có 2 sản phẩm được kiểm tra, và lô bị

loại nếu cả hai sản phẩm được kiểm tra là phế phẩm.

a. Hãy tìm xác suất lô có 5 phế phẩm. lặp lại với lô có 10 phế phẩm.

b. Tính lại xác suất trong phần a nếu 3 sản phẩm được kiểm tra và lô được chấp nhận khi có nhiều nhất 1 phế phẩm trong 3 sản phẩm được kiểm tra.

54. Hãy tìm xác suất để 2 sinh viên hoặc hơn nữa trong lớp có 20 sinh viên có cùng ngày sinh. *Gợi ý:* Dùng Hệ quả 1.

55. Hãy dùng xác suất có điều kiện và sơ đồ hình cây để tìm xác suất của các biến cố sơ cấp trong các thí nghiệm ngẫu nhiên được xác định trong các phần a tới d của Bài tập 5.

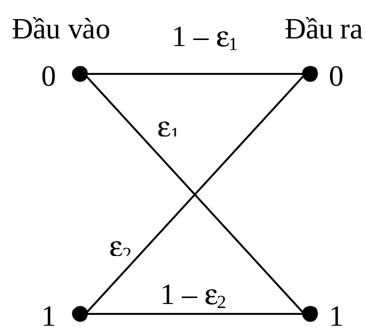
56. Thời gian đến nơi làm việc của một giáo sư có phân phối đều trong khoảng từ 8 đến 9 giờ sáng. Hãy tìm xác suất để giáo sư đến nơi làm việc trong khoảng một phút, biết rằng giáo sư không đến nơi làm việc trước 8 giờ 30 phút. Lặp lại với 8 giờ 50 phút. Hãy giải thích các kết quả.

57. Kênh truyền thông nhị thức không đối xứng như được chỉ ra trong Hình P2.2. Giả sử rằng các tín hiệu vào đồng xác suất.

a. Hãy tìm xác suất để tín hiệu ra là 0.

b. Hãy tìm xác suất để tín hiệu vào là 0 khi biết rằng tín hiệu ra là 1. Hãy tìm xác suất để tín hiệu vào là 1 biết rằng tín hiệu ra là 1. Tín hiệu vào nào có xác suất lớn hơn?

HÌNH P2.2



58. Một con xúc xắc được tung và số các điểm là N_1 được ghi lại; khi đó số nguyên N_2 được lấy một cách ngẫu nhiên từ $\{1, \ldots, N_1\}$.

a. Dùng sơ đồ hình cây để mô tả không gian mẫu.

b. Hãy tìm xác suất của biến cố $\{N_2 = 3\}$.

c. Hãy tìm xác suất của biến cố $\{N_1 = 4\}$ biết $\{N_2 = 3\}$.

d. Hãy tìm xác suất của biến cố $\{N_1 = 4\}$ biết $\{N_2 = 5\}$.

59. Một trong hai đồng xu được lấy một cách ngẫu nhiên và được tung. Đồng xu thứ nhất có xác suất hiện mặt ngửa là p_1 và đồng xu thứ hai có xác suất xuất hiện mặt ngửa là p_2.

    a. Xác suất xuất hiện kết cục của phép tung là mặt ngửa bằng bao nhiêu?

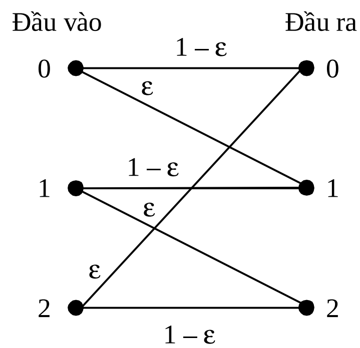
    b. Xác suất sử dụng đồng xu thứ 2 bằng bao nhiêu khi biết rằng mặt ngửa đã xảy ra.

60. Một kênh truyền thông tạm phân được chỉ ra ở Hình P2.3. Giả sử rằng các tín hiệu vào là 0, 1, và 2 xảy ra với xác suất 1/2, 1/4, và 1/4 một cách tương ứng.

    a. Hãy tìm xác suất của các tín hiệu ra.

    b. Giả sử rằng tín hiệu ra quan sát được là 1. Xác suất để tín hiệu vào là 1, 2, 3 bằng bao nhiêu?

