CHƯƠNG 3

XÁC SUẤT VÀ NHỮNG PHÂN PHỐI XÁC SUẤT RỜI RẠC

*V*ề chương này:

Mục tiêu của chương này là đặt nền tảng cho sự suy luận thống kê. Chúng ta sẽ giới thiệu một số khái niệm cơ bản về xác suất, và sau đó sẽ trình bày một số mô hình để tính toán xác suất của các kết quả của mẫu.

NGHIÊN CỨU TÌNH HUỐNG

NHỮNG NGƯỜI SINH TRONG THỜI KỲ BÙNG NỔ SINH ĐẺ SẼ TỪ BỎ CÁC CỬA HÀNG BÁCH HÓA?

Có phải hầu hết việc mua sắm của anh/chị được thực hiện ở các cửa hàng bách hóa, hay có phải anh/chị tiêu xài những đô la mua sắm của mình tại một số các cửa hàng đặc sản chọn lọc? Nghiên cứu cho thấy rằng những người trung niên là những khách hàng tốt nhất của các cửa hàng bách hóa. Các nhà quản lý cửa hàng bách hóa đang hy vọng doanh thu bán hàng chậm sẽ được đẩy nhanh khi thế hệ bùng nổ sinh đẻ bước vào tuổi trung niên. Trong bài báo nhan đề "Có phải những người sinh trong thời kỳ bùng nổ sinh đẻ sẽ từ bỏ các cửa hàng bách hóa?" (Schwartz, 1990), tác giả xem xét vấn đề này dưới ánh sáng của một nghiên cứu của Marvin J. Rothenberg, Inc., một công ty tư vấn đặt trụ sở ở Fair Lawn, New Jersey. Các nhà nghiên cứu tại công ty Rothenberg đã phỏng vấn 60.000 khách hàng thường xuyên của các cửa hàng bách hóa thân chủ của họ trong giai đoạn từ 1988 đến 1990. Những cửa hàng này được sắp xếp theo quy mô từ nhỏ đến lớn và tọa lạc khắp các bang "từ New York ở miền Đông sang California ở miền Tây, và từ Wisconsin ở miền Bắc đến Louisiana ở Miền Nam." Theo báo cáo của Rothenberg, "các cửa hàng bách hóa chiếm lĩnh một tỷ trọng lớn hơn trong số đô la mua lẻ được chi tiêu bởi những khách hàng già hơn của chúng, so với tỷ trọng chúng chiếm lĩnh từ những khách hàng trẻ hơn."

Tuy nhiên, khi thế hệ bùng nổ sinh đẻ bước vào trung niên, thì mối quan ngại là những số lượng của họ có thể không tự động chuyển thành doanh thu bán hàng tăng cho các cửa hàng bách hóa. Mặc dù những người già hơn chi tiêu một tỷ lệ lớn hơn trong số đô la mua lẻ của mình ở các cửa hàng bách hóa, nhưng họ chi tiêu ít hơn những người tương ứng trẻ hơn của họ. Phân phối của tỷ trọng chi tiêu mua lẻ phân theo các khách hàng của cửa hàng bách hóa (dựa trên 60.000 khách hàng đã được lấy mẫu) được tóm lược trong Bảng **3.1**.

Bảng **3.1**
Chi tiêu Mua lẻ

Độ tuổi	**Tỷ trọng Chi tiêu Mua lẻ**	**Tỷ trọng Doanh thu Bán hàng**	**Số tiền được Chi tiêu Hàng tháng bình quân mỗi Khách hàng**
18−24	0,22	0,15	50$
25−34	0,24	0,27	108
35−44	0,26	0,25	126
45−54	0,28	0,16	116
55−64	0,28	0,12	89
65 và già hơn	0,39	0,05	53

Nguồn: Marvin J. Rothenberg, Inc., Fair Lawn, N.J.

Trong chương này, anh/chị sẽ được giới thiệu về vai trò của xác suất trong thống kê. Về đề tài này, anh/chị sẽ phát hiện rằng những cuộc thăm dò dư luận và những nghiên cứu đều tạo ra các biến ngẫu nhiên. Gọi là biến ngẫu nhiên bởi vì các giá trị của chúng biến đổi từ mẫu này sang mẫu khác, và chúng rất thường biến đổi và thay đổi theo thời gian. Các Mục 3.1 đến 3.5 trình bày phần giới thiệu tuy ngắn gọn nhưng có hệ thống về các khái niệm căn bản của xác suất. Mục 3.6 bàn về những biến ngẫu nhiên và các phân phối xác suất của chúng. Chúng ta sẽ tiếp tục thảo luận về tỷ trọng của các cửa hàng bách hóa trong chi tiêu mua lẻ và thế hệ bùng nổ sinh đẻ trong Mục 3.7. (trong nguyên bản tiếng Anh)

3.1 Vai trò của Xác suất trong Thống kê

Xác suất và thống kê liên quan với nhau về một khía cạnh quan trọng. **Xác suất là công cụ cho phép nhà thống kê sử dụng thông tin của một mẫu để đưa ra những suy luận về, hay để mô tả, tổng thể mà từ đó mẫu này được lấy ra.** Chúng ta có thể minh họa mối quan hệ này bằng một thí dụ đơn giản sau đây.

Hãy xét một con súc sắc cân bằng, với sáu mặt quen thuộc của nó. Khi con súc sắc được thảy ra, bất kỳ mặt nào trong sáu mặt đều có khả năng ngang nhau trở thành mặt trên. Nếu mà chúng ta thảy con súc sắc này lặp đi lặp lại nhiều lần, thì chúng ta sẽ tạo ra một tổng thể các số, trong đó mặt trên, x, sẽ là 1, 2, 3, 4, 5, hay 6. Tổng thể này trông ra sao? Mặc dù nó sẽ lớn vô hạn, nhưng chúng ta vẫn có thể nói nó sẽ gồm có một số lượng bằng nhau của các mặt 1, 2, …, 6. Bây giờ chúng ta hãy thảy con súc sắc này một lần và quan sát giá trị của x. Điều này tương đương với việc lấy một mẫu có cỡ $n = 1$ từ tổng thể. Xác suất để $x = 2$ là bao nhiêu? Do biết được cấu trúc của tổng thể, nên chúng ta biết rằng mỗi trong sáu giá trị của x đều có khả năng xảy ra bằng nhau, do đó xác suất để $x = 2$ là 1/6. Đây là một ứng dụng đơn giản của *xác suất*. Khi tổng thể được biết, thì chúng ta có thể tính xác suất của việc quan sát được một mẫu nào đó.

Bây giờ giả sử tổng thể là *chưa* biết. Nghĩa là, chúng ta không biết liệu con súc sắc có cân bằng hay không. Tổng thể vẫn gồm có các mặt 1, 2, 3, …, 6, nhưng chúng ta không biết chúng xảy ra theo tỷ lệ nào! Để cố gắng tìm ra điều đó, chúng ta thảy con súc sắc $n = 10$ lần và ghi nhận mặt trên, x, sau mỗi lần thảy. Điều này tương đương với việc lấy một mẫu có cỡ $n = 10$ từ tổng thể. Giả sử rằng, trong mỗi lần thảy, mặt trên hóa ra đều là $x =1$. Anh/Chị sẽ suy luận gì về tổng thể? Anh/Chị có cho là con súc sắc này cân bằng không? Có lẽ không, bởi vì, nếu như con súc sắc này cân bằng, thì khả năng quan sát một con số "1" mười lần liên tiếp sẽ rất thấp. Hoặc là chúng ta đã quan sát được biến cố rất không chắc xảy ra này, hoặc là con súc sắc này không cân bằng. Chúng ta có thể nghiêng về kết luận thứ hai (con súc sắc không cân bằng), bởi vì nó *có khả năng xảy ra* (*probable*) hơn trong hai chọn lựa đó.

Thí dụ này cho thấy mối quan hệ giữa xác suất và thống kê. Khi tổng thể được biết, thì xác suất được sử dụng để mô tả khả năng xảy ra của các kết quả khác nhau của mẫu. Khi tổng thể là *chưa biết* và chúng ta chỉ có một mẫu, thì chúng ta có bài toán thống kê, đó là cố gắng đưa ra những suy luận về tổng thể chưa biết. Xác suất là công cụ chúng ta sử dụng để đưa ra những suy luận này. **Như thế xác suất suy luận từ tổng thể đến mẫu, trong khi đó thống kê hành động ngược lại, đi từ mẫu đến tổng thể.**

Như chúng ta giải thích ngôn ngữ xác suất, chúng ta sẽ giả định tổng thể được biết và sẽ tính xác suất của việc rút những mẫu khác nhau. Khi làm thế, chúng ta thật sự đang chọn một **mô hình** (**model**) cho một tình trạng vật thể, bởi vì thành phần thật sự của một tổng thể hiếm khi được biết trong thực tiễn. Như thế, nhà xác suất học lập mô hình cho một tình trạng vật thể (tổng thể) bằng xác suất rất giống nhà điêu khắc tạc tượng bằng đất sét. Trong những mục sau đây, chúng tôi sẽ sử dụng những thí dụ đơn giản để giúp anh/chị nắm vững khái niệm xác suất. Những ứng dụng thực tế sẽ được trình bày tiếp theo những thí dụ đơn giản này.

3.2 Xác suất của một Biến cố

Dữ liệu được thu nhận bằng cách quan sát những biến cố không được kiểm soát trong thiên nhiên hoặc bằng cách quan sát những tình trạng được kiểm soát trong phòng thí nghiệm. Chúng ta sẽ dùng thuật ngữ **thí nghiệm** (**experiment**) để mô tả cả hai phương pháp thu thập dữ liệu nói trên.

ĐỊNH NGHĨA • Một **thí nghiệm** là một quy trình qua đó một quan sát (hay một giá trị đo lường) được thu nhận. •

Lưu ý rằng quan sát không nhất thiết mang lại một giá trị bằng số. Đây là một số thí dụ điển hình về các thí nghiệm.

1 Ghi sản lượng hàng ngày của một nhà máy chế tạo.

2 Ghi tỷ giá hối đoái giữa đô la Mỹ và đồng bảng Anh.

3 Phỏng vấn một người tiêu dùng để xác định sự ưa thích sản phẩm trong số một nhóm gồm mười loại xe hơi.

4 Kiểm tra một bóng đèn để xác định xem liệu nó là một sản phẩm có khuyết tật hay chấp nhận được.

5 Tung một đồng xu và quan sát mặt xuất hiện.

Các thí nghiệm có thể dẫn đến một kết quả hay nhiều hơn được gọi là **biến cố** (**events**) và được ký hiệu bằng chữ hoa.

ĐỊNH NGHĨA • Một **biến cố** (**event**) là kết quả của một thí nghiệm. •

THÍ DỤ **3.1** Thí nghiệm: Hãy thảy một con súc sắc và quan sát con số xuất hiện ở mặt trên. Một số biến cố sẽ như sau:

biến cố A :	quan sát được một số lẻ
biến cố B :	quan sát được một số nhỏ hơn 4
biến cố E_1 :	quan sát được số 1
biến cố E_2 :	quan sát được số 2
biến cố E_3 :	quan sát được số 3
biến cố E_4 :	quan sát được số 4
biến cố E_5 :	quan sát được số 5
biến cố E_6 :	quan sát được số 6 •

Trong thí dụ 3.1, có sự khác biệt rõ ràng giữa các biến cố A và B với các biến cố E_1, E_2, E_3, E_4, E_5, và E_6. Biến cố A sẽ xảy ra nếu biến cố E_1, E_3, hay E_5, xảy ra—nghĩa là, nếu chúng ta quan sát được số 1, 3, hay 5. Như thế A có thể được chia tách thành một tập hợp các biến cố đơn giản hơn, đó là E_1, E_3, và E_5. Tương tự, biến cố B sẽ xảy ra nếu E_1, E_2, hay E_3 xảy ra và có thể được xem là một tập hợp các biến cố nhỏ hơn hoặc đơn giản hơn. Ngược lại, không thể chia tách các biến cố E_1, E_2, E_3, …, E_6. Những biến cố này được gọi là **biến cố đơn (simple event)**.

ĐỊNH NGHĨA • Một biến cố mà không thể chia tách được thì được gọi là một **biến cố đơn**. Các biến cố đơn sẽ được biểu thị bằng ký hiệu E với một chỉ số dưới dòng. •

Các biến cố E_1, E_2, …, E_6 là một danh sách hoàn chỉnh tất cả các biến cố đơn gắn liền với thí nghiệm trong Thí dụ 3.1. **Một thí nghiệm sẽ dẫn đến một và chỉ một trong những biến cố đơn đó**. Chẳng hạn như, nếu thảy một con súc sắc, thì chúng ta sẽ quan sát được một con số 1, 2, 3, 4, 5, hay 6, nhưng mà chúng ta hẳn nhiên không thể quan sát được hơn một trong những biến cố đơn này cùng một lúc. Như thế, danh sách các biến cố đơn cung cấp một bản phân tích thống kê về tất cả các kết quả không thể chia tách có thể có của thí nghiệm này.

ĐỊNH NGHĨA • **Không gian mẫu *S* (sample space)** là tập hợp tất cả các kết quả có thể có của một thí nghiệm. •

Cuối cùng, chúng ta có thể định nghĩa các kết quả của một thí nghiệm theo các biến cố đơn gắn liền với nó.

ĐỊNH NGHĨA • Một **biến cố** là một tập hợp của một hay nhiều hơn một biến cố đơn •

Xác suất của một biến cố *A* là một thước đo về sự tin tưởng của chúng ta rằng một thí nghiệm sẽ dẫn đến biến cố A. Để gắn ý nghĩa cho khái niệm này, chúng ta lưu ý rằng các tổng thể và các quan sát được tạo ra bằng cách lặp đi lặp lại một thí nghiệm rất nhiều lần. Nếu trong số N rất lớn những lần lặp đi lặp lại thí nghiệm này mà biến cố A được quan sát n lần, thì chúng ta xem xác suất của biến cố A là

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

Thực ra, $P(A)$ là giá trị giới hạn của phân số n/N khi N trở nên lớn vô hạn. Cách diễn giải thực tế này về ý nghĩa của xác suất − đây là cách diễn giải của hầu hết những người không chuyên môn − được gọi là **khái niệm tần suất tương đối về xác suất (relative frequency concept of probability)**.

