Capítulo 5

Probabilidades

5.1 Introdução

Na primeira parte deste livro, vimos que a análise de um conjunto de dados por meio de técnicas numéricas e gráficas permite que tenhamos uma boa idéia da distribuição desse conjunto. Em particular, a distribuição de freqüências é um instrumento importante para avaliarmos a variabilidade das observações de um fenômeno aleatório. A partir dessas freqüências observadas podemos calcular medidas de posição e variabilidade, como média, mediana, desvio padrão etc. Essas freqüências e medidas calculadas a partir dos dados são *estimativas* de quantidades desconhecidas, associadas em geral a populações das quais os dados foram extraídos na forma de *amostras*. Em particular, as freqüências (relativas) são estimativas de *probabilidades* de ocorrências de certos eventos de interesse. Com suposições adequadas, e sem observarmos diretamente o fenômeno aleatório de interesse, podemos criar um *modelo teórico* que reproduza de maneira razoável a distribuição das freqüências, quando o fenômeno é observado diretamente. Tais modelos são chamados *modelos probabilísticos* e serão objeto de estudo neste capítulo e nos subseqüentes.

Exemplo 5.1. Queremos estudar as freqüências de ocorrências das faces de um dado. Um procedimento a adotar seria lançar o dado certo número de vezes, n, e depois contar o número n_i de vezes em que ocorre a face i, $i = 1, 2, ..., 6$. As proporções n_i/n determinam a distribuição de freqüências do experimento realizado. Lançando o dado um número $n'(n' \neq n)$ de vezes, teríamos outra distribuição de freqüências, mas com um padrão que esperamos ser muito próximo do anterior.

O modelo probabilístico pode ser construído por meio de premissas, como se segue.

Primeiro, observamos que só podem ocorrer seis faces; a segunda consideração que se faz é que o dado seja perfeitamente equilibrado, de modo a não favorecer alguma face em particular. Com essas suposições, cada face deve ocorrer o mesmo número de vezes quando o dado é lançado n vezes, e, portanto, a proporção de ocorrência de cada face deve ser 1/6. Nessas condições, o modelo teórico (ou probabilístico) para o experimento é dado na Tabela 5.1.

5.3 Probabilidade Condicional e Independência

Voltemos à Tabela 5.3 do Exemplo 5.6. Dado que um estudante, escolhido ao acaso, esteja matriculado no curso de Estatística, a probabilidade de que seja mulher é 20/30 = 2/3. Isso porque, do total de 30 alunos que estudam Estatística, 20 são mulheres. Escrevemos

$$P(\text{mulher} \mid \text{Estatística}) = \frac{2}{3}.$$

Para dois eventos quaisquer A e B, sendo $P(B) > 0$, definimos a *probabilidade condicional de A dado B*, $P(A|B)$, como sendo

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{5.7}$$

Para o exemplo mencionado, se B e A indicam, respectivamente, os eventos "aluno matriculado em Estatística" e "aluno é mulher", então

$$P(A|B) = \frac{20/200}{30/200} = \frac{2}{3},$$

como havíamos obtido.

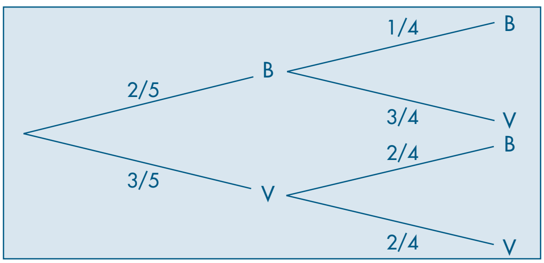
Observe que $P(A) = P(\text{mulher}) = 85/200 = 17/40$, e com a informação de que B ocorreu (o aluno é matriculado em Estatística), obtemos $P(A|B) = 2/3$. Podemos dizer que $P(A)$ é a probabilidade *a priori* de A e, com a informação adicional de que B ocorreu, obtemos a probabilidade *a posteriori* $P(A|B)$. Note que, nesse caso, $P(A|B) > P(A)$, logo a informação de que B ocorreu aumentou a chance de A ocorrer.