HÌNH P2.3



61. Một nhà sản xuất máy tính sử dụng các chip từ 3 nguồn. Các chip từ nguồn A, B và C là phế phẩm với xác suất .001, .005, và .01 một cách tương ứng. Nếu chip được lấy ngẫu nhiên và là phế phẩm, hãy tìm xác suất để chip đó là của nhà sản xuất A; từ nhà sản xuất C.

PHẦN 2.5
Sự Độc lập của các Biến cố

62. Chứng minh rằng, nếu A và B là các biến cố độc lập thì khi đó các cặp A và B^c; A^c và B, và A^c và B^c cũng độc lập.

63. Chứng minh rằng, các biến cố A và B độc lập nếu P[A | B] = = P[A | B^c].

64. Giả sử A và B là những biến cố có xác suất P[A] và P[B].

    a. Hãy tìm xác suất p[A ∪ B] nếu A và B độc lập.

    b. Hãy tìm xác suất P[A ∪ B] nếu A và B xung khắc.

65. Một thí nghiệm gồm phép lấy một hộp một cách ngẫu nhiên và lấy một viên bi từ hộp và ghi lại màu của nó (đen hoặc trắng). Đặt A là biến cố “hộp 1 được chọn” và B là biến cố “bi màu đen được lấy”. Với điều kiện nào thì A và B sẽ độc lập?

66. Hãy tìm xác suất trong Bài tập 13 giả thiết rằng A, B và C độc lập?

67. Hãy tìm xác suất để hệ là “sẵn sàng” trong Bài tập 14 phần a và b. Giả thiết là các đơn vị trong hệ thống bị hỏng một cách độc lập với nhau và đơn vị thứ *k* bị hỏng với xác suất p_k.

68. Có một thí nghiệm được lặp lại một số lớn lần, và sự xuất hiện biến cố A và B được ghi lại. Bạn có thể kiểm tra sự độc lập của các biến cố A và B như thế nào?

69. Hãy tính xác suất để hệ trong Ví dụ 2.32 là ‘sẵn sàng” khi bộ điều kiển thứ hai được gắn vào hệ thống.

70. Trong một hệ truyền thông nhị phân trong Ví dụ 2.23, hãy tìm giá trị của ε để với giá trị này tín hiệu vào của kênh là độc lập với tín hiệu ra của kênh. Có thể sử dụng một kênh như vậy để truyền thông tin hay không?

PHẦN 2.6

Dãy Thí nghiệm Liên tiếp

71. Một block gồm 100 bit được truyền kênh trên nhị thức với xác suất bit sai là $p = 10^{-3}$. Hãy tìm xác suất để block gồm 3 hoặc hơn nữa số bit sai.

72. Mười phần trăm sản phẩm của một dây chuyền sản xuất là phế phẩm. Xác suất để có hơn một phế phẩm trong lô gồm *n* sản phẩm bằng bao nhiêu?

73. Một sinh viên cần 10 chip dạng nào đó để làm một mạch điện. Biết rằng 5% các chip dạng này là phế phẩm. Anh ta cần phải mua bao nhiêu chip để với xác suất lớn hơn 90% có đủ chip để làm mạch điện?

74. Đặt *k* là số các loa hoạt động trong nhóm *n* loa không tương tác (tức là độc lập với nhau). Hãy viết một chương trình để tính P[*k*] với *k* = 0, 1, …, *n* và với *n* = 8, 24 và 48. Dùng kết quả của bạn để tính lại Hình 1.6 trong Chương 1.

75. Một hệ gồm 10 chip. Thời gian sống của mỗi chip tuân theo phân phối mũ với tham số. Hãy tìm xác suất để ít nhất một nửa số bit làm việc sau 1/α giây.

76. Một máy mắc lỗi trong thao tác nào đó với xác suất p. Có hai dạng lỗi. Tỉ lệ của loại I là a, và tỉ lệ của lỗi loại II là $1 - a$.

    a. Xác suất để mắc k lỗi trong n thao tác bằng bao nhiêu?

    b. Xác suất để mắc k_1 lỗi loại I trong n thao tác bằng bao nhiêu?

    c. Xác suất để mắc k_2 lỗi loại II trong n thao tác bằng bao nhiêu?

    d. Xác suất để xảy ra đồng thời k_1 lỗi loại I, k_2 lỗi loại II trong n thao tác bằng bao nhiêu?