Không cần phải nói là việc lặp đi lặp lại một thí nghiệm rất nhiều lần sẽ rất tốn thời gian. Vì lý do này, chúng ta sẽ sử dụng những phương pháp thay thế khác nhau để tính toán xác suất mà phù hợp với khái niệm tần suất tương đối này. Thí dụ, định nghĩa về xác suất dựa trên tần suất tương đối hàm ý rằng $P(A)$ phải là một phân số nằm trong khoảng từ 0 đến 1, bao gồm cả 0 và 1, với $P(A) = 0$ nếu biến cố A không bao giờ xảy ra và $P(A) = 1$ nếu A luôn luôn xảy ra. Vì thế, $P(A)$ càng gần với số 1, thì A càng có khả năng sẽ xảy ra.

Một số biến cố A và B có đặc tính độc đáo là khi biến cố này xảy ra thì biến cố kia không thể xảy ra (và ngược lại). Các biến cố có đặc tính này thì được gọi là **loại trừ lẫn nhau (mutually exclusive)**. Thí dụ, giả sử một thí nghiệm là quan sát hành động trong tháng tiếp theo của Hội đồng Thống đốc của Hệ thống Dự trữ Liên bang (Hoa Kỳ) đối với lãi suất cơ bản. Định rõ các biến cố sau đây:

A : lãi suất cơ bản được tăng lên.
B : lãi suất cơ bản được hạ xuống.
C : lãi suất cơ bản không thay đổi.

Như thế, các biến cố *A* và *B* là các biến cố loại trừ lẫn nhau, bởi vì một thí nghiệm duy nhất không thể dẫn đến cả *A* lẫn *B*. Nếu việc công bố tiếp theo về chính sách lãi suất cơ bản dẫn đến *A*, thì *B* không thể đồng thời xảy ra. Nếu lãi suất cơ bản đã được tăng lên, thì nó không có thể đã được hạ xuống đồng thời. Các biến cố *A*, *B*, và *C* cũng được gọi là các biến cố loại trừ lẫn nhau. Nếu bất cứ một biến cố nào trong nhóm này được quan sát, thì không biến cố nào trong cả hai biến cố còn lại có thể đã xảy ra.

ĐỊNH NGHĨA • Hai biến cố *A* và *B* được gọi là **loại trừ lẫn nhau** nếu như khi *A* xảy ra thì *B* không thể xảy ra (và ngược lại). •

Các biến cố đơn có tính loại trừ lẫn nhau, và vì thế cho nên xác suất gắn liền với các biến cố đơn thỏa mãn những điều kiện sau đây.

Yêu cầu đối với các Xác suất của Biến cố-Đơn

1. Mỗi xác suất phải nằm trong khoảng từ 0 đến 1, bao gồm cả 0 và 1.
2. Tổng số của các xác suất đối với tất cả các biến cố đơn trong *S* bằng 1.

Khi có thể viết ra các biến cố đơn gắn liền với một thí nghiệm và đánh giá các xác suất riêng của chúng, thì chúng ta có thể tìm xác suất của biến cố *A* bằng cách cộng các xác suất đối với các biến cố đơn chứa đựng trong biến cố *A*.

ĐỊNH NGHĨA • **Xác suất của biến cố *A*** bằng tổng các xác suất của những biến cố đơn chứa đựng trong *A*. •

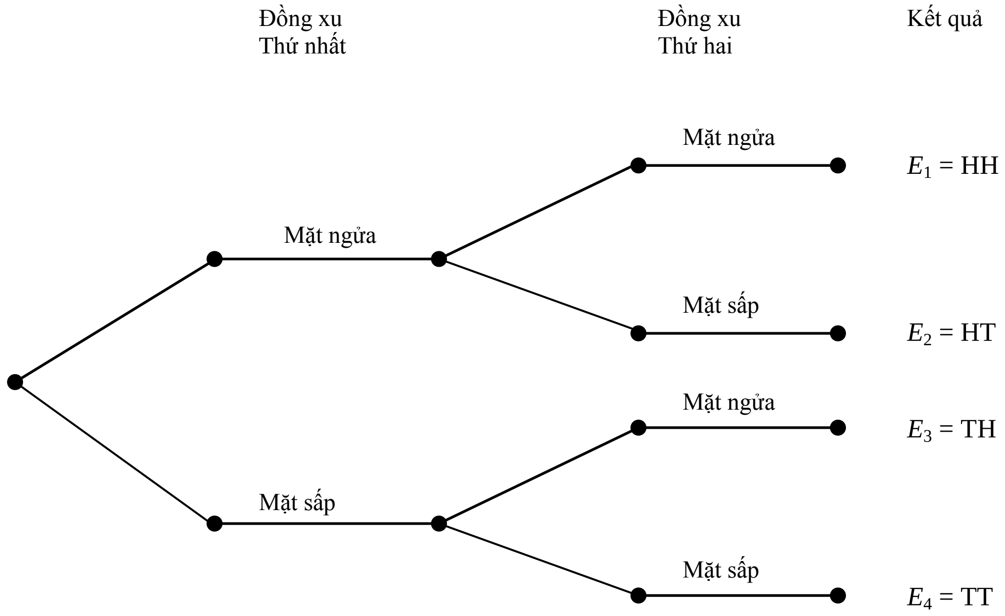
THÍ DỤ **3.2** Sử dụng phương pháp biến cố đơn, hãy tính xác suất của việc quan sát chính xác một mặt ngửa trong một lần tung hai đồng xu.

Lời giải Hãy xây dựng không gian mẫu và cho *H* biểu thị mặt ngửa và *T* biểu thị mặt sấp. Các kết quả gắn liền với thí nghiệm này có thể được trình bày bằng cách dùng một **giản đồ hình cây (tree diagram)**. Trong một giản đồ hình cây, mỗi cấp phân nhánh kế tiếp tương ứng với bước kế tiếp cần phải có để tạo ra các kết quả có thể có của thí nghiệm. Giản đồ hình cây được trình bày trong Hình 3.1. Các biến cố đơn được tạo ra thì được trình bày ở cột cuối cùng của Hình 3.1 và trong Bảng 3.2.

BẢNG **3.2** Không gian mẫu đối với Thí dụ 3.2

Biến cố	Đồng xu Thứ nhất	Đồng xu Thứ hai	$P(E_i)$
E_1	H	H	1/4
E_2	H	T	1/4
E_3	T	H	1/4
E_4	T	T	1/4

HÌNH **3.1**
Giản đồ hình cây
đối với Thí dụ 3.2



Không gian mẫu *S* đối với thí nghiệm này có chứa bốn biến cố đơn được liệt kê trong Bảng 3.2. Bởi vì, một giả định hợp lý là, bất kỳ biến cố đơn nào trong các biến cố đơn này đều có khả năng xảy ra như bất kỳ biến cố đơn nào khác, nên chúng ta ấn định xác suất là 1/4 cho mỗi biến cố đơn này. Chúng ta quan tâm đến

biến cố *A*: quan sát được chính xác một mặt ngửa

Biến cố *A* sẽ xảy ra nếu và chỉ nếu E_2 hoặc E_3 xảy ra. Vì thế cho nên,

$$\begin{aligned} P(A) &= P(E_2) + P(E_3) \\ &= \frac{1}{4} + \frac{1}{4} \\ &= \frac{1}{2} \quad \bullet \end{aligned}$$

THÍ DỤ **3.3** Căn cứ vào quan sát của chúng ta về hành động của Hội đồng Thống đốc của Hệ thống Dự trữ Liên bang đối với lãi suất cơ bản, và định rõ các biến cố *A*, *B*, và *C* như sau:

A : lãi suất cơ bản được tăng lên.

B : lãi suất cơ bản được giảm xuống.

C : lãi suất cơ bản không thay đổi.

Giả sử chúng ta biết rằng $P(A) = 0{,}2$ và $P(B) = 0{,}3$. Hãy tìm $P(C)$.

Lời giải Bằng việc xem xét *A*, *B*, và *C*, chúng ta có thể thấy thí nghiệm này có thể dẫn đến một và chỉ một trong ba biến cố loại trừ lẫn nhau này. Do đó, *A*, *B*, và *C* là những biến cố đơn biểu hiện không gian mẫu *S* cho thí nghiệm này, và

$$P(A)+P(B)+P(C)=1$$

Thay thế các giá trị của $P(A)$ và $P(B)$ vào phương trình này sẽ cho ra

$$0{,}2+0{,}3+P(C)=1$$

Giải để tìm $P(C)$, chúng ta có

$$P(C)=1-0{,}3-0{,}2=0{,}5 \quad \bullet$$

Những Lời Gợi ý về Giải Bài toán

Việc tính toán xác suất của một biến cố: phương pháp biến cố đơn.

- Hãy sử dụng các bước sau đây để tính xác suất của một biến cố bằng cách cộng các xác suất của các biến cố đơn
  1 Định rõ thí nghiệm
  2 Nhận dạng một biến cố đơn tiêu biểu. Liệt kê những biến cố đơn gắn liền với thí nghiệm, và kiểm định mỗi biến cố đơn để nắm chắc rằng nó không thể chia tách được. Điều này định rõ không gian mẫu *S*.
  3 Ấn định các xác suất hợp lý cho các biến cố đơn trong *S*, nắm chắc rằng mỗi xác suất đều nằm trong khoảng từ 0 đến 1 (bao gồm cả 0 và 1) và tổng số các xác suất của các biến cố đơn bằng 1.
  4 Định rõ các biến cố được quan tâm *A* như là một tập hợp cụ thể các biến cố đơn. (Một biến cố đơn ở trong *A* nếu *A* xảy ra khi biến cố đơn này xảy ra. Hãy kiểm định *tất cả* các biến cố đơn trong *S* để xác định những biến cố đơn nào ở trong *A*).
  5 Tìm $P(A)$ bằng cách cộng các xác suất của các biến cố đơn trong *A*.
- Khi các biến cố đơn có khả năng xảy ra ngang nhau, hãy lấy tổng số các xác suất của các biến cố đơn trong *A* (bước 5) bằng cách đếm những điểm trong *A* và nhân với xác suất của mỗi biến cố đơn.
- Việc tính toán xác suất của một biến cố bằng cách dùng thủ tục gồm năm bước được mô tả ở trên là có hệ thống và sẽ dẫn đến lời giải đúng, nếu tất cả các bước được theo đúng. Những nguồn sai số chính bao gồm những điều sau đây:
  1 Việc không định rõ thí nghiệm (bước 1).
  2 Việc không nêu rõ các biến cố đơn (bước 2).
  3 Việc không liệt kê tất cả các biến cố đơn (bước 2).
  4 Việc không ấn định các xác suất hợp lệ cho các biến cố đơn (bước 3).

Bài tập

Các kỹ thuật căn bản

3.1 Một thí nghiệm liên quan đến việc thảy một con súc sắc. Hãy nêu rõ các biến cố đơn trong những biến cố sau đây:

A : quan sát được số 2
B : quan sát được một số lẻ
C : quan sát được một số nhỏ hơn 4
D : quan sát được cả A và B
E : quan sát được A hoặc B hoặc cả hai
F : quan sát được cả A và C.

Hãy tính xác suất của các biến cố từ *A* đến *F* bằng cách cộng các xác suất của những biến cố đơn thích hợp.

3.2 Một thí nghiệm có thể dẫn đến một trong bốn biến cố đơn, với

$$P(E_1) = 0{,}1 \qquad P(E_2) = 0{,}15 \qquad P(E_3) = 0{,}6 \qquad P(E_4) = 0{,}15$$

Hãy giải thích tại sao việc ấn định các xác suất này cho những biến cố đơn là hợp lý hay là không hợp lý.

3.3 Một thí nghiệm có thể dẫn đến một trong năm biến cố, $E_1, E_2, E_3, E_4,$ và E_5. Nếu $P(E_1) = P(E_2) = P(E_3) = P(E_4) = 0{,}15$, hãy tìm $P(E_5)$.

3.4 Một không gian mẫu có chứa năm biến cố đơn $E_1, E_2, E_3, E_4,$ và E_5. Nếu $P(E_3) = 0{,}4$, $P(E_4) = 2P(E_5)$, và $P(E_1) = P(E_2) = 0{,}15$, hãy tìm các xác suất của E_4 và E_5.

3.5 Một không gian mẫu có chứa mười biến cố đơn, E_1, E_2, …, E_{10}. Nếu $P(E_1) = 3P(E_2) = 0{,}45$, với những biến cố đơn còn lại có khả năng xảy ra ngang nhau, hãy tìm xác suất của những biến cố đơn còn lại này.

3.6 Trò chơi ru-lét sử dụng một vòng ru-lét có 38 ngăn. Ba mươi sáu ngăn được đánh số 1, 2, …, 36, và hai ngăn còn lại được đánh dấu 0 và 00. Vòng ru-lét được quay tròn và một trong những ngăn này được xác định là “người thắng cuộc.” (Quả cầu nhỏ dừng lại ở ngăn nào thì người dự đoán đúng ngăn đó sẽ thắng). Giả định rằng việc quan sát bất kỳ ngăn nào cũng có khả năng xảy ra như bất kỳ ngăn nào khác.

a Hãy xác định các biến cố đơn trong một lần quay vòng ru-lét.

b Hãy ấn định xác suất cho các biến cố đơn này.

c Cho *A* là biến cố anh/chị quan sát được 0 hoặc 00. Hãy liệt kê các biến cố đơn trong biến cố *A* , và hày tìm *P*(*A*).

d Giả định anh/chị đánh cược vào các số 1 đến 18. Xác suất để một trong các con số của anh/chị sẽ thành người thắng cuộc là bao nhiêu?