Da relação (5.7) obtemos a chamada *regra do produto de probabilidades*,

$$P(A \cap B) = P(B)\, P(A|B). \tag{5.8}$$

Exemplo 5.10. Uma urna contém duas bolas brancas (B) e três vermelhas (V). Suponha que são sorteadas duas bolas ao acaso, *sem reposição*. Isso significa que escolhemos a primeira bola, verificamos sua cor e não a devolvemos à urna; misturamos as bolas restantes e retiramos a segunda. O diagrama em árvore da Figura 5.2 ilustra as possibilidades. Em cada "galho" da árvore estão indicadas as probabilidades de ocorrência, sendo que para as segundas bolas as probabilidades são condicionais. A probabilidade do resultado conjunto é dada, então, por (5.8). Veja a Tabela 5.4.

Figura 5.2: Diagrama em árvore para a extração de duas bolas de uma urna, sem reposição.



Se A indicar o evento "bola branca na segunda extração", então

$$P(A) = P(BB) + P(VB) = \frac{2}{20} + \frac{6}{20} = \frac{2}{5}.$$

Tabela 5.4: Resultados e probabilidades para o experimento do Exemplo 5.10.

Resultados	Probabilidades
BB	$2/5 \times 1/4 = 2/20$
BV	$2/5 \times 3/4 = 6/20$
VB	$3/5 \times 2/4 = 6/20$
VV	$3/5 \times 2/4 = 6/20$
Total	1

Exemplo 5.11. Imagine, agora, que as duas extrações são feitas da mesma urna do exemplo anterior, mas a primeira bola é *reposta* na urna antes da extração da segunda. Nessas condições, as extrações são independentes, pois o resultado de uma extração não tem influência no resultado da outra. Obtemos a situação da Figura 5.3 e da Tabela 5.5.

Figura 5.3: Diagrama em árvore para a extração de duas bolas de uma urna, com reposição.

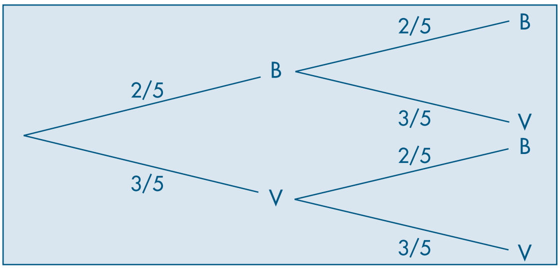


Tabela 5.5: Resultados e probabilidades para o experimento do Exemplo 5.11.

Resultados	Probabilidades
BB	$2/5 \times 2/5 = 4/25$
BV	$2/5 \times 3/5 = 6/25$
VB	$3/5 \times 2/5 = 6/25$
VV	$3/5 \times 3/5 = 9/25$
Total	1

Observe que, aqui,

$$P(\text{branca na } 2^{\text{a}} \mid \text{branca na } 1^{\text{a}}) = 2/5 = P(\text{branca na } 2^{\text{a}}),$$

ou seja, se indicarmos por A e B os eventos "bola branca na segunda extração" e "bola branca na primeira extração", respectivamente, então $P(A|B) = P(A)$. Nesse caso, dizemos que o evento A *independe* do evento B e, usando (5.8), temos

$$P(A \cap B) = P(A)\ P(B). \tag{5.9}$$

É fácil ver que se A independe de B, então B independe de A — dizemos que A e B são independentes. A fórmula (5.9) pode ser tomada como definição de independência entre dois eventos, ou seja, A e B são *independentes* se, e somente se, (5.9) for válida.

Exemplo 5.12. Considere ainda a urna dos dois exemplos anteriores, mas vamos fazer três extrações *sem reposição*. Indiquemos por V_i ou B_i a obtenção de bola vermelha ou branca na i-ésima extração, respectivamente, i = 1, 2, 3. Obtemos a Figura 5.4 e a Tabela 5.6.

Figura 5.4: Diagrama em árvore para a extração de três bolas de uma urna, sem reposição.

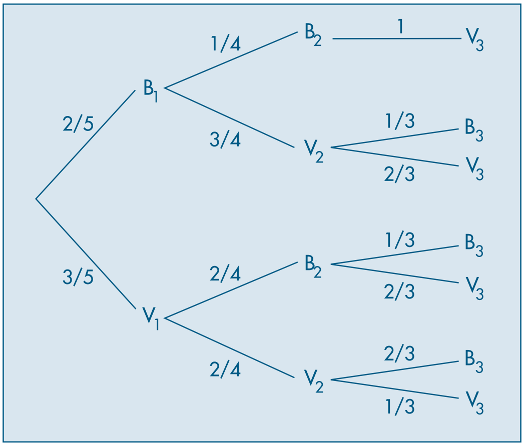


Tabela 5.6: Resultados e probabilidades para o experimento do Exemplo 5.12.