77. Có ba dạng điện báo đến tại một trung tâm điện báo. 10% là điện báo "ưu tiên cao", 40% "ưu tiên trung bình", và 50% là "ưu tiên thấp".

    a. Hãy tìm xác suất để k điện báo trong N điện báo không phải là ưu tiên cao.

    b. Giả sử rằng mỗi bức điện gửi đến một lần. Hãy tìm xác suất để có k điện báo ưu tiên cao đến.

    c. Hãy tìm xác suất để trong 20 điện báo có 5 ưu tiên cao, 10 ưu tiên thường và 5 ưu tiên thấp.

78. Một bộ lập mã một dãy thông tin nhị phân thành một dãy gồm một đoạn k số "0" được kết thúc bởi số "1", với $k = 0, \ldots, m - 1$, hoặc một dòng gồm m số "0". Trường hợp $m = 3$ là:

dòng mã	*ký hiệu số k*
1	0
01	1
001	2
000	3

    Giả sử rằng dãy thông tin được tạo bởi dãy các phép thử Bernoulli độc lập với P["1"] = P[thành công] = p.

    a. Hãy tìm xác suất để có k ký tự 0 trong trường hợp $m = 3$.

    b. Hãy tìm xác suất để có k ký tự 0 trong trường hợp m tổng quát.

79. Thời gian ô tô đỗ ở bãi xe tuân theo luật phân phối mũ với tham số là 1. Tiền phải trả để đỗ ở bãi trong thời gian nửa tiếng trở lại là \$1.

    a. Hãy tìm xác suất để ô tô phải trả \$$k$.

    b. Giả sử rằng số tiền phải trả tối đa là \$5. Hãy tìm xác suất để ô tô phải trả \$$k$.

80. Một đồng xu không đối xứng được tung lặp lại cho đến khi mặt ngửa xuất hiện hai lần. Hãy tìm xác suất để k lần được thực hiện. *Gợi ý:* Chứng minh rằng {"k lần tung được thực hiện"} = A $\cap$ B, ở

đây A = {“lần tung thứ k được mặt ngửa”} và B = {“trong k – 1 lần tung đầu tiên xuất hiện mặt ngửa 1 lần”}.

81. Một hộp ban đầu chứa hai viên bi đen và hai viên bi trắng. Thí nghiệm sau được lặp lại vô hạn lần, một viên bi được lấy từ hộp; nếu bi màu trắng thì nó được đặt trả lại hộp nếu khác nó được đặt ra ngoài.

   a. Hãy vẽ sơ đồ tam phân cho thí nghiệm này và gắn nhãn các cành bằng xác suất chuyển. *Gợi ý:* Hộp chỉ có thể ở 1 trong 3 trạng thái.

   b. Hãy tìm xác suất của các kết cục : *www*, *bww*, *bbw* và *bbwww*.

   c. Hãy tìm xác suất để hộp không còn bi đen sau 3 lần lấy.

   d. Hãy tìm xác suất để hộp còn 2 bi đen sau n phép thử.

82. Trong Ví dụ 2.42, đặt $p_0(n)$ và $p_1(n)$ là xác suất để hộp 0 hay hộp 1 được sử dụng trong thí nghiệm đơn thứ n.

   a. Hãy tìm $p_0(1)$ và $p_1(1)$.

   b. Biểu diễn $p_0(n + 1)$ và $p_1(n + 1)$ qua $p_0(n)$ và $p_1(n)$.

   c Hãy tính $p_0(n)$ và $p_1(n)$ với n = 2, 3, 4.

   d. Hãy tìm nghiệm cho phép đệ quy trong phần b với điều kiện ban đầu được cho trong phần a.

   e. Các xác suất trên bằng bao nhiêu khi $n \to \infty$?

PHẦN 2.7

Phương pháp Tạo giả Số Ngẫu nhiên bằng Máy tính: Sự Sản sinh Số Ngẫu nhiên

83. Thí nghiệm hộp được sử dụng để mô phỏng thí nghiệm ngẫu nhiên với không gian mẫu S = {1, 2, 3 4, 5} và các xác suất $p_1 = 1/3$, $p_2 = 1/5$, $p_3 = 1/4$, $p_4 = 1/7$ và $p_5 = 1 - (p_1 + p_2 + p_3 + p_4)$. Hộp có thể gồm bao nhiêu viên bi? Tổng quát hoá chứng tỏ rằng thí nghiệm hộp có thể được dùng để mô tả thí nghiệm ngẫu nhiên bất kỳ với không gian mẫu hữu hạn và với xác suất được cho bởi các số tỉ lệ.