Ứng dụng

3.7 Một công ty thăm dò dầu mỏ khoan trúng dầu hoặc khí trong 10% những giếng khoan của mình. Nếu công ty này khoan hai giếng, thì bốn biến cố đơn có thể xảy ra và ba trong các xác suất liên quan của chúng được trình bày trong bảng dưới đây:

Biến cố Đơn	**Kết quả của Giếng khoan Thứ nhất**	**Kết quả của Giếng khoan Thứ hai**	**Xác suất**
1	Trúng (dầu hay khí)	Trúng (dầu hay khí)	0,01
2	Trúng	Không trúng	?
3	Không trúng	Trúng	0,09
4	Không trúng	Không trúng	0,81

a Hãy xây dựng một giản đồ hình cây để biểu thị thí nghiệm này.

b Hãy tìm xác suất để công ty này sẽ khoan trúng dầu hay khí trong giếng khoan thứ nhất và không trúng trong giếng khoan thứ hai.

c Hãy tìm xác suất để công ty này sẽ khoan trúng dầu hay khí trong ít nhất là một trong hai giếng khoan này.

3.8 Một cuộc điều tra tiếp thị cho một cửa hàng bách hóa lớn đã phân loại khách hàng của cửa hàng này dựa theo việc liệu họ là nam hay nữ và dựa theo nơi cư trú của họ, vùng ngoại ô hay thành phố. Các tỷ lệ của khách hàng rơi vào bốn loại đó được trình bày trong *bảng xác suất* sau đây. Mỗi số liệu trong bảng này thể hiện một kết quả có thể có của thí nghiệm này (một biến cố đơn).

	Giới tính	
Nơi cư trú	**Nam**	**Nữ**
Vùng ngoại ô	0,17	0,67
Thành phố	0,04	0,12

Giả sử một người trưởng thành đơn lẻ được chọn từ trong nhóm các khách hàng này. Hãy tìm các xác suất sau đây.

a xác suất để khách hàng này cư trú ở vùng ngoại ô.

b xác suất để khách hàng này là một người nữ sống ở thành phố.

c xác suất để khách hàng này là một người nam.

3.9 Mỗi một trong hai nhà phân tích thị trường cổ phiếu được yêu cầu dự báo xem đến cuối 12 tháng tới, liệu số trung bình cổ phiếu Dow-Jones sẽ tăng 100 điểm hay nhiều hơn, sẽ giảm 100 điểm hay nhiều hơn, hoặc sẽ thay đổi ít hơn 100 điểm. Như thế thí nghiệm này gồm có việc quan sát cặp dự báo do hai nhà phân tích thị trường đưa ra. Giả sử mỗi nhà phân tích có khả năng chọn bất kỳ một phương án nào trong ba phương án chọn lựa này ngang bằng khả năng chọn bất kỳ phương án còn lại nào.

a Hãy liệt kê các biến cố đơn trong không gian mẫu *S*. Hãy sử dụng giản đồ hình cây.

b Cho *A* là biến cố ít nhất là một trong hai nhà phân tích dự báo số trung bình Dow-Jones sẽ tăng 100 điểm hay nhiều hơn. Hãy tìm các biến cố đơn trong *A*.

c Cho *B* là biến cố cả hai nhà phân tích đều dự báo giống nhau. Hãy tìm các biến cố đơn trong *B*.

d Hãy ấn định xác suất cho các biến cố đơn trong *S*, và tìm *P*(*A*).

e Hãy tìm *P*(*B*).

3.10 Một công ty thực phẩm dự định tiến hành một thí nghiệm để so sánh nhãn hiệu trà của mình với nhãn hiệu trà của hai đối thủ cạnh tranh. Khi thực hành thực sự thì một số chuyên viên nếm trà sẽ được thuê. Đối với thí dụ này, chúng ta sẽ giả định rằng chỉ một chuyên viên nếm trà nếm từng nhãn hiệu trong ba nhãn hiệu trà này, mà chúng không được đánh dấu gì ngoại trừ các ký hiệu nhận dạng A, B, và C.

a Hãy định rõ thí nghiệm này

b Hãy liệt kê các biến cố đơn trong *S*.

c Nếu chuyên viên nếm trà này không có năng lực phân biệt sự khác nhau về vị giữa các loại trà, thì xác suất để chuyên viên nếm trà này sẽ xếp loại trà A vào hạng ngon nhất là bao nhiêu? Vào hạng ít đáng mong muốn nhất là bao nhiêu?

3.11 Bốn người trong công đoàn, trong đó hai người từ một nhóm thiểu số, được phân công bốn công việc khác nhau rõ ràng và mỗi công việc dành cho chỉ một người làm.

a Hãy định rõ thí nghiệm.

b Hãy liệt kê các biến cố đơn trong *S*.

c Nếu việc phân công làm những công việc này không thiên lệch—nghĩa là nếu bất kỳ một việc ra lệnh phân công nào cũng có khả năng xảy ra như bất kỳ việc ra lệnh nào khác—thì xác suất để hai người từ nhóm thiểu số được phân công làm hai công việc ít đáng mong muốn nhất là bao nhiêu?

3.3 Sự Hợp thành của Biến cố và Quan hệ của Biến cố

Thường thì chúng ta quan tâm đến các kết quả thí nghiệm được tạo nên bởi sự hợp thành (compotisition) nào đó của hai, hay nhiều hơn, biến cố. Các biến cố thường có thể được tạo nên bởi **hội** (**unions**) hay **giao** (**intersections**) của những biến cố khác, hay bởi một kết hợp nào đó của cả hai.

ĐỊNH NGHĨA • **Giao** của các biến cố *A* và *B*, được ký hiệu bằng *AB*, là biến cố cả *A* và *B* đều xảy ra.[†] •

ĐỊNH NGHĨA • **Hội** của các biến cố *A* và *B*, được ký hiệu bằng $A \cup B$, là biến cố *A* hoặc *B* hoặc cả hai xảy ra •

THÍ DỤ **3.4** Căn cứ vào thí nghiệm trong Thí dụ 3.2, trong đó hai đồng xu được tung lên, và định rõ.

biến cố *A* : ít nhất một mặt ngửa

biến cố *B* : ít nhất một mặt sấp

Định rõ các biến cố *A, B, AB,* và $A \cup B$ như là những tập hợp của các biến đơn, và hãy tìm các xác suất của chúng.

[†] Một số tác giả sử dụng ký hiệu $A \cap B$.

Lời giải Hãy nhớ lại rằng các biến cố đơn đối với thí nghiệm này là

$E_1 =$ HH (mặt ngửa trên đồng xu thứ nhất, mặt ngửa trên đồng xu thứ hai)

$E_2 =$ HT

$E_3 =$ TH

$E_4 =$ TT

Sự xuất hiện các biến cố đơn E_1, E_2, hay E_3 hàm ý ít nhất là có một mặt ngửa và như thế định rõ biến cố *A*, và $P(A) = \frac{1}{4} = \frac{1}{4} = \frac{1}{4} = \frac{3}{4}$. Các biến cố khác được định rõ một cách tương tự.

biến cố *B*: E_2, E_3, E_4, và $P(B) = \frac{3}{4}$

biến cố *AB:* E_2, E_3, và $P(AB) = \frac{2}{4} = \frac{1}{2}$

biến cố $A \cup B$: E_1, E_2, E_3, E_4, và $P(A \cup B) = \frac{4}{4} = 1$

Lưu ý rằng $A \cup B$ bao gồm toàn bộ không gian mẫu và do đó chắc chắn xảy ra.

Như thế $P(A \cup B) = 1$ •

Khi hai biến cố *A* và *B* có tính loại trừ lẫn nhau, nghĩa là khi *A* xảy ra thì *B* không thể xảy ra, và ngược lại. Các biến cố loại trừ lẫn nhau còn được gọi là **các biến cố cách biệt nhau** (**disjoint events**). Khi *A* và *B* có tính loại trừ lẫn nhau, thì

1 $P(AB) = 0$

2 $P(A \cup B) = P(A) + P(B)$

Cụ thể là, nếu *P*(*A*) và *P*(*B*) được biết, thì chúng ta không cần liệt kê các biến đơn tạo thành $A \cup B$ và cộng các xác suất riêng của chúng lại; mà chúng ta có thể đơn giản là cộng *P*(*A*) và *P*(*B*).

THÍ DỤ **3.5** Một bản tóm lược về hoạt động (tổng số tiền tính bằng đô la của những séc được viết) của các tài khoản sử dụng séc cá nhân với số dư tối thiểu 1000$ của một ngân hàng được trình bày trong bảng sau đây.

Hoạt động	Dưới 1000$	1000−2999$	3000−4999$	5000−7999$	8000$ và cao hơn
Tỷ lệ phần trăm	17%	43%	28%	9%	3%

Một tài khoản sử dụng séc với số dư tối thiểu 1000$ và đơn lẻ được chọn ngẫu nhiên từ những tài khoản tại ngân hàng này.

a Hãy tìm xác suất của các biến cố sau đây:

A : hoạt động sử dụng séc ít hơn 3.000$

B : hoạt động sử dụng séc từ 1000$ đến 4.999$

C : hoạt động sử dụng séc vượt quá 4.999$

b Hãy tìm $P(B \cup C)$.

c Các biến cố *A*, *B*, và *C* có loại trừ lẫn nhau không?

Lời giải Các biến cố đơn trong thí nghiệm này được cho trong bảng, với E_1 = (dưới 1000$), E_2 = (1000 − 2999$) và v.v..

a Như thế,

$$P(A) = P(E_1) + P(E_2) = 0{,}17 + 0{,}43 = 0{,}60$$
$$P(B) = P(E_2) + P(E_3) = 0{,}43 + 0{,}28 = 0{,}71$$
$$P(C) = P(E_4) + P(E_5) = 0{,}09 + 0{,}03 = 0{,}12$$

b Biến cố $B \cup C$ gồm có các biến cố đơn E_2, E_3, E_4, và E_5, do đó

$$P(B \cup C) = 0{,}43 + 0{,}28 + 0{,}09 + 0{,}03 = 0{,}83$$

c *A* và *C* loại trừ lẫn nhau, *B* và *C* cũng thế. Tuy nhiên, *A* và *B* có chứa biến cố đơn chung E_2 và như thế không loại trừ lẫn nhau. Vì thế cho nên, *A*, *B*, và *C* không loại trừ lẫn nhau. •

Sự bù (**Complementation**) là một mối quan hệ khác của biến cố mà thường đơn giản hóa việc tính toán xác suất.

ĐỊNH NGHĨA • **Phần bù** (**complement**) của một biến cố *A*, được ký hiệu bằng $\overline{A}$, gồm có tất cả những biến cố đơn trong không gian mẫu mà không ở trong *A*. •

Phần bù của *A* là biến cố *A* không xảy ra. Vì thế cho nên, *A* và $\overline{A}$ loại trừ lẫn nhau, và $A \cup \overline{A}$ tạo thành toàn bộ không gian mẫu. Tất yếu là $P(A) + P(\overline{A}) = 1$ và

$$P(A) = 1 - P(\overline{A})$$

Thí dụ, nếu biến cố *A* là sự xuất hiện của ít nhất một mặt ngửa trong việc tung hai đồng xu không thiên lệch (xem Thí dụ 3.4), thì biến cố $\overline{A}$ là sự xuất hiện của không có mặt ngửa nào trong việc tung đồng xu, và

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{1}{4} = \frac{3}{4}$$

3.4 Xác suất có Điều kiện và những Biến cố Độc lập

Hai biến cố thường quan hệ với nhau theo một cách nào đó để xác suất của việc xảy ra biến cố này phụ thuộc vào việc liệu biến cố kia đã xảy ra hay không. Thí dụ, giả sử anh/chị là một người mua bán tiền tệ mà anh/chị giao dịch đô la Mỹ với bảng Anh và thí nghiệm chủ yếu là quan sát liệu đô la Mỹ có lên giá (đã tăng giá trị) so với bảng Anh hay không. Gọi *A* là biến cố đô la Mỹ lên giá so với bảng Anh trên thị trường tiền tệ của Anh trước khi thị trường của chúng ta (thị trường Anh) mở cửa vào đúng 9 giờ sáng, và gọi *B* là biến cố đô la Mỹ lên giá trên thị trường tiền tệ Hoa Kỳ sau khi thị trường này mở cửa.

Các biến cố A và B chắc chắn có quan hệ với nhau, bởi vì hai thị trường này sẽ có khả năng lên xuống cùng chiều trong hầu hết (nhưng không nhất thiết tất cả) mọi ngày. Vì thế, xác suất $P(B)$ để đô la Mỹ sẽ lên giá so với bảng Anh trên thị trường Hoa Kỳ không giống với xác suất để B sẽ xảy ra trên cơ sở anh/chị biết rằng đô la Mỹ đã lên giá (biến cố A) trên thị trường Anh rồi.

Thí dụ, giả sử đô la Mỹ lên giá so với bảng Anh trong 60% của tất cả mọi ngày trên thị trường Anh, và trong 50% của tất cả mọi ngày, nó lên giá trên cả hai thị trường Mỹ và Anh (nghĩa là $P(A) = 0,6$ và $P(AB) = 0,5$). Như thế, nếu anh/chị đã biết đô la Mỹ lên giá so với bảng Anh trên thị trường Anh, thì xác suất để nó sẽ lên giá trên thị trường Mỹ là 5/6. Phân số này, $P(AB)/P(A)$, được gọi là **xác suất có điều kiện** (**conditional probability**) của B trên cơ sở biến cố A đã xảy ra, và được biểu thị bằng ký hiệu

$$P(B \mid A)$$

Gạch thẳng đứng trong biểu thức $P(B \mid A)$ được hiểu là "cho trước" và các biến cố xuất hiện bên phải của gạch này là những biến cố đã xảy ra.