Resultados	Probabilidades
$B_1B_2V_3$	$2/5 \times 1/4 \times 1 = 2/20 = 6/60$
$B_1V_2B_3$	$2/5 \times 3/4 \times 1/3 = 6/60$
$B_1V_2V_3$	$2/5 \times 3/4 \times 2/3 = 12/60$
$V_1B_2B_3$	$3/5 \times 2/4 \times 1/3 = 6/60$
$V_1B_2V_3$	$3/5 \times 2/4 \times 2/3 = 12/60$
$V_1V_2B_3$	$3/5 \times 2/4 \times 2/3 = 12/60$
$V_1V_2V_3$	$3/5 \times 2/4 \times 1/3 = 6/60$
Total	$60/60 = 1$

Observe que $P(B_2|B_1) = 1/4$, ao passo que $P(V_3|B_1 \cap B_2) = 1$; daí,

$$P(B_1 \cap B_2 \cap V_3) = P(B_1)\ P(B_2|B_1)\ P(V_3|B_1 \cap B_2) = 2/5 \times 1/4 \times 1 = 1/10.$$

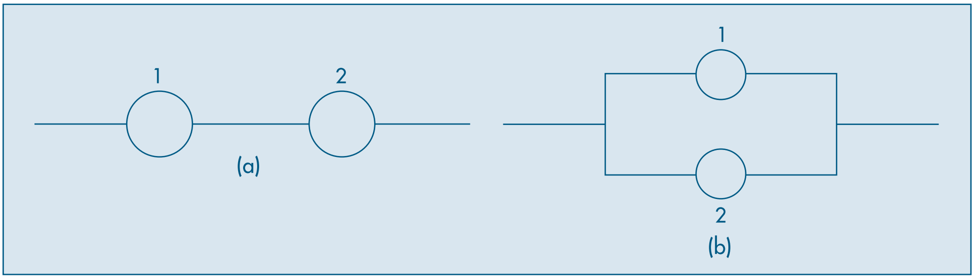
De modo geral, dados três eventos A, B e C, temos que

$$P(A \cap B \cap C) = P(A)\ P(B|A)\ P(C|A \cap B). \tag{5.10}$$

Essa relação pode ser estendida para um número finito qualquer de eventos. Veja o Problema 60.

Exemplo 5.13. A *teoria da confiabilidade* estuda sistemas e seus componentes, como por exemplo sistemas mecânicos e eletrônicos (um automóvel ou um computador) e sistemas biológicos, como o corpo humano. O objetivo da teoria é estudar as relações entre o funcionamento dos componentes e do sistema. A Figura 5.5 (a) ilustra um sistema composto de dois componentes ligados em *série*.

Figura 5.5: Sistema com dois componentes (a) em série (b) em paralelo.



O sistema da figura funcionará se os componentes 1 e 2 funcionarem simultaneamente. Se um dos componentes falhar, o sistema também falhará. Supondo que os componentes funcionem *independentemente*, e se p_i for a probabilidade de o componente i (i = 1,2) funcionar, então a probabilidade de o sistema funcionar será

$$P(F) = P(A_1 \cap A_2) = P(A_1)P(A_2) = p_1 p_2,$$

onde indicamos por F o evento "o sistema funciona" e por A_i o evento "o componente i funciona", i = 1, 2.

A probabilidade p_i é a chamada *confiabilidade do componente i* e $P(F) = h(p_1, p_2) = p_1 p_2$ a *confiabilidade do sistema*.

Se os componentes 1 e 2 estiverem em *paralelo*, como na Figura 5.5 (b), então o sistema funcionará se *pelo menos um* dos dois componentes funcionar. Ou seja,

$$P(F) = P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = p_1 + p_2 - p_1 p_2$$

e a confiabilidade do sistema é $h(p_1, p_2) = p_1 + p_2 - p_1 p_2$.

Vejamos agora o conceito de independência para três eventos: dizemos que os eventos A, B e C são *independentes* se, e somente se,

$$\begin{aligned} &P(A \cap B) = P(A)\ P(B), \\ &P(A \cap C) = P(A)\ P(C), \\ &P(B \cap C) = P(B)\ P(C), \\ &P(A \cap B \cap C) = P(A)\ P(B)\ P(C). \end{aligned} \tag{5.11}$$

Se apenas as três primeiras relações de (5.11) estiverem satisfeitas, dizemos que os eventos A, B e C são *mutuamente independentes*. É possível que três eventos sejam mutuamente independentes, mas não sejam completamente independentes. Veja o Problema 59.