84. Giả sử chúng ta quan tâm tới việc sử dụng phép tung đồng xu cân đối để mô phỏng thí nghiệm ngẫu nhiên mà ở đó có 6 kết cục đồng khả năng, ở đây S = {0, 1, 2, 3, 4, 5}. Thuật toán sau được đưa ra:

   1. Tung đồng xu cân đối ba lần và nhận được số nhịi thức bằng việc đồng nhất mặt sấp với 0 và mặt ngửa với 1.

   2. Nếu kết cục của phép tung đồng xu trong bước 1 là biểu diễn nhị thức của một số trong S, số được lấy ra nếu khác quay lại bước 1.

   Thuật toán này là sự đơn giản hoá “phương pháp loại bỏ” được thảo luận trong Phần 3.12.

a. Hãy tìm xác suất để một số được lấy ra ở bước 2.

b. Chứng minh rằng các số được lấy ra ở bước 2 là đồng xác suất.

c. Tổng quát hoá thuật toán trên để chứng tỏ rằng phép tung đồng xu có thể được sử dụng như thế nào để mô phỏng thí nghiệm hộp ngẫu nhiên bất kỳ.

85. Hãy viết chương trình máy tính để thực hiện việc tạo ra số ngẫu nhiên được mô tả bởi Phương trình (2.32).

    a. Để kiểm tra chương trình của bạn, hãy tìm Z_{1000}; với hạt giống ban đầu $Z_0 = 1$, nó sẽ phải bằng 522,329,230.

    b. Hãy tạo ra 10,000 số ngẫu nhiên trong khoảng đơn vị và cắt các số này đến 1 chữ số sau dấu phảy. Hãy tính biểu đồ thu gọn tần số cho các số đơn, bộ đôi các số và bộ ba các số. Chú ý rằng mỗi số sẽ tham gia tích cực vào tích của hai bộ đôi, một lần như là thành phần thứ nhất, một lần như là thành phần thứ hai, và tương tự như đối với bộ ba. Các biểu đồ có phù hợp với các kết quả được mong đợi?

86. Giả sử bạn có một chương trình cho bạn các số U_n có phân phối đều trong khoảng [0, 1]. Đặt $Y_n = \alpha U_n + \beta$. Hãy tìm α và β sao cho Y_n có phân phối đều trong khoảng $[a, b]$. Lấy $a = -5$ và $b = 15$. Hãy viết chương trình máy tính để tạo ra Y_n và hãy tính trung bình mẫu và phương sai mẫu trong trường hợp $n = 1000$. Hãy so sánh trung bình mẫu và phương sai mẫu với $(a + b)/2$ và $(b - a)^2/12$ một cách tương ứng.

87. Giả sử bạn có một chương trình cho bạn các số U_n có phân phối đều trong khoảng [0, 1].

    a. Giả sử chúng ta tạo ra dãy số nhị thức B_n theo cách như sau: Nếu $U_n \le 1/2$, khi đó $B_n = 0$; nếu khác $B_n = 1$. Bằng cách nào để dãy B_n là sự mô phỏng phép tung đồng xu cân đối?

    b. Bạn có thể phỏng theo cách làm ở phần a để mô phỏng dãy phép thử Bernoulli?

    c. Bạn có thể phỏng theo cách làm ở phần a để mô phỏng phép tung con xúc xắc cân đối?

    d. Bạn có thể phỏng theo cách làm ở phần a để mô phỏng một thí nghiệm ngẫu nhiên *bất kỳ* có hữu hạn kết cục? Có thể có phương pháp tạo ra một số vô hạn đếm được các kết cục hay không?

88. Hãy viết phương trình mô phỏng thí nghiệm hộp được thảo luận

trong Phần 1.3. Hãy tính tần suất tương đối của các kết cục trong 1000 lần lấy từ hộp.