ĐỊNH NGHĨA • **Xác suất có điều kiện** của B trên cơ sở biến cố A đã xảy ra là

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

và xác suất có điều kiện của A trên cơ sở biến cố B đã xảy ra là

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

THÍ DỤ **3.6** Hãy tính $P(A \mid B)$ đối với thí nghiệm thảy con súc sắc được mô tả trong Hình 3.1, trong đó

A : quan sát được một số lẻ

B : quan sát được một số nhỏ hơn 4

Lời giải Trên cơ sở biến cố B (một số nhỏ hơn 4) đã xảy ra, chúng ta đã quan sát được một số 1, 2, hay 3, tất cả các con số này xảy ra với tần suất bằng nhau. Trong ba biến cố đơn này, chính xác là hai biến cố (1 và 3) dẫn đến biến cố A (một số lẻ). Như thế,

$$P(A \mid B) = \frac{2}{3}$$

Hoặc chúng ta có thể có được $P(A \mid B)$ bằng cách thay thế các giá trị vào phương trình

$$P(A \mid B) = \frac{P(AB)}{P(B)} = \frac{1/3}{1/2} = \frac{2}{3}$$

Lưu ý rằng $P(A \mid B) = 2/3$ và $P(A) = 1/2$, điều này chỉ ra A và B phụ thuộc lẫn nhau. •

ĐỊNH NGHĨA • Hai biến cố A và B được gọi là **độc lập (independent**) nếu và chỉ nếu

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

hoặc

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

Nếu không thì các biến cố này được gọi là **phụ thuộc** (**dependent**). •

Chuyển định nghĩa này thành lời: Hai biến cố là **độc lập** nếu sự xuất hiện hay sự không xuất hiện của biến cố này không làm thay đổi xác suất của sự xuất hiện biến cố kia. Nếu $P(A \mid B) = P(A)$, thì $P(B \mid A)$ cũng sẽ bằng $P(B)$. Tương tự, nếu $P(A \mid B)$ và $P(A)$ không bằng nhau, thì $P(B \mid A)$ và $P(B)$ cũng sẽ không bằng nhau.

THÍ DỤ **3.7** Căn cứ vào thí nghiệm thảy con súc sắc trong Thí dụ 3.1, trong đó

A : quan sát được một số lẻ

B : quan sát được một số nhỏ hơn 4

Hai biến cố A và B có loại trừ lẫn nhau không? Chúng có bù nhau không? Chúng có độc lập không?

Lời giải Thể hiện theo các biến cố đơn, chúng ta có thể viết

biến cố A : E_1, E_3, E_5

biến cố B : E_1, E_2, E_3

Biến cố AB là tập hợp các biến cố đơn trong cả A và B. Bởi vì AB bao gồm các biến cố E_1 và E_3, nên A và B không loại trừ lẫn nhau. Chúng không bù nhau bởi vì B không phải là tập hợp của tất cả các kết quả trong S mà không ở trong A. Kiểm định về sự độc lập nằm trong chính định nghĩa về độc lập; nói rõ ra là chúng ta sẽ kiểm tra để xem liệu $P(A \mid B) = P(A)$ hay không. Từ Thí dụ 3.6, $P(A \mid B) = 2/3$. Như thế, vì $P(A) = 1/2$, nên $P(A \mid B) \neq P(A)$ và theo định nghĩa, các biến cố A và B là phụ thuộc. •

Một phương pháp khác để giải các bài toán xác suất là dựa vào những quan hệ của biến cố và hai quy tắc xác suất, mà chúng ta sẽ trình bày chứ không chứng minh. Quy tắc thứ nhất được gọi là **quy tắc cộng xác suất** (**additive rule of probability**), và nó đề cập đến hội của các biến cố.

Quy tắc Cộng Xác suất

Cho trước hai biến cố A và B, xác suất của hội $A \cup B$ của chúng là

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

Nếu A và B loại trừ lẫn nhau, thì $P(AB) = 0$ và

$$P(A \cup B) = P(A) + P(B)$$

Hãy lưu ý rằng tổng số $P(A) + P(B)$ tính hai lần các biến cố đơn nằm trong cả A lẫn B; do đó việc trừ đi $P(AB)$ sẽ cho kết quả đúng.

Quy tắc thứ hai về xác suất được gọi là **quy tắc nhân xác suất** (**multiplicative rule of probability**) và là kết quả tất yếu của định nghĩa về xác suất có điều kiện.

Quy tắc Nhân Xác suất

Xác suất để cả hai biến cố *A* và *B* xảy ra là

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

Nếu *A* và *B* là các biến cố độc lập, thì

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

Tương tự, nếu *A*, *B*, và *C* là các biến cố độc lập với nhau, thì xác suất để *A*, *B*, và *C* đều sẽ xảy ra là

$$P(ABC) = P(A)P(B)P(C)$$

THÍ DỤ **3.8** Giả sử anh/chị chọn lựa ngẫu nhiên ba cổ phiếu thường, A, B, và C, từ trong số 500 cổ phiếu được sử dụng để tính số trung bình Standard & Poor's về 500 cổ phiếu. Xác suất để mức tăng hằng năm của cả ba cổ phiếu này sẽ cao hơn mức tăng của số trung bình S&P là bao nhiêu? Cho *A*, *B*, và *C* biểu thị các biến cố rằng từng cổ phiếu A, B, và C có kết quả vượt trội số trung bình S&P, và giả định $P(A) = P(B) = P(C) = 1/2$.

Lời giải Biến cố cả ba chọn lựa cổ phiếu đều sẽ vượt trội số trung bình S&P là giao của các biến cố *A*, *B*, và *C*. Chúng ta, không biết các xác suất có điều kiện của *A*, *B*, và C, do đó chúng ta không thể dùng định nghĩa của xác suất có điều kiện để kiểm định sự độc lập. Mà đúng ra, chúng ta phải dựa vào trực giác của mình. Bởi vì chúng ta đã chọn ba cổ phiếu này theo cách thức không liên quan với nhau, nên dường như không có khả năng rằng việc chọn lựa một cổ phiếu từ trong số 500 cổ phiếu nói trên sẽ ảnh hưởng nhiều đến việc chọn lựa một cổ phiếu khác. Vì lý do này, chúng ta sẽ tuyên bố các biến cố này độc lập và sẽ tính

$$\begin{aligned} P(ABC) &= P(A)P(B)P(C) \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8} \end{aligned}$$

Như thế chúng ta tính ra xác suất của việc chọn lựa ba cổ phiếu mà sẽ đều vượt trội số trung bình S&P, là 1/8. •

Một **bảng xác suất** (**probability table**) đối với các biến cố *A* và *B* là một bảng hai chiều, mà bốn số liệu trong bảng là bốn xác suất của những giao $P(AB)$, $P(A\overline{B})$, $P(\overline{A}B)$, và $P(\overline{AB})$, và những tổng số ở cột cuối cùng bên phải và hàng dưới cùng của bảng này tương ứng với các xác suất không điều kiện $P(A), P(\overline{A}), P(B)$, và $P(\overline{B})$, như trong Bảng 3.3. Có một số điểm thú vị và khá hiển nhiên từ bảng xác suất. Một điểm là, một biến cố *A* có thể được biểu hiện như là gồm có hai phần—đó là *AB*, những biến cố đơn nào trong *A* mà cũng ở trong *B*, và $A\overline{B}$, những biến cố đơn nào trong *A* mà không ở trong *B*—do đó

$$P(A) = P(AB \cup A\overline{B}) = P(AB) + P(A\overline{B})$$

Một điểm khác là, $P(A \mid B)$, xác suất có điều kiện của A, với B được cho trước, thể hiện phần của các biến cố đơn trong B mà cũng ở trong A, do đó $P(A \mid B) = P(AB) / P(B)$ chỉ đơn thuần là xác suất đã được thay đổi tỷ lệ (rescaled probability) đối với một không gian mẫu đã giảm và tương đương với B.

BẢNG **3.3**

	B	$\overline{B}$	**Tổng số**
A	$P(AB)$	$P(A\overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A}B)$	$P(\overline{AB})$	$P(\overline{A})$
Tổng số	$P(B)$	$P(\overline{B})$	1

THÍ DỤ **3.9** Một thương gia nhận đặt và giao hàng may mặc qua đường bưu điện. Thương gia này bán hai dòng sản phẩm, một dòng tương đối đắt, dòng kia không đắt. Một cuộc điều tra về các đơn đặt hàng đã cho thấy tần suất tương đối của các đơn đặt hàng, phân theo dòng sản phẩm và giới tính của khách hàng. Trong Bảng 3.4, A là biến cố khách hàng là nữ và B là biến cố đơn đặt hàng là từ dòng sản phẩm 1. Giả sử rằng một khách hàng, người đã đặt một đơn đặt hàng, được chọn ngẫu nhiên.

a Tìm xác suất để khách hàng này là nữ.

b Tìm xác suất để đơn đặt hàng này là từ dòng sản phẩm 1.

c Tìm xác suất để đơn đặt hàng này là từ dòng sản phẩm 1 và khách hàng này là nữ.

d Tìm $P(A \cup B)$, xác suất để khách hàng này là nữ và đơn đặt hàng này là từ dòng sản phẩm 1, hoặc cả hai.

e Tìm $P(B \mid A)$.

f Hãy chứng tỏ A và B là phụ thuộc và $P(AB) = P(A)P(B \mid A)$.

BẢNG **3.4**

	Dòng sản phẩm		
Giới tính	$1(B)$	$2(\overline{B})$	**Tổng số**
Nữ (A)	0,516	0,205	0,721
Nam $(\overline{A})$	0,132	0,147	0,279
Tổng	0,648	0,352	1,000

Lời giải $\overline{A}$ là biến cố khách hàng này là nam, và $\overline{B}$ là biến cố đơn đặt hàng này là từ dòng sản phẩm 2.

a $P(A) = 0{,}721$ là tổng số ở hàng thứ nhất

b $P(B) = 0{,}648$ là tổng số ở cột thứ nhất

c $P(AB)$, xác suất để khách hàng này là nữ và đơn đặt hàng này là từ dòng sản phẩm 1, là 0,516, đây là số liệu ở hàng thứ nhất và cột thứ nhất.

d Bằng việc sử dụng những số liệu trong Bảng 3.4

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(AB) \\ &= 0{,}721 + 0{,}648 - 0{,}516 = 0{,}853 \end{aligned}$$

e Để tìm $P(B \mid A)$, chúng ta sử dụng các mục ghi ở hàng thứ nhất để tìm

$$P(B \mid A) = \frac{P(AB)}{P(A)} = \frac{0{,}516}{0{,}721} = 0{,}716$$

f. Vì $P(B \mid A) \neq P(B)$ (xem các phần (b) và (e)), nên A và B là phụ thuộc. Ngoài ra,

$P(AB) = 0{,}516$ có thể thấy trực tiếp từ bảng này

trong khi

$$P(B \mid A)P(A) = (0{,}716)(0{,}721) = 0{,}516$$

cho nên $P(AB) = P(A)P(B \mid A)$ •

THÍ DỤ **3.10** Một công ty phát hiện rằng 85% số người được chọn cho chương trình thực tập của công ty đã hoàn tất khóa học. Trong số những người này, 60% đã trở thành những người bán hàng có hiệu quả, so với chỉ 10% trong số những người thực tập mà đã không hoàn tất chương trình thực tập.

a Xác suất để một người tham dự chương trình thực tập trở thành người bán hàng có hiệu quả là bao nhiêu?

b Nếu một người bán hàng, đã tham dự chương trình thực tập, được cho là có hiệu quả, thì xác suất để người này hoàn tất chương trình thực tập là bao nhiêu.

Lời giải Định rõ các biến cố sau đây:

A : người thực tập này hoàn tất chương trình thực tập

B : người thực tập này trở thành một người bán hàng có hiệu quả

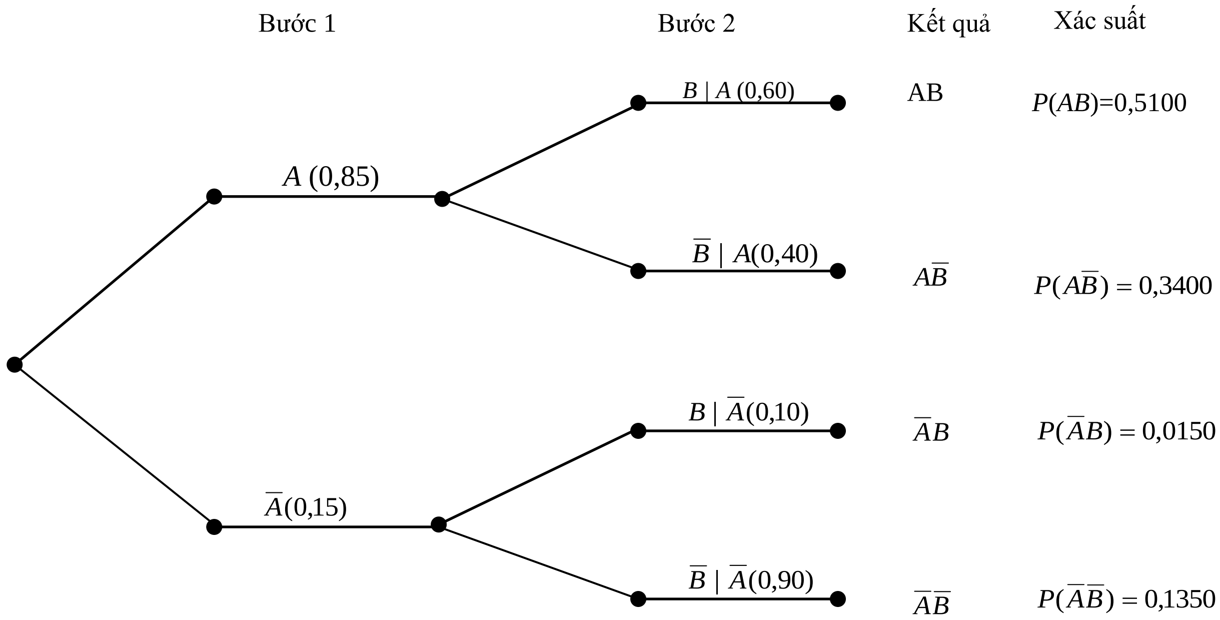
Sử dụng một giản đồ hình cây như trong Hình 3.2 để biểu hiện các xác suất thích hợp. Các xác suất ở bước thứ hai, được trình bày trong các ngoặc đơn, là có điều kiện, cho trước bước thứ nhất.