A definição pode ser estendida facilmente para um número finito qualquer de eventos. Veja o Problema 61.

Problemas

15. Considere uma urna contendo três bolas pretas e cinco bolas vermelhas. Retire duas bolas da urna, sem reposição.
    (a) Obtenha os resultados possíveis e as respectivas probabilidades.
    (b) Mesmo problema, para extrações com reposição.

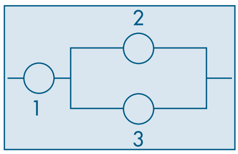
16. No problema anterior, calcule as probabilidades dos eventos:
    (a) Bola preta na primeira e segunda extrações.
    (b) Bola preta na segunda extração.
    (c) Bola vermelha na primeira extração.

17. A probabilidade de que A resolva um problema é de 2/3, e a probabilidade de que B o resolva é de 3/4. Se ambos tentarem independentemente, qual a probabilidade de o problema ser resolvido?

18. Um dado é viciado, de tal forma que a probabilidade de sair um certo ponto é proporcional ao seu valor (por exemplo, o ponto 6 é três vezes mais provável de sair do que o ponto 2). Calcular:
    (a) a probabilidade de sair 5, sabendo-se que o ponto que saiu é ímpar;
    (b) a probabilidade de tirar um número par, sabendo-se que saiu um número maior que 3.

19. As probabilidades de que dois eventos independentes ocorram são p e q, respectivamente. Qual a probabilidade:
    (a) de que nenhum desses eventos ocorra?
    (b) de que pelo menos um desses eventos ocorra?

20. Na figura ao lado temos um sistema com três componentes funcionando independentemente, com confiabilidades p_1, p_2 e p_3. Obtenha a confiabilidade do sistema.

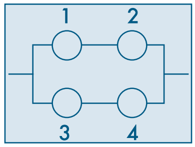


21. Na tabela abaixo, os números que aparecem são probabilidades relacionadas com a ocorrência de A, B, $A \cap B$ etc. Assim, $P(A) = 0{,}10$, enquanto $P(A \cap B) = 0{,}04$.

	B	B^c	Total
A	0,04	0,06	0,10
A^c	0,08	0,82	0,90
Total	0,12	0,88	1,00

Verifique se A e B são independentes.

22. Supondo que todos os componentes do sistema da figura ao lado tenham a mesma confiabilidade p e funcionem independentemente, obtenha a confiabilidade do sistema.



5.4 O Teorema de Bayes

Uma das relações mais importantes envolvendo probabilidades condicionais é dada pelo Teorema de Bayes. A versão mais simples desse teorema é dada pela fórmula (5.12):

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(B)}. \tag{5.12}$$

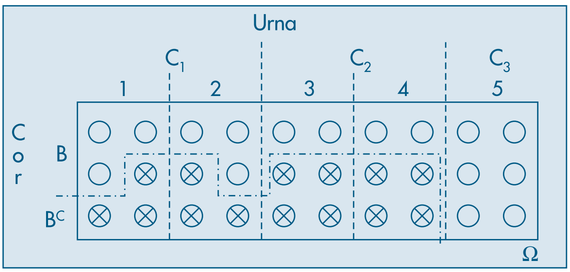
Como salientamos na seção anterior, temos a probabilidade inicial $P(A)$ e, dada a informação de que B ocorreu (ou dada a suposição de que B venha a ocorrer), obtemos a probabilidade *a posteriori* $P(A|B)$, dada por (5.12). Ou seja, *atualizamos* a probabilidade inicial, multiplicando-a por $\frac{P(B|A)}{P(B)}$. Observe que $P(A|B) > P(A)$ se $P(B|A) > P(B)$.

A forma geral do Teorema de Bayes será introduzida por um exemplo.

Exemplo 5.14. Temos cinco urnas, cada uma com seis bolas. Duas dessas urnas (tipo C_1) têm 3 bolas brancas, duas outras (tipo C_2) têm 2 bolas brancas, e a última urna (tipo C_3) tem 6 bolas brancas. Escolhemos uma urna ao acaso e dela retiramos uma bola. Qual a probabilidade de a urna escolhida ser do tipo C_3, sabendo-se que a bola sorteada é branca?

Na Figura 5.6 temos esquematizados o espaço amostral e os eventos de interesse.