89. Hày tìm phương pháp sử dụng các số ngẫu nhiên có phân phối đều trong khoảng [0, 1] để tạo ra dãy các số nguyên tuân theo luật xác suất hình học. Hãy viết chương trình máy tính để mô phỏng thí nghiệm tung đồng xu được thảo luận trong Ví dụ 2.8. Lặp lại thí nghiệm 100 lần để và tính sơ đồ hình cây được chỉ ra ở Hình 2.6.

90. Hãy tìm phương pháp sử dụng dãy các số ngẫu nhiên U_n có phân phối đều trong khoảng [0, 1] để tạo ra dãy các số nguyên luật xác suất nhị thức. Phỏng theo phương pháp này để tạo ra vectơ (k_1, …, k_M) tuân theo luật xác suất đa thức.

CÁC BÀI TOÁN YÊU CẦU KIẾN THỨC TỔNG HỢP

91. Có hai người, mỗi người tung đồng xu cân đối 3 lần, xác suất để họ nhận được cùng một số lần xuất hiện mặt ngửa bằng bao nhiêu?

92. Giả sử rằng trong Ví dụ 2.40, máy tính A gửi từng gói tin tới máy tính B cùng một lúc trên hai kênh điện thoại không tin cậy. Máy tính B có thể phát hiện khi có lỗi trên cả hai đường truyền. Giả sử xác suất gói tin bị truyền sai trên kênh 1 và 2 là q_1 và q_2 một cách tương ứng. Máy tính B đề nghị phát lại cho đến khi nó nhận được gói tin không lỗi trên cả hai đường truyền.

    a. Hãy tìm xác suất để có *k* lần truyền được yêu cầu

    b. Hãy tìm xác suất để trong lần truyền cuối cùng, gói tin trên kênh 2 được nhận không lỗi.

93. Để một bảng mạch điện làm việc, cần phải có 7 chip giống nhau làm việc. Để tăng độ tin cậy ta gắn thêm một chip vào bảng điện và thiết kế cho phép nó có thể thay thế 1 trong 7 chip bất kỳ, khi nó bị hỏng.

    a. Hãy tìm xác suất P_b để bảng điện làm việc, biết rằng xác suất để từng chip riêng lẻ làm việc là *p*.

    b. Giả sử rằng *n* bảng mạch điện được mắc song song, và chúng ta yêu cầu với xác suất 99.9% có ít nhất 1 bảng mạch điện làm việc. Cần thiết phải có bao nhiêu bảng mạch điện?

94. Một từ hai đồng xu được lấy ra một cách ngẫu nhiên và được tung ba lần. Biết rằng các đồng xu có xác suất xuất hiện mặt ngửa bằng p_1 và p_2 một cách tương ứng, với $p_1 > p_2$.

    a. Hãy tìm xác suất để đồng xu 1 được tung, biết rằng số lần xuất hiện mặt ngửa bằng *k*, với *k* = 0, 1, 2, 3.

    b. Trong phần a, đồng xu nào có xác suất lớn hơn khi có *k* lần xuất hiện mặt ngửa?

c. Tổng quát hoá lời giải trong phần b cho trường hợp đồng xu được chọn được tung *m* lần. Trong thực tế hãy tìm giá trị ngưỡng *T* sao cho khi $k > T$ (số lần xuất hiện mặt ngửa), đồng xu 1 có xác suất được chọn lớn hơn, và khi $k \le T$, đồng xu 2 có xác suất được chọn là lớn hơn.

95. Giả sử rằng một bộ bài gồm 52 quân khác nhau được xáo trộn kỹ, trong đó có 4 quân Át và 4 quân K.

a. Hãy tìm xác suất để nhận được một quân Át ở lần rút đầu tiên.

b. Hãy rút một quân bài từ bộ bài và quan sát quân bài lấy được. Xác suất nhận được một quân bài Át ở lần lấy thứ hai bằng bao nhiêu? Câu trả lời có thay đổi không, nếu bạn đã không quan sát lần rút thứ nhất ?

c. Giả sử ta rút 7 quân bài từ bộ bài. Xác suất có 3 quân Át trong 7 quân bài bằng bao nhiêu? Xác suất để trong 7 quân bài có 2 quân K bằng bao nhiêu? Xác suất để trong 7 quân bài có 3 quân Át và (hoặc) 2 quân K bằng bao nhiêu?

d. Giả sử rằng toàn bộ bộ bài được chia đều cho 4 người chơi. Xác suất bằng bao nhiêu để mỗi người chơi nhận được một quân Át?