a Một người tham dự chương trình thực tập sẽ hoặc sẽ không hoàn tất chương trình này. Như thế, $P(B) = P(AB) + P(\bar{A}B) = 0{,}5100 + 0{,}0150 = 0{,}5250$

b Để tìm $P(A \mid B)$, chúng ta cần $P(AB)$ = 0,5100 và $P(B)$ = 0,5250 từ phần (a). Như thế,

$$P(A \mid B) = \frac{P(AB)}{P(B)} = \frac{0{,}5100}{0{,}5250} = 0{,}9714 \quad \bullet$$

HÌNH **3.2** Giản đồ hình cây cho Thí dụ 3.10



Những Lời Gọi ý về Giải Bài toán

Tính xác suất của một biến cố: phương pháp sự hợp thành-biến cố (event-composition approach)

- Hãy sử dụng những bước sau đây để tính xác suất của một biến cố bằng cách sử dụng phương pháp sự hợp thành-biến cố:
  - **1** Định rõ thí nghiệm.
  - **2** Hãy hình dung rõ ràng bản chất của các biến cố đơn. Hãy nhận dạng một ít biến cố đơn để làm rõ suy nghĩ của anh/chị.
  - **3** Khi có thể, hãy viết một phương trình thể hiện biến cố được quan tâm (ví dụ, *A*) như là một sự hợp thành của hai hay hơn hai biến cố, bằng việc sử dụng một trong hai, hoặc cả hai, dạng của sự hợp thành (hội và giao). Hãy lưu ý rằng, điều này đánh đồng các tập hợp điểm. Hãy đảm bảo rằng biến cố được hàm ý bởi sự hợp thành này và biến cố *A* đều thể hiện cùng một tập hợp các biến cố đơn.
  - **4** Hãy áp dụng các Quy tắc Cộng và Nhân Xác suất vào bước 3 và tìm *P*(*A*).
- Hãy cẩn thận với bước 3. Anh/Chị thường có thể cấu tạo nhiều sự hợp thành mà sự hợp thành nào cũng sẽ tương đương với biến cố *A*. Kỹ xảo là cấu tạo một sự hợp thành trong đó tất cả các xác suất xuất hiện trong bước 4 đều sẽ được biết. Hãy hình dung ra các kết quả của bước 4 đối với bất kỳ sự hợp thành nào và hãy chọn cái sự hợp thành mà đối với nó các xác suất thành phần đều được biết.
- Hãy luôn luôn viết ra những chữ cái biểu hiện các biến cố được mô tả trong một bài tập. Sau đó, hãy viết ra các xác suất được cho, và gắn chúng với các biến cố. Hãy xác định xác suất được yêu cầu trong bài tập. Điều này có thể giúp anh/chị đạt được sự hợp thành biến cố thích hợp.

BÀI TẬP

Các Kỹ thuật Căn bản

3.12 Một thí nghiệm có thể dẫn đến một trong năm biến cố đơn có khả năng xảy ra ngang nhau. Các biến cố *A, B,* và *C* bao gồm những biến cố đơn sau đây:

$$A: E_1, E_3 \qquad B: E_1, E_2, E_4, E_5 \qquad C: E_3, E_4$$

Hãy liệt kê các biến cố đơn trong những biến cố sau đây, và tìm xác suất riêng của chúng

a	S	**b**	A	**c**	B
d	C	**e**	$\overline{A}$	**f**	$\overline{B}$
g	AB	**h**	AC	**i**	$A \mid B$
j	$A \cup B$	**k**	$A \cup B$	**l**	$A \mid C$

3.13 Căn cứ vào Bài tập 3.12. Có phải các biến cố *A* và *B* loại trừ lẫn nhau? Độc lập? Hãy giải thích.

3.14 Một thí nghiệm tạo ra không gian mẫu có chứa tám biến cố đơn E_1, E_2, …, E_8, với $P(E_1) = 0{,}1$, $P(E_2) = P(E_3) = P(E_4) = 0{,}05$, $P(E_5) = 0{,}3$, $P(E_6) = 0{,}2$, $P(E_7) = 0{,}1$, và $P(E_8) = 0{,}15$. Các biến cố *A* và *B* là

$$A: \; E_1, E_4, E_6 \qquad B: \; E_3, E_4, E_5, E_6, E_7$$

- **a** Tìm $P(A)$.
- **b** Tìm $P(B)$.
- **c** Tìm $P(AB)$.
- **d** Tìm $P(A \cup B)$.
- **e** Có phải các biến cố *A* bà *B* loại trừ lẫn nhau? Hãy giải thích.
- **f** Tìm $P(\mathrm{A} \mid B)$.
- **g** Tìm $P(B \mid A)$.
- **h** Có phải *A* và *B* là các biến độc lập? Hãy giải thích.

3.15 Một thí nghiệm có thể dẫn đến biến cố *A*, với xác suất $P(A) = 0{,}2$, và biến cố *B*, với xác suất $P(B) = 0{,}6$. Xác suất của giao của *A* và *B* là $P(AB) = 0{,}15$.

- **a** Hãy lập một bảng xác suất cho thí nghiệm này.
- **b** Hãy sử dụng bảng xác suất trong phần (a) để tìm $P\left(\overline{A}B\right)$, $P\left(A \mid \overline{B}\right)$, và $P(A \cup B)$.
- **c** Có phải *A* và *B* là các biến độc lập? Hãy giải thích.

Ứng dụng

3.16 Những đơn đặt hàng mới, đối với các sản phẩm của một công ty, biến đổi về quy mô tính bằng đô la dựa theo phân phối xác suất sau đây.

Quy mô doanh số bán ($)	0−1000	1001−2000	2001−3000	3001−4000	4001−5000
Xác suất	0,10	0,35	0,25	0,20	0,10

a Tìm xác suất để một đơn đặt hàng mới sẽ vượt quá 2.000$.

b Tìm xác suất để một đơn đặt hàng mới sẽ là 2.000$ hay ít hơn, cho trước đơn đặt hàng này vượt quá 1.000$.

c Tìm xác suất để một đơn đặt hàng mới sẽ lớn hơn 3.000$, cho trước doanh số bán vượt quá 2.000$.

3.17 Các nhà nghiên cứu quảng cáo luôn luôn quan tâm muốn biết liệu người tiêu dùng có nhận thấy các mẫu quảng cáo của họ có thể tin được hay không, đặc biệt là khi các mẫu quảng cáo này chứa đựng sự ủng hộ của người nổi tiếng, các cuộc phỏng vấn có máy quay phim ngầm, hay những lời khẳng định rằng các sản phẩm "mới và được cải thiện". Dữ liệu trong bảng sau đây thể hiện một phần nhỏ của một cuộc điều tra do Tổ chức Roper tiến hành. Bảng này cho thấy tỷ lệ những người trưởng thành nhận thấy những mẫu quảng cáo có sự ủng hộ của người nổi tiếng là có thể tin được, được liệt kê dựa theo trình độ giáo dục của người trưởng thành.

	Thấp hơn Trung học	**Tốt nghiệp Trung học Phổ thông**	**Một số năm Đại học**	**Tốt nghiệp Đại học**
Có thể tin được	0,27	0,27	0,25	0,18

Nguồn: Tạp chí Nhân khẩu học Hoa Kỳ, tháng 12, 1990, trang 14.

Giả sử một người trưởng thành đơn lẻ được chọn ngẫu nhiêu từ trong số những người được liệt kê trong cuộc điều tra này.

a Nếu tỷ lệ những người tốt nghiệp đại học trong nhóm được điều tra là 0,24, thì xác suất để người trưởng thành được chọn sẽ là một người tốt nghiệp đại học và không tin mẫu quảng cáo này, là bao nhiêu?

b Nếu người trưởng thành được chọn đã học xong một số năm đại học, thì xác suất để người ấy không tin mẫu quảng cáo này là bao nhiêu?

c Nếu tỷ lệ những người trưởng thành trong nhóm được điều tra chưa bao giờ vào đại học là 0,4, thì xác suất để người trưởng thành được chọn sẽ là người chưa vào đại học và tin mẫu quảng cáo này là bao nhiêu?

[*Gợi ý*: những mục ghi trong bảng là các xác suất có điều kiện của việc tin vào mẫu quảng cáo này, cho trước trình độ giáo dục của người trưởng thành.]

3.18 Một cuộc điều tra gần đây về các doanh nghiệp Hoa Kỳ đã chỉ ra rằng, 27% các doanh nghiệp đều duy trì một chương trình nghỉ phép sinh đẻ (vợ sinh thì chồng cũng được nghỉ) cho các bậc cha mẹ của những đứa con mới sinh. Trong số những doanh nghiệp này, một phần ba cung cấp một kiểu tiếp tục trả tiền lương nào đó trong suốt thời gian nghỉ phép, và ba phần tư tiếp tục trợ cấp chăm sóc y tế (*Tạp chí về Nghề Kế toán*, tháng 11/1990, trang 23).

a Xác suất để một doanh nghiệp được chọn ngẫu nhiên cho nghỉ phép sinh đẻ, với một hình thức tiếp tục trả lương nào đó, là bao nhiêu?

b Xác suất để một doanh nghiệp được chọn ngẫu nhiên cho nghỉ phép sinh đẻ, mà không tiếp tục trợ cấp chi phí chăm sóc y tế, là bao nhiêu?

3.19 Một cuộc điều tra bao gồm 100 xe hơi, mỗi xe được phân loại dựa theo việc liệu nó đã có bộ phanh antilock (phanh giúp điều khiển xe khi ngừng xe thật nhanh) hay không và liệu nó đã dính dáng đến một tai nạn trong năm vừa qua hay không. Giả sử một trong những chiếc xe hơi này được chọn ngẫu nhiên để kiểm tra.

a Xác suất để chiếc xe hơi này đã dính dáng đến một tai nạn trong năm vừa qua là bao nhiêu?

b Xác suất để chiếc xe hơi này đã dính dáng đến một tai nạn và có bộ phanh antilock là bao nhiêu?

c Cho trước rằng chiếc xe hơi này đã dính dáng đến một tai nạn trong năm vừa qua, xác suất để nó có bộ phanh antilock là bao nhiêu?

Bảng cho Bài tập 3.19

	Bộ phanh antilock	**Không có bộ phanh antilock**
Tai nạn	0,03	0,12
Không tai nạn	0,40	0,45

Nguồn: Dữ liệu phỏng theo Tạp chí *Nghiên cứu của Người Tiêu dùng*, tháng 3/1994.

3.20 Một bài viết trong Tạp chí *Nghiên cứu của Người Tiêu dùng* ("Antilock Brakes", 1994) trình bày 43% của tất cả xe hơi kiểu 1993 đều có trang bị bộ phanh antilock. Giả sử ba chiếc xe hơi kiểu 1993 được chọn ngẫu nhiên để kiểm tra.

a Cho trước chiếc xe hơi thứ nhất được chọn có bộ phanh antilock, xác suất để chiếc xe hơi thứ hai được chọn có bộ phanh antilock là bao nhiêu?

b Có phải hai biến cố được mô tả trong phần (a) là độc lập? Hãy giải thích.

c Xác suất để cả ba chiếc xe hơi được chọn đều có bộ phân antilock là bao nhiêu?

d Xác suất để chỉ một trong ba chiếc xe hơi được chọn có bộ phanh antilock là bao nhiêu?

3.21 Một cuộc điều tra đã được tiến hành để đánh giá tác động của siêu xa lộ thông tin đối với các doanh nghiệp ở Hoa Kỳ. Dựa trên cuộc điều tra các nhà điều hành tiếp thị cao cấp này, 40% nói họ sẽ sử dụng siêu xa lộ thông tin để tương tác trực tiếp với khách hàng, 36% nói họ sẽ không, và 24% trả lời không biết.

a Nếu một nhà điều hành tiếp thị cao cấp được chọn ngẫu nhiên, thì xác suất để nhà điều hành này sẽ sử dụng siêu xa lộ thông tin để tương tác trực tiếp với khách hàng là bao nhiêu?

b Nếu hai nhà điều hành tiếp thị cao cấp được chọn ngẫu nhiêu, thì xác suất để chỉ một nhà điều hành sẽ sử dụng siêu xa lộ thông tin để tương tác trực tiếp với khách hàng là bao nhiêu?

3.22 Một bài viết trong *Báo cáo Người Tiêu dùng* ("Xếp hạng: Sơn Nội thất Latex," 1994) đã xếp hạng 35 nhãn hiệu sơn nội thất latex bằng cách sử dụng sự phân loại định tính được trình bày dưới đây.

Thứ hạng	**Số lượng Nhãn hiệu**
Tuyệt hảo	2
Rất tốt	21
Tốt	11
Trung bình	1
Kém	0

Giả sử một người tiêu dùng chọn một trong 35 nhãn hiệu này một cách ngẫu nhiên.

a Xác suất để người tiêu dùng này chọn một nhãn hiệu "tuyệt hảo" là bao nhiêu?

b Xác suất để người tiêu dùng này chọn một nhãn hiệu được xếp hạng ít ra là "tốt" là bao nhiêu?

c Xác suất để người tiêu dùng này chọn một nhãn hiệu *không* được xếp hạng "rất tốt" hoặc "tuyệt hảo" là bao nhiêu?

d Nếu người tiêu dùng này chọn hai nhãn hiệu *khác nhau* để so sánh, thì xác suất để cả hai nhãn hiệu này đều được xếp hạng "rất tốt" là bao nhiêu?