Figura 5.6: Espaço amostral e eventos para o Exemplo 5.14.



Queremos encontrar $P(C_3|B)$, sabendo que

$$P(C_1) = 2/5,\ P(B|C_1) = 1/2,$$
$$P(C_2) = 2/5,\ P(B|C_2) = 1/3,$$
$$P(C_3) = 1/5,\ P(B|C_3) = 1.$$

Da definição de probabilidade condicional, temos

$$P(C_3|B) = \frac{P(C_3 \cap B)}{P(B)} = \frac{P(C_3)P(B|C_3)}{P(B)} . \tag{5.13}$$

A segunda igualdade é devida à fórmula (5.8).

Precisamos encontrar o valor de $P(B)$, já que o numerador é conhecido. Como C_1, C_2 e C_3 são eventos mutuamente exclusivos, e reunidos formam o espaço amostral completo, podemos decompor o evento B na reunião de três outros, também mutuamente exclusivos, como segue (ver também a Figura 5.6):

$$B = (C_1 \cap B) \cup (C_2 \cap B) \cup (C_3 \cap B), \tag{5.14}$$

e então

$$\begin{aligned} P(B) &= P(C_1 \cap B) + P(C_2 \cap B) + P(C_3 \cap B) \\ &= P(C_1)\ P(B|C_1) + P(C_2)\ P(B|C_2) + P(C_3)\ P(B|C_3) \\ &= \frac{2}{5} \times \frac{1}{2} + \frac{2}{5} \times \frac{1}{3} + \frac{1}{5} \times 1 = \frac{8}{15}. \end{aligned}$$

Substituindo esse resultado em (5.13), obtemos

$$P(C_3|B) = \frac{1/5 \times 1}{8/15} = \frac{3}{8} .$$

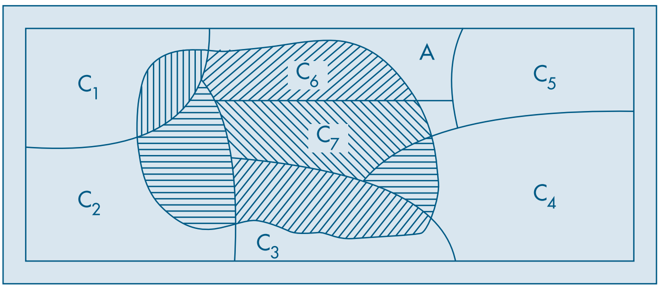
Podemos, agora, generalizar os resultados acima do seguinte modo: seja $\{C_1, C_2, ..., C_n\}$ uma partição do espaço amostral Ω, isto é,

$$C_i \cap C_j = \emptyset, \quad \text{sempre que} \quad i \neq j,$$
$$C_1 \cup C_2 \cup ... \cup C_n = \Omega.$$

Considere um evento qualquer A em Ω. Supomos conhecidas as probabilidades $P(C_i)$ e $P(A|C_i)$, $i = 1, 2, ..., n$.

Então, temos o seguinte resultado, ilustrado pela Figura 5.7.

Figura 5.7: Partição de um espaço amostral.



Teorema 5.1 (Bayes). A probabilidade de ocorrência do evento C_i, supondo-se a ocorrência do evento A, é dada por

$$P(C_i|A) = \frac{P(C_i)P(A|C_i)}{\sum_{j=1}^{n} P(C_j)P(A|C_j)}, \tag{5.15}$$

para todo $i = 1, 2, ..., n$.

Podemos pensar $C_1, ..., C_n$ como um conjunto de *hipóteses*, sendo somente uma delas verdadeira. Dado que A ocorreu, a probabilidade inicial de C_i, $P(C_i)$, é modificada de modo a se obter $P(C_i|A)$, dada por (5.15). Passamos da probabilidade *a priori* $P(C_i)$ para a probabilidade *a posteriori* $P(C_i|A)$, multiplicando a primeira por

$$\frac{P(A|C_i)}{\sum_{j=1}^{n} P(C_j)P(A|C_j)}. \tag{5.16}$$

Para A fixado, as probabilidades $P(A|C_i)$ em (5.15) são denominadas *verossimilhanças* das hipóteses $C_1, C_2, ..., C_n$. Vemos que $P(C_i|A) > P(C_i)$ se (5.16) for maior do que um, isto é, se $P(A|C_i) > P(A)$, onde $P(A)$ é o denominador de (5.16). Observe que esse denominador é uma média ponderada dos $P(A|C_j)$ e os pesos são as probabilidades $P(C_j)$, que têm soma unitária. Como o numerador é sempre uma das parcelas do denominador $P(A)$, torna-se indispensável o uso de um novo índice, j, na decomposição deste.

Exemplo 5.15. Para selecionar seus funcionários, uma empresa oferece aos candidatos um curso de treinamento durante uma semana. No final do curso, eles são submetidos a uma prova e 25% são classificados como bons (B), 50% como médios (M) e os restantes 25% como fracos (F). Para facilitar a seleção, a empresa pretende substituir o treinamento por um teste contendo questões referentes a conhecimentos gerais e específicos. Para isso, gostaria de conhecer qual a probabilidade de um indivíduo aprovado no teste ser considerado fraco, caso fizesse o curso. Assim, neste ano, antes do início do curso, os candidatos foram submetidos ao teste e receberam o conceito aprovado (A) ou reprovado (R). No final do curso, obtiveram-se as seguintes probabilidades condicionais:

$$P(A|B) = 0{,}80, \qquad P(A|M) = 0{,}50, \qquad P(A|F) = 0{,}20.$$

Queremos encontrar $P(F|A)$ e, pelo Teorema de Bayes, essa probabilidade é dada por

$$P(F|A) = \frac{P(A|F)P(F)}{P(A|B)P(B) + P(A|M)P(M) + P(A|F)P(F)}$$

$$= \frac{(0{,}20)(0{,}25)}{(0{,}80)(0{,}25) + (0{,}50)(0{,}50) + (0{,}20)(0{,}25)} = 0{,}10.$$

Então, apenas 10% dos aprovados é que seriam classificados como fracos durante o curso. De modo análogo podemos encontrar $P(B|A) = 0{,}40$ e $P(M|A) = 0{,}50$, que poderiam fornecer subsídios para ajudar na decisão de substituir o treinamento pelo teste.

Um gráfico em árvore pode ajudar bastante na solução de um problema envolvendo o Teorema de Bayes. Desse modo, para o Exemplo 5.15, teremos a Figura 5.8 e a Tabela 5.7. Assim, o numerador de $P(F|A)$ está assinalado com um pequeno círculo, ao passo que o denominador é a soma das três parcelas assinaladas com asterisco.

Figura 5.8: Diagrama em árvore para o Exemplo 5.15.

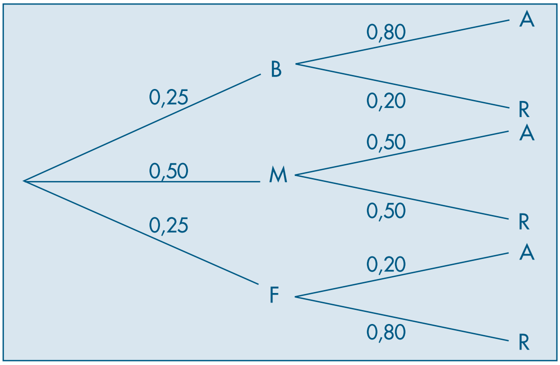


Tabela 5.7: Resultados e probabilidades para o Exemplo 5.15.

Resultados	Probabilidades
BA	(0,25) (0,80) = 0,20 *
BR	(0,25) (0,20) = 0,05
MA	(0,50) (0,50) = 0,25 *
MR	(0,50) (0,50) = 0,25
FA	(0,25) (0,20) = 0,05 * °
FR	(0,25) (0,80) = 0,20

O Teorema de Bayes, que aparentemente poderia ser encarado como mais um resultado na teoria de probabilidades, tem importância fundamental, pois fornece a base para uma abordagem da inferência estatística conhecida como *inferência bayesiana*. Esse ponto será abordado brevemente no Capítulo 11.

O Teorema de Bayes fornece um mecanismo formal para atualizar probabilidades, como já vimos acima. Vejamos mais um exemplo para ilustrar esse ponto.

Exemplo 5.16. A administração de um fundo de investimentos em ações pretende divulgar, após o encerramento do pregão, a probabilidade de queda de um índice da bolsa no dia seguinte, baseando-se nas informações disponíveis até aquele momento. Suponha que a previsão inicial seja de 0,10. Após encerrado o pregão, nova informação sugere uma alta do dólar frente ao real. A experiência passada indica que, quando houve queda da bolsa no dia seguinte, 20% das vezes foram precedidas por esse tipo de notícia, enquanto, nos dias em que a bolsa esteve em alta, apenas em 5% das vezes houve esse tipo de notícia no dia anterior.