3.23 Một hệ thống vòi cứu hỏa sử dụng trên các cao ốc thương mại được thiết kế để cho mỗi vòi cứu hỏa có thể được kích hoạt thông qua hai thiết bị độc lập. Vòi cứu hỏa sẽ hoạt động khi một trong hai thiết bị này (hoặc cả hai) được kích hoạt. Độ tin cậy của thiết bị thứ nhất (xác suất để nó được kích hoạt khi đạt đến nhiệt độ nhất định) là 0,91, trong khi độ tin cậy của thiết bị thứ hai là 0,95. Xác suất để vòi cứu hỏa này sẽ hoạt động đúng khi đạt đến một nhiệt độ nhất định là bao nhiêu?

3.5 Quy tắc Bayes và Xác suất có Điều kiện (Tùy chọn)

Các xác suất có điều kiện cho phép chúng ta cập nhật xác suất bằng cách dùng thông tin khi nó trở nên có sẵn. Thí dụ, chúng ta có thể biết tiên nghiệm, hay trước khi sự thật xảy ra, rằng tất cả các đơn vị đầu vào sử dụng trong một quy trình sản xuất là do ba nhà cung ứng khác nhau cung cấp, và rằng nhà cung ứng S_1 cung cấp 20% những đơn vị đầu vào này, nhà cung ứng S_2 cung cấp 30%, và nhà cung ứng S_3 cung cấp 50% còn lại. Vì vậy, nếu không có sẵn thông tin nào khác và một đơn vị đầu vào được chọn ngẫu nhiên từ quy trình sản xuất này, thì xác suất để đơn vị đầu vào này đến từ nhà cung ứng S_1 là 0,20, từ nhà cung ứng S_2 là 0,30, và từ nhà cung ứng S_2 là 0,50.

Từ kết quả thực hiện trong quá khứ, các tỷ lệ phần trăm đơn vị đầu vào có khuyết tật được cung cấp bởi ba nhà cung ứng này là S_1 : 0,05, S_2 : 0,02 và S_3 : 0,01. Nếu một đơn vị đầu vào được chọn ngẫu nhiên từ quy trình sản xuất này được phát hiện là có khuyết tật, thì xác suất để đơn vị đầu vào này được cung cấp bởi S_1 là bao nhiêu? bởi S_2 là bao nhiêu? bởi S_3 là bao nhiêu? Nghĩa là, thông tin này làm thay đổi ra sao các xác suất *tiên nghiệm* (*prior* probabilities) để đơn vị đầu vào này đến từ một trong ba nhà cung ứng?

Nếu D là biến cố một đơn vị đầu vào có khuyết tật được quan sát thấy, thì bất kỳ một trong ba nhà cung ứng đều có thể đã cung cấp đơn vị đầu vào này. Chúng ta mong muốn cập nhật các xác suất *tiên nghiệm*, $P(S_i)$, bằng cách sử dụng các xác suất có điều kiện dưới dạng $P(S_i|D)$. Thông tin liên quan được cho trong giản đồ hình cây trong Hình 3.3. Các xác suất ở bước thứ hai là có điều kiện, trên cơ sở bước thứ nhất. Thí dụ, xác suất để một đơn vị đầu vào có khuyết tật, cho trước nhà cung ứng S_1, là $P(D \mid S_1) = 0,05$.

Giả sử chúng ta mong muốn tìm các xác suất *hậu suy/hậu nghiệm(posterior probabilities*) đã được cập nhật với thông tin của mẫu. Thí dụ, để tìm $P(S_1 \mid D)$, chúng ta có

$$P(S_1 \mid D) = \frac{P(S_1 D)}{P(D)}$$

Từ giản đồ hình cây trong Hình 3.3, $P(D)$ là tổng số của ba kết quả dẫn đến D. Như thế,

$$\begin{aligned} P(D) &= P(S_1 D) + P(S_2 D) + P(S_3 D) \\ &= P(S_1)P(D \mid S_1) + P(S_2)P(D \mid S_2) + P(S_3)P(D \mid S_3) \\ &= 0{,}010 + 0{,}006 + 0{,}005 = 0{,}021 \end{aligned}$$

Vì $P(S_1D) = P(S_1)P(D \mid S_1) = 0{,}20(0{,}05) = 0{,}10$, do đó

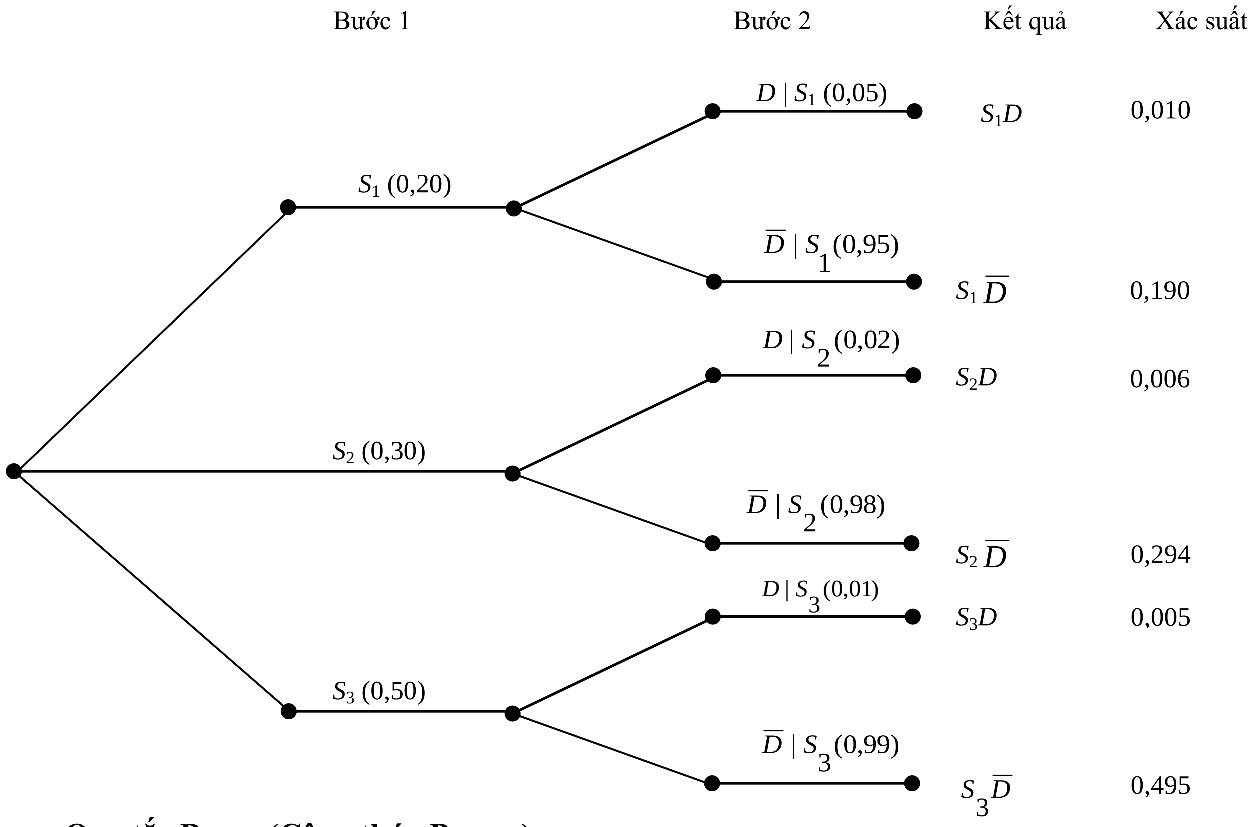
$$P(S_1 \mid D) = \frac{0{,}010}{0{,}021} = 0{,}476$$

Hơn nữa, hai xác suất hậu suy còn lại là

$$P(S_2 \mid D) = \frac{0{,}006}{0{,}021} = 0.286 \quad \text{và} \quad P(S_3 \mid D) = \frac{0.005}{0{,}021} = 0{,}238$$

Bằng viêc sử dụng các xác suất *tiên nghiệm*, một đơn vị đầu vào nhiều khả năng đã được cung ứng bởi nhà cung ứng S_3. Bằng việc sử dụng các xác suất *hậu suy*, đơn vị đầu vào được chọn làm mẫu nhiều khả năng đã đến từ nhà cung ứng S_1.

Phương pháp tính toán xác suất hậu suy thường được gọi là *Quy tắc Bayes*, đặt theo tên nhà toán học Anh, Thomas Bayes.



Quy tắc Bayes (Công thức Bayess)

Cho S_1, S_2, …, S_k biểu hiện k trạng thái của tự nhiên (states of nature) có thể có duy nhất và loại trừ lẫn nhau, với các xác suất tiên nghiệm $P(S_1)$, $P(S_2)$, …, $P(S_k)$. Nếu một biến cố A xảy ra, thì xác suất hậu suy của S_i trên cơ sở A đã xảy ra là xác suất có điều kiện.

$$P(S_i \mid A) = \frac{P(S_i)P(A \mid S_i)}{\sum_{j=1}^{k} P(S_j)P(A \mid S_j)}$$

với $i = 1, 2, \ldots, k$.

THÍ DỤ **3.11** Một cửa hàng bách hóa đang xem xét việc thực hiện một chính sách quản lý tín dụng mới, trong nỗ lực nhằm cắt giảm số lượng khách hàng tín dụng không trả các khoản thanh toán nợ của họ. Bà giám đốc tín dụng đã đề nghị rằng trong tương lai, phải đình chỉ tín dụng đối với bất kỳ khách hàng nào đã có hai lần trễ hạn một tuần hay lâu hơn, trong việc trả góp nợ hàng tháng. Bà hỗ trợ lời yêu cầu của mình bằng cách lưu ý rằng, hồ sơ tín dụng trong quá khứ cho thấy 90% trong tất cả những người không trả các khoản thanh toán nợ của họ đều đã trễ hạn đối với ít nhất là hai khoản thanh toán hàng tháng.

Giả sử từ việc điều tra riêng của mình, chúng ta tìm ra rằng có 2% trong tất cả khách hàng tín dụng đã thật sự không trả các khoản thanh toán nợ và 45% số người không quịt nợ thì đã có ít nhất hai khoản thanh toán hàng tháng trễ hạn. Hãy tìm xác suất để một khách hàng có hai hay nhiều hơn các khoản thanh toán trễ hạn sẽ thật sự không trả những khoản thanh toán của mình, và, dưới ánh sáng của xác suất này, hãy phê phán kế hoạch tín dụng của giám đốc tín dụng.

Lời giải Định rõ các biến cố L và D như sau:

biến cố L : một khách hàng tín dụng bị trễ hạn hai tuần hay nhiều hơn với ít nhất hai khoản thanh toán hàng tháng

biến cố D : một khách hàng tín dụng không trả các khoản thanh toán nợ của mình

và cho $\overline{D}$ biểu thị phần bù của biến cố D. Chúng ta tìm xác suất có điều kiện

$$P(D \mid L) = \frac{P(DL)}{P(L)} = \frac{P(L \mid D)P(D)}{P(L \mid D)P(D) + P(L \mid \overline{D})P(\overline{D})}$$

Từ thông tin được cho trong nội dung mô tả bài toán, chúng ta tìm ra rằng

$$P(D \mid L) = \frac{(0{,}90)(0{,}02)}{(0{,}90)(0{,}02) + (0{,}45)(0{,}98)} = \frac{0{,}0180}{0{,}0180 + 0{,}4410} = 0{,}0392$$

Vì thế, nếu như kế hoạch của giám đốc tín dụng được thực hiện, thì xác suất chỉ vào khoảng 0,04—hay vào khoảng 1 trong 25—để một khách hàng, mà bị mất những đặc quyền về tín dụng của mình, lẽ ra đã thật sự không trả các khoản thanh toán nợ của mình. Kế hoạch của giám đốc tín dụng sẽ là một chính sách kinh doanh kém, trừ khi ban giám đốc cho là đáng bỏ công phát hiện một khách hàng không trả nợ triển vọng dù phải chịu phí tổn là mất đi 24 khách hàng tín dụng tốt. •

BÀI TẬP

Các Kỹ thuật Căn bản

3.24 Giả sử ba trạng thái của tự nhiên loại trừ lẫn nhau, S_1, S_2, S_3, có thể tồn tại với các xác suất

$$P(S_1) = 0{,}4 \qquad P(S_2) = 0{,}5 \qquad P(S_3) = 0{,}1$$

và trên cơ sở những trạng trái của tự nhiên này, một biến cố I có thể xảy ra với các xác suất

$$P(I \mid S_1) = 0{,}1 \qquad P(I \mid S_2) = 0{,}3 \qquad P(I \mid S_3) = 0{,}2$$

Cho trước biến cố I được quan sát thấy, hãy tìm $P(S_1 \mid I)$, $P(S_2 \mid I)$, và $P(S_3 \mid I)$.

3.25 Giả sử bốn trạng thái của tự nhiên loại trừ lẫn nhau, S_1, S_2, S_3, và S_4 có thể tồn tại với các xác suất

$$P(S_1)=0,1 \qquad P(S_2)=0,4 \qquad P(S_3)=0,3 \qquad P(S_4)=0,2$$

và trên cơ sở những trạng trái của tự nhiên này, một biến cố I có thể xảy ra với các xác suất

$$P(I\mid S_1)=0,6 \qquad P(I\mid S_2)=0,2 \qquad P(I\mid S_3)=0,2 \qquad P(I\mid S_4)=0,5$$

Cho trước biến cố I được quan sát thấy, hãy tìm $P(S_1\mid I)$, $P(S_2\mid I)$, $P(S_3\mid I)$, và $P(S_4\mid I)$.