Chamando de E o evento que indica "queda da bolsa", a sua probabilidade *a priori* é $P(E) = 0{,}10$, enquanto a probabilidade de alta é $P(E^c) = 0{,}90$. Se B indicar "alta do dólar", então as verossimilhanças são dadas por

$$P(B|E) = 0{,}20, \qquad P(B|E^c) = 0{,}05.$$

Logo, pelo Teorema de Bayes, teremos que

$$P(E|B) = \frac{P(E)\ P(B|E)}{P(E)P(B|E) + P(E^c)P(B|E^c)},$$

ou seja,

$$P(E|B) = \frac{(0{,}10)(0{,}20)}{(0{,}10)(0{,}20) + (0{,}90)(0{,}05)} = \frac{0{,}02}{0{,}065} = \frac{4}{13} = 0{,}31.$$

Portanto, a nova informação aumenta a probabilidade de que haja queda na bolsa de 10% para 31%.

Suponha, agora, que horas depois surja nova informação relevante: o Banco Central irá reduzir a taxa de juros vigente a partir do dia seguinte. Denotando-se, agora, por B_1 o evento "alta do dólar" e por B_2 o evento "queda na taxa de juros", o interesse será saber como essa nova informação, B_2, afetará a probabilidade calculada, $P(E|B_1)$. Segue-se que essa é agora a probabilidade *a priori* para E com respeito a B_2.

Novamente, informações passadas mostram que, dado que tenha havido alta do dólar e queda da bolsa, 10% das vezes foram precedidas por notícias de queda de juros, enquanto, dado que tenha havido alta do dólar e alta da bolsa, 60% das vezes foram precedidas de queda dos juros. Então, as verossimilhanças agora serão dadas por

$$P(B_2|E,\ B_1) = 0{,}10, \qquad P(B_2|E^c,\ B_1) = 0{,}60.$$

O Teorema de Bayes fica escrito agora na forma

$$P(E|B_1,\ B_2) = \frac{P(E|B_1)\ P(B_2|E,\ B_1)}{P(E|B_1)\ P(B_2|E,\ B_1) + P(E^c|B_1)\ P(B_2|E^c,\ B_1)},$$

do que segue que

$$P(E|B_1,\ B_2) = \frac{(0{,}31)(0{,}10)}{(0{,}31)(0{,}10) + (0{,}69)(0{,}60)} = \frac{0{,}031}{0{,}445} = 0{,}07.$$

Ou seja, a informação B_2 causa um decréscimo na probabilidade de queda da bolsa, de 0,31 para 0,07, que é menor ainda do que a probabilidade *a priori* inicial, P(E) = 0,10.

Observe que a probabilidade $P(E|B_1,\ B_2)$ pode ser escrita também como $P(E|B_1 \cap B_2)$, ou seja, temos a ocorrência simultânea dos eventos B_1 e B_2.

Problemas

23. Uma companhia produz circuitos em três fábricas, I, II e III. A fábrica I produz 40% dos circuitos, enquanto a II e a III produzem 30% cada uma. As probabilidades de que um circuito integrado produzido por essas fábricas não funcione são 0,01, 0,04 e 0,03, respectivamente. Escolhido um circuito da produção conjunta das três fábricas, qual a probabilidade de o mesmo não funcionar?
24. Considere a situação do problema anterior, mas suponha agora que um circuito escolhido ao acaso seja defeituoso. Determine qual a probabilidade de ele ter sido fabricado por I.
25. A urna I contém duas bolas pretas e três brancas, ao passo que a urna II contém três bolas pretas e três brancas. Escolhemos uma urna ao acaso e dela extraímos uma bola que tem cor branca. Se a bola é recolocada na urna, qual é a probabilidade de se retirar novamente uma bola branca da mesma urna?