Ứng dụng

3.26 Khi các đơn vị sản phẩm đến tận cuối dây chuyền sản xuất, người kiểm tra chọn những đơn vị sản phẩm nào phải đi qua một cuộc kiểm tra toàn diện. Mười phần trăm trong tất cả các đơn vị sản phẩm được sản xuất ra có khuyết tật. Sáu mươi phần trăm trong tất cả các đơn vị sản phẩm có khuyết tật đi qua cuộc kiểm tra toàn diện, và 20% trong tất cả đơn vị sản phẩm tốt đi qua cuộc kiểm tra toàn diện. Cho trước một đơn vị sản phẩm được kiểm tra toàn diện, xác suất nó có khuyết tật là bao nhiêu?

3.27 Với những đe dọa gần đây của chủ nghĩa khủng bố, các sân bay ngày càng lo lắng về việc phát hiện vũ khí tại cổng đón hành khách lên máy bay. Các tỷ lệ phát hiện đối với những vũ khí mang trên người hay trong hành lý xách tay của họ phải cực kỳ cao. Ở một thành phố cụ thể, sân bay A xử lý 50% trong toàn bộ vận tải đường không, trong khi các sân bay B và C xử lý lần lượt là 30% và 20%. Các tỷ lệ phát hiện tại sân bay A là 0,99, sân bay B là 0,95, và sân bay C là 0,80. Nếu một hành khách tại một trong các sân bay này bị phát hiện đang mang theo vũ khí qua cổng lên máy bay, thì xác suất hành khách này đang sử dụng sân bay A là bao nhiêu? sân bay C là bao nhiêu?

3.28 Giả sử 50% trong tất cả những người điền vào mẫu khai thuế thu nhập dài đều cố tìm được những khoản khấu trừ mà họ biết là bất hợp pháp và 2% nữa sẽ liệt kê không đúng các khoản khấu trừ, do thiếu hiểu biết về các quy định về thuế thu nhập. Trong số 5% phạm tội gian lận, có 80% sẽ không chịu nhận là mình biết rõ về sai lầm này khi bị đối chất bởi nhà điều tra. Nếu một người điền mẫu khai thuế dài đó bị đối chất với một khoản khấu trừ không chính đáng và người ấy không chịu nhận là mình biết rõ về sai lầm này, thì xác suất người ấy phạm tội là bao nhiêu?

3.6 Những Biến Ngẫu nhiên Rời rạc và Phân phối Xác suất của Chúng

Trong Chương 2, chúng ta đã định nghĩa *biến* là một đặc trưng thay đổi hay biến đổi theo thời gian và/hoặc đối với các cá nhân hay các đối tượng khác nhau được xem xét. Các *biến định lượng* dẫn đến những giá trị đo lường bằng số, trong khi đó các *biến định tính* dẫn đến những giá trị đo lường phân loại. Tuy nhiên, ngay cả biến định tính cũng có thể dẫn đến giá trị đo lường bằng số, nếu chúng ta đếm số lượng các cá nhân hay những phần tử trong mỗi loại đã định rõ.

Nếu chúng ta biểu thị biến được đo lường bằng x, thì giá trị mà x nhận sẽ thay đổi hay biến đổi, phụ thuộc vào kết quả cụ thể của thí nghiệm. Thí dụ, giả sử chúng ta cho x biểu thị sản lượng hàng ngày trong một nhà máy chế tạo công nghiệp. Biến x có thể nhận một số giá trị khác nhau phụ thuộc vào ngày mà chúng ta đo lường x. Tương tự, doanh số bán khóa sổ mỗi ngày do một người bán thực hiện là một biến định lượng x, mà giá trị của nó phụ thuộc vào ngày lấy giá trị đo lường. Nếu giá trị của một biến *x phụ thuộc vào kết quả ngẫu nhiên của một thí nghiệm*, thì chúng ta gọi biến *x* là **biến ngẫu nhiên** (**random variable**)

ĐỊNH NGHĨA • Biến *x* là **biến ngẫu nhiên** nếu giá trị nó nhận được, tương ứng với kết quả của một thí nghiệm, là một biến cố ngẫu nhiên. •

Việc quan sát số lượng khuyết tật trên một món đồ đạc được chọn ngẫu nhiên, việc chọn lựa ngẫu nhiên một người nộp đơn xin vào đại học và việc quan sát điểm Sát hạch Khả năng Học tập (SAT) của một người, và việc đo lường số lượng cuộc gọi điện thoại nhận được bởi một đường dây nóng can thiệp vào lúc nguy kịch trong suốt một thời kỳ được chọn ngẫu nhiên, tất cả đều dẫn đến những biến cố bằng số ngẫu nhiên.

Các biến ngẫu nhiên định lượng được phân loại thành *rời rạc* hoặc *liên tục* dựa theo giá trị mà *x* có thể nhận. Sự phân biệt giữa các biến ngẫu nhiên rời rạc và liên tục là quan trọng, bởi vì các mô hình xác suất khác nhau cần được sử dụng đối với mỗi loại biến. Chúng ta sẽ tập trung chú ý vào biến ngẫu nhiên rời rạc trong phần còn lại của chương này. Biến ngẫu nhiên liên tục là chủ đề của Chương 5.

ĐỊNH NGHĨA • **Phân phối xác suất (probability distribution)** của một biến ngẫu nhiên rời rạc là một công thức, bảng hay đồ thị mà cung cấp $p(x)$, xác suất gắn với mỗi một trong các giá trị của *x*. •

Các biến cố gắn với các giá trị khác nhau của *x* không thể trùng lắp, bởi vì một và chỉ một giá trị của *x* được chỉ định cho mỗi biến cố đơn; như thế các giá trị của *x* biểu thị những biến cố bằng số loại trừ lẫn nhau. Tính tổng số $p(x)$ đối với tất cả các giá trị của *x* bằng tổng số của các xác suất của tất cả các biến cố đơn, và do vậy sẽ bằng 1.

Những Yêu cầu đối với một Phân phối Xác suất Rời rạc

1 $0 \le p(x) \le 1$

2 $\sum_{x} p(x) = 1$

THÍ DỤ **3.12** Xét một thí nghiệm bao gồm việc tung hai đồng xu, và cho *x* bằng số lượng mặt ngửa quan sát được. Hãy tìm phân phối xác suất của *x*.

Lời giải Đối với thí nghiệm này, các biến cố đơn cùng với xác suất tương ứng của chúng được trình bày như sau:

Biến cố Đơn	**Đồng xu 1**	**Đồng xu 2**	$P(E_i)$	x
E_1	H	H	1/4	2
E_2	H	T	1/4	1
E_3	T	H	1/4	1
E_4	T	T	1/4	0

Vì E_1 đi đôi với biến cố đơn "quan sát một mặt ngửa trên đồng xu 1 và một mặt ngửa trên đồng xu 2," nên chúng ta gán cho nó giá trị $x = 2$. Tương tự, chúng ta gán $x = 1$ cho biến cố E_2, và v.v.. Chúng ta có thể tính xác suất của mỗi giá trị của x bằng cách cộng các xác suất của các biến cố đơn trong biến cố bằng số đó. Biến cố bằng số $x = 0$ chứa đựng một biến cố đơn, E_4; biến cố $x = 1$ chứa đựng hai biến cố đơn, E_2 và E_3; và $x = 2$ chứa đựng một biến cố đơn, E_1. Các giá trị của x cùng với xác suất tương ứng của chúng được trình bày trong Bảng 3.5.

Nhận thấy rằng

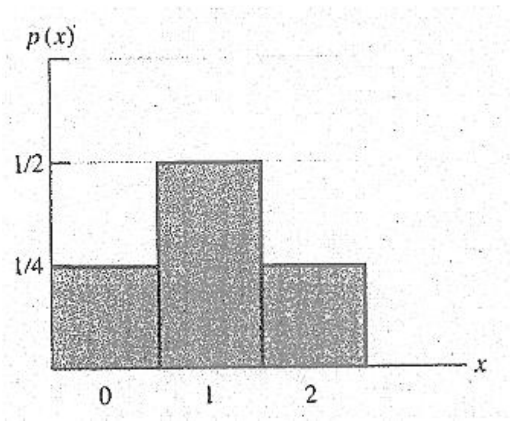
$$\sum_{x=0}^{2} p(x) = 1$$

BẢNG **3.5**
Phân phối xác suất của x
(x = số mặt ngửa)

x	**Các Biến cố Đơn trong x**	$p(x)$
0	E_4	¼
1	E_2, E_3	½
2	E_1	¼
	$\sum_{x=0}^{2} p(x) =$	1

Phân phối xác suất trong Bảng 3.5 có thể được trình bày bằng đồ thị dưới dạng biểu đồ tần suất tương đối (xem Mục 2.4 trong nguyên bản tiếng Anh)†. Biểu đồ tần suất của biến ngẫu nhiên x sẽ có ba lớp, tương ứng với $x = 0$, $x = 1$, và $x = 2$. Bởi vì $p(0) = 1/4$, nên tần suất tương đối trên lý thuyết đối với $x = 0$ là 1/4; $p(1) = 1/2$, và như thế tần suất tương đối trên lý thuyết đối với $x = 1$ là 1/2. Hình 3.4 trình bày biểu đồ tần suất này.

HÌNH **3.4**
Biểu đồ xác suất hình thanh của $p(x)$ cho Thí dụ 3.12



Có lẽ anh/chị đã chú ý sự tương tự giữa phân phối xác suất của một biến ngẫu nhiên rời rạc và biểu đồ tần suất tương đối được thảo luận trong Chương 2 (trong nguyên bản tiếng Anh). Sự khác biệt là các biểu đồ tần suất tương đối được xây dựng cho một *mẫu* gồm *n* giá trị đo lường rút ra từ tổng thể, trong khi đó biểu đồ xác suất được xây dựng như là một mô hình cho *toàn bộ tổng thể* các giá trị đo lường. Giống như chúng ta tính số

† Phân phối xác suất trong Bảng 3.5 cũng có thể được trình bày bằng cách sử dụng một công thức, được cho trong Mục 4.2.

trung bình và độ lệch chuẩn cho một mẫu gồm *n* giá trị đo lường để đo lường vị trí và độ biến thiên của phân phối tần suất tương đối, chúng ta cũng có thể tính một số trung bình và độ lệch chuẩn để mô tả phân phối xác suất cho một biến ngẫu nhiên. Trung bình tổng thể, đo lường giá trị trung bình của *x* trong tổng thể, còn được gọi là **giá trị kỳ vọng** (**expected value**) của biến ngẫu nhiên *x*.

Phương pháp tính toán trung bình tổng thể hay giá trị kỳ vọng của một biến ngẫu nhiên có thể dễ hiểu hơn bằng cách xem xét một thí dụ. Cho *x* bằng số lượng mặt ngửa quan sát được trong việc tung hai đồng xu. Để cho thuận tiện, *p*(*x*) được cho như sau

x	0	1	2
$p(x)$	1/4	1/2	1/4

Giả sử thí nghiệm được lặp lại một số lượng lớn lần—chẳng hạn như, *n* = 4.000.000 lần. Theo trực giác, chúng ta kỳ vọng quan sát được xấp xỉ 1 triệu con số 0, 2 triệu con số 1, và 1 triệu con số 2. Như thế giá trị trung bình của *x* sẽ bằng

$$\frac{\text{Tổng các giá trị đo lường}}{n} = \frac{1.000.000(0) + 2.000.000(1) + 1.000.000(2)}{4.000.000}$$

$$= \frac{1.000.000(0)}{4.000.000} + \frac{2.000.000(1)}{4.000.000} + \frac{1.000.000(2)}{4.000.000}$$

$$= (1/4)(0) + (1/2)(1) + (1/4)(2)$$

Lưu ý rằng số hạng đầu tiên trong tổng này bằng (0)*p*(0), số hạng thứ hai bằng (1)*p*(1), và số hạng thứ ba bằng (2)*p*(2). Như thế, giá trị trung bình của *x* là

$$\sum_{x=0}^{2} xp(x) = 1$$

Kết quả này không phải là tình cờ, và nó mang lại sự biện minh theo trực giác cho định nghĩa về giá trị kỳ vọng của một biến ngẫu nhiên rời rạc *x*.

ĐỊNH NGHĨA • Gọi *x* là môt biến ngẫu nhiên rời rạc với phân phối xác suất *p*(*x*). Số trung bình hay **giá trị kỳ vọng của *x*** (**expected value of *x***) được cho như sau:

$$\mu = E(x) = \sum_{x} xp(x)$$

trong đó các phần tử được tính tổng số đối với mọi giá trị của biến ngẫu nhiên *x*. •

THÍ DỤ **3.13** Trong một cuộc xổ số được tiến hành để giúp quỹ từ thiện địa phương, 10.000 vé số sẽ được bán với giá 5USD mỗi vé. Giải thưởng là một chiếc xe ôtô trị giá 12.000USD. Nếu anh/chị mua hai vé số, thì khoản lợi (gain) kỳ vọng của anh/chị là bao nhiêu?

Lời giải Khoản lợi *x* của anh/chị có thể nhận một trong hai giá trị. Anh/Chị sẽ hoặc là mất 10USD (nghĩa là khoản lợi sẽ là −10USD) hoặc là nhận được 11.990USD, với các xác suất lần lượt là 0,9998 và 0,0002. Phân phối xác suất đối với khoản lợi *x* như sau:

x	$p(x)$
−10$	0,9998
11.990$	0,0002

Khoản lợi kỳ vọng sẽ là

$$E(x) = \sum_x xp(x)$$
$$= (-10\$)(0{,}9998) + (11.990\$)(0{,}0002)$$
$$= -7{,}60\$$$

Hãy nhớ lại rằng giá trị kỳ vọng của x là số trung bình của tổng thể trên lý thuyết mà sẽ được tạo ra nếu cuộc xổ số được lặp lại một số lượng lớn vô hạn lần. Nếu như thế, thì khoản lợi trung bình hay kỳ vọng mỗi lần xổ số sẽ là một khoản lợi âm (khoản lỗ) 7,6USD. •

THÍ DỤ **3.14** Hãy xác định phí bảo hiểm hàng năm cho một hợp đồng bảo hiểm có giá trị 1000USD và bảo hiểm một biến cố đã xảy ra với tỷ lệ 2 lần trong 100, trong suốt một thời kỳ dài. Gọi x bằng khoản lợi tài chính hàng năm của công ty bảo hiểm do việc bán hợp đồng bảo hiểm này, và gọi C bằng mức phí bảo hiểm hàng năm chưa biết. Chúng ta sẽ tính giá trị của C sao cho khoản lợi kỳ vọng, $E(x)$, sẽ bằng 0. Như thế C là mức phí bảo hiểm cần thiết để được hòa vốn. Công ty sẽ cộng thêm chi phí hành chính và lợi nhuận vào con số này.

Lời giải Bước thứ nhất trong lời giải là xác định các giá trị mà khoản lợi x có thể nhận và sau đó xác định $p(x)$. Nếu biến cố này không xảy ra trong suốt năm đang xét, thì công ty bảo hiểm sẽ nhận được phí bảo hiểm $x = C$ đô la. Nếu biến cố này thực sự xảy ra, thì khoản lợi sẽ âm. Cụ thể là, công ty sẽ mất 1000USD trừ đi phí bảo hiểm là C đô la đã thu. Như thế, $x = -(1000 - C)$ đô la. Các xác suất đi đôi với hai giá trị này của x lần lượt là 98/100 và 2/100. Phân phối xác suất đối với khoản lợi này là như sau:

x = khoản lợi (gain)	$p(x)$
C	$\frac{98}{100}$
$-(1.000 - C)$	$\frac{2}{100}$

Bởi vì chúng ta muốn phí bảo hiểm C sao cho, trong dài hạn (đối với nhiều hợp đồng bảo hiểm tương tự), khoản lợi trung bình sẽ bằng 0, nên chúng ta sẽ ấn định giá trị kỳ vọng của x bằng 0 và giải phương trình để tìm C. Như thế,

$$E(x) = \sum_x xp(x)$$
$$= C\left(\frac{98}{100}\right) + [-(1.000 - C)]\left(\frac{2}{100}\right) = 0$$

hoặc

$$\frac{98}{100}C+\left(\frac{2}{100}\right)C-20=0$$

Giải phương trình này để tìm *C*, chúng ta có

$$C = 20\$$$

Như thế, nếu mà công ty bảo hiểm này tính mức phí bảo hiểm hàng năm là 20USD, thì khoản lợi được tính cho một số lượng lớn những hợp đồng bảo hiểm tương tự sẽ bằng 0. Mức phí bảo hiểm thực sự sẽ bằng 20$ cộng với các chi phí hành chính và lợi nhuận. •

Trong Chương 2, chúng ta đã định nghĩa phương sai của tổng thể σ^2 là trung bình của bình phương các độ lệch của những giá trị đo lường so với trung bình của chúng. Bởi vì có sự kỳ vọng thì tương đương với "tính trung bình," nên chúng ta định nghĩa **phương sai** và độ lệch chuẩn như sau.

ĐỊNH NGHĨA • Gọi *x* là một biến ngẫu nhiên rời rạc với phân phối xác suất $p(x)$ và giá trị kỳ vọng $E(x) = \mu$. **Phương sai của *x* (variance of x)** là

$$\sigma^2 = E[(x-\mu)^2] = \sum_x (x-\mu)^2 p(x)$$

trong đó việc tính tổng số được thực hiện đối với mọi giá trị của biến ngẫu nhiên *x*.[†] •

ĐỊNH NGHĨA • **Độ lệch chuẩn σ của một biến ngẫu nhiên** bằng căn bậc hai của phương sai của nó. •

THÍ DỤ **3.15** Gọi *x* là một biến ngẫu nhiên với phân phối xác suất được cho trong bảng sau đây:

x	0	1	2	3	4	5
$p(x)$	0,10	0,40	0,20	0,15	0,10	0,05

Hãy tìm μ, σ^2 và σ. Vẽ đồ thị của $p(x)$ và xác định vị trí khoảng $\mu \pm 2\sigma$ trên đồ thị này. Xác suất để *x* sẽ nằm trong khoảng $\mu \pm 2\sigma$ là bao nhiêu?

Lời giải

[†] Có thể cho thấy (bỏ qua việc chứng minh) rằng

$$\sigma^2 = \sum_x (x-\mu)^2 p(x) = \sum_x x^2 p(x) - \mu^2$$

Kết quả này tương đồng với công thức đi tắt của s^2 được cho trong Chương 2.

$$\begin{aligned}
\mu &= E(x) = \sum_{x=0}^{5} xp(x) \\
&= (0)(0{,}10) + (1)(0{,}40) + \ldots + (4)(0{,}10) + (5)(0{,}05) \\
&= 1{,}90
\end{aligned}$$

$$\begin{aligned}
\sigma^2 &= E[x-\mu]^2 = \sum_{x=0}^{5} (x-\mu)^2 p(x) \\
&= (0-1{,}9)^2(0{,}10) + (1-1{,}9)^2(0{,}40) + \ldots + (5-1{,}9)^2(0{,}05) \\
&= 1{,}79
\end{aligned}$$

và

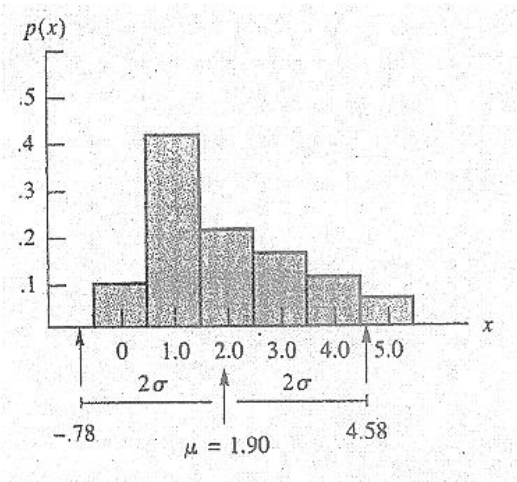
$$\sigma = \sqrt{\sigma^2} = \sqrt{1{,}79} = 1{,}34$$

Khoảng ($\mu \pm 2\sigma$) là [1,90 ± (2) (1,34)], hay − 0,78 đến 4,58.

Đồ thị của $p(x)$ và khoảng $\mu \pm 2\sigma$ được trình bày trong Hình 3.5. Anh/Chị có thể thấy x = 0, 1, 2, 3, 4 nằm trong khoảng này. Vì thế,

$$\begin{aligned}
p(\mu - 2\sigma < x < \mu + 2\sigma) &= p(0) + p(1) + \ldots + p(4) \\
&= 0{,}10 + 0{,}40 + 0{,}20 + 0{,}15 + 0{,}10 \\
&= 0{,}95 \quad \bullet
\end{aligned}$$

HÌNH **3.5**
Biểu đồ xác suất hình thanh của $p(x)$ đối với Thí dụ 3.5



Những lời Gợi ý về Giải Bài toán

1 Để tìm giá trị kỳ vọng của một biến ngẫu nhiên rời rạc *x*, hãy lập một bảng có ba cột, cột thứ nhất dành cho *x* và cột thứ hai dành cho $p(x)$. Sau đó, hãy nhân mỗi giá trị của *x* với xác suất tương ứng của nó và ghi các kết quả vào cột thứ ba. Tổng số của cột thứ ba này, tức là tổng số của $xp(x)$, sẽ cho anh/chị giá trị kỳ vọng của *x*.

2 Để tìm σ^2 cho một biến ngẫu nhiên rời rạc *x*, hãy bắt đầu bằng một bảng có bốn cột, cột thứ nhất dành cho *x*, cột thứ hai dành cho $(x - \mu)^2$, cột thứ ba dành cho $p(x)$, và cột thứ tư dành cho tích chéo (cross products), $(x-\mu)^2\, p(x)$. Trước hết, hãy tính giá trị của $(x-\mu)^2$ cho mỗi giá trị của x và ghi các kết quả vào cột 2. Kế đó, tính tích chéo $(x-\mu)^2\, p(x)$ cho mỗi giá trị của x, và ghi các kết quả vào cột 4. Tổng số của cột 4 sẽ cho anh/chị giá trị của σ^2.

Bài tập

Các Kỹ thuật Căn bản

3.29 Một biến ngẫu nhiên có phân phối xác suất sau đây:

x	0	1	2	3	4	5
$p(x)$	0,1	0,3	0,4	0,1	?	0,05

a Tìm $p(4)$

b Xây dựng một biểu đồ xác suất để mô tả $p(x)$.

3.30 Căn cứ vào Bài tập 3.29.

a Xác suất của $x = 2$ là bao nhiêu? ít nhất là bằng 2 là bao nhiêu?

b Xác suất để x không lớn hơn 3 là bao nhiêu?

c Tìm $\mu = E(x)$ và σ^2.

3.31 Phân phối nào trong các phân phối sau đây không thể hiện một phân phối xác suất. Hãy giải thích.

a

x	$p(x)$
1,5	0,60
2,0	0,25
6,5	0,15

b

x	$p(x)$
0	−1
1	2
2	3
3	4

c

x	$p(x)$
1	0,2
2	0,4
3	0,2

3.32 Cho x là một biến ngẫu nhiên rời rạc với phân phối xác suất được cho trong bảng sau đây:

x	1	2	3	4	5	6	7
$p(x)$	0,05	0,2	0,35	0,2	0,1	0,05	0,05

a Tìm μ, σ^2, và σ.

b Xây dựng biểu đồ xác suất hình thanh cho $p(x)$.

c Xác định vị trí khoảng $(\mu \pm 2\sigma)$ trên trục x của biểu đồ này. Xác suất để x sẽ nằm trong khoảng này là bao nhiêu?

d Nếu mà anh/chị chọn một số rất lớn các giá trị của x từ tổng thể này, có phải hầu hết sẽ nằm trong khoảng $(\mu \pm 2\sigma)$? Hãy giải thích.

Ứng dụng

3.33 Một viên kim cương trị giá 50.000USD được bảo hiểm cho toàn bộ giá trị của nó, bằng cách trả một mức phí bảo hiểm *D* đô la. Nếu xác suất của vụ trộm trong một năm cho trước được ước lượng là 0,01, thì công ty bảo hiểm phải tính mức phí bảo hiểm là bao nhiêu nếu công ty này muốn khoản lợi (gain) kỳ vọng bằng 1000USD?

3.34 Thời gian tồn tại tối đa của bằng sáng chế đối với một thuốc mới là 17 năm. Trừ đi khoảng thời gian cần thiết cho việc thử nghiệm và chấp thuận thuốc mới này bởi Cơ quan Quản lý Thực phẩm và Dược phẩm, anh/chị có được thời gian tồn tại thực tế của bằng sáng chế thuốc này—nghĩa là, khoảng thời gian mà một công ty phải thu hồi các chi phí nghiên cứu và phát triển cũng như kiếm lợi nhuận. Giả sử phân phối của khoảng thời gian tồn tại của bằng sáng chế đối với những thuốc mới được cho thấy trong bảng sau đây:

Năm, x	3	4	5	6	7	8	9	10	11	12	13
$p(x)$	0,03	0,05	0,07	0,10	0,14	0,20	0,18	0,12	0,07	0,03	0,01

a Tìm số năm kỳ vọng của thời gian tồn tại của bằng sáng chế đối với một thuốc mới.
b Tìm độ lệch chuẩn của *x*.
c Tìm xác suất để *x* nằm trong khoảng ($\mu \pm 2\sigma$).

3.35 Khi được hỏi liệu có nên giới hạn thời gian giữ nhiệm vụ của các nghị sĩ Quốc hội trong 12 năm hay không, thì 61% người Mỹ được hỏi đều tán thành một giới hạn nhiệm kỳ trong Quốc hội 12 năm (Galifianakis, 1994). Giả sử ba người được chọn ngẫu nhiên và được hỏi liệu họ có tán thành một giới hạn nhiệm kỳ trong Quốc hội 12 năm hay không. Gọi *x* là số người tán thành một giới hạn nhiệm kỳ trong quốc hội 12 năm.

a Tìm phân phối xác suất cho *x*.
b Xây dựng một biểu đồ xác suất hình thanh cho *x*.
c Xác suất để ít nhất là hai người tán thành một giới hạn nhiệm kỳ trong quốc hội 12 năm là bao nhiêu?
d Tìm trung bình và độ lệch chuẩn của *x*.

3.36 Một người đại diện của nhà sản xuất đang xem xét phương án nhận được một hợp đồng bảo hiểm, để bảo hiểm những tổn thất có thể có phát sinh bởi việc tiếp thị một sản phẩm mới. Nếu sản phẩm này là một thất bại hoàn toàn, thì người đại diện cảm nhận sẽ phải chịu một khoản lỗ là 80.000USD; nếu nó chỉ thành công vừa phải, thì sẽ phải chịu một khoản lỗ 25.000USD. Các chuyên viên tính toán bảo hiểm đã xác định, dựa trên những cuộc điều tra thị trường và thông tin có sẵn khác, rằng xác suất để sản phẩm này sẽ là một thất bại là 0,01 và xác suất để sản phẩm này chỉ thành công vừa phải là 0,05. Giả định rằng người đại diện của nhà sản xuất sẵn lòng bỏ qua tất cả những tổn thất có thể có khác, công ty bảo hiểm nên tính mức phí bảo hiểm đối với hợp đồng bảo hiểm này là bao nhiêu để được hòa vốn?

3.37 Một công ty chế tạo vận chuyển sản phẩm của mình bằng xe moóc theo xe tải với hai cỡ khác nhau, 8 x 10 x 30 và 8 x 10 x 40. Nếu 30% các chuyến hàng của nó được thực hiện bằng cách sử dụng xe moóc 30 bộ và 70% bằng cách sử dụng xe moóc 40 bộ, hãy tìm thể tích trung bình được vận chuyển trên mỗi xe moóc chở hàng (giả định các xe moóc luôn luôn đầy).