5.5 Probabilidades Subjetivas

Na seção 5.1 vimos como associar probabilidades a eventos. Utilizamos um enfoque chamado *freqüentista*, pois se baseia na estabilidade das freqüências relativas e no fato de podermos, hipoteticamente, repetir um experimento várias vezes. Mas é óbvio que nem sempre podemos considerar replicações. Suponha que queiramos calcular a probabilidade de chover no dia 12 de janeiro do próximo ano, na cidade de São Paulo. Evidentemente, se considerarmos o evento A = chover em São Paulo no dia 12 de janeiro do próximo ano, ele não pode ser replicado. O que poderemos eventualmente considerar é em quantos dias 12 de janeiro de anos anteriores choveu e calcular uma freqüência relativa. Se tivermos essa informação, ela evidentemente poderá ser usada. Mas suponha que uma pessoa morando em Fortaleza tenha de calcular essa probabilidade. Se ela não tiver informação sobre o tempo em São Paulo, poderá simplesmente dizer que essa probabilidade é de 1/2. Por outro lado, uma pessoa vivendo em São Paulo terá informações adicionais. Por exemplo, saberá que normalmente janeiro, fevereiro e março são meses com muita chuva. Esse morador de São Paulo poderá arriscar uma probabilidade, digamos de 2/3 para o evento A. Vemos, portanto, que a associação de probabilidades a um evento depende de cada indivíduo, de sua informação a respeito desse evento. Esse tipo de apreciação é particularmente recomendável quando o indivíduo julga que as replicações anteriores não sejam comparáveis com a próxima. Por exemplo, o fenômeno El Niño pode ter ocorrido com grande intensidade em janeiro de 1999, provocando muita chuva no sudeste do Brasil, e sua intensidade nos anos seguintes talvez seja menor.

Respostas a questões como essa envolvem o que chamamos de *probabilidade subjetiva*. Ou seja, cada indivíduo, baseado em informações anteriores e na sua opinião pessoal a respeito do evento em questão, pode ter uma resposta para a probabilidade desse evento. A Inferência Bayesiana, de que trataremos brevemente neste livro (veja o Capítulo 11), toma como uma de suas bases o fato de que todas as probabilidades são subjetivas. O Teorema de Bayes tem papel importante nesse tipo de inferência, pois passa a ser visto como um mecanismo de atualização de opiniões. Ou seja, o indivíduo aprende B e passa a ter opinião $P(A|B)$ sobre A.

Um ingrediente básico quando se associam probabilidades é a *coerência*. Se um indivíduo julgar que um evento A é mais provável que seu complementar, então ele deverá, como que apostando na ocorrência de A, associar uma probabilidade maior do que 1/2 ao evento A. Por exemplo, se ele julgar que uma proporção 3 : 1 a favor de A é razoável, então ele deverá sugerir $P(A) = 3/4$. A fórmula de Bayes fornece uma maneira coerente de atualizar opiniões.

As probabilidades associadas a eventos de modo subjetivo têm propriedades análogas àquelas vistas em seções anteriores e podem ser obtidas a partir do princípio da coerência. Há outras maneiras de se associar probabilidades a eventos e os interessados poderão consultar O'Hagan (1994), por exemplo, para obter mais informações sobre esse assunto e outros ligados à Inferência Bayesiana.

5.6 Problemas e Complementos

26. Um restaurante popular apresenta apenas dois tipos de refeições: salada completa ou um prato à base de carne. Considere que 20% dos fregueses do sexo masculino preferem a salada, 30% das mulheres escolhem carne, 75% dos fregueses são homens e os seguintes eventos:

    H: freguês é homem | A: freguês prefere salada

    M: freguês é mulher | B: freguês prefere carne

    Calcular:

    (a) $P(H), P(A|H), P(B|M)$; (b) $P(A \cap H), P(A \cup H)$; (c) $P(M|A)$.

27. Uma companhia de seguros analisou a freqüência com que 2.000 segurados (1.000 homens e 1.000 mulheres) usaram o hospital. Os resultados são apresentados na tabela:

	Homens	Mulheres
Usaram o hospital	100	150
Não usaram o hospital	900	850

    (a) Qual a probabilidade de que uma pessoa segurada use o hospital?
    (b) O uso do hospital independe do sexo do segurado?

28. As probabilidades de três motoristas serem capazes de guiar até em casa com segurança, depois de beber, são de 1/3, 1/4 e 1/5, respectivamente. Se decidirem guiar até em casa, depois de beber numa festa, qual a probabilidade de todos os três motoristas sofrerem acidentes? Qual a probabilidade de pelo menos um dos motoristas guiar até em casa a salvo?

29. Duas lâmpadas queimadas foram acidentalmente misturadas com seis lâmpadas boas. Se vamos testando as lâmpadas, uma por uma, até encontrar duas defeituosas, qual é a probabilidade de que a última defeituosa seja encontrada no quarto teste?

30. Suponhamos que 10.000 bilhetes sejam vendidos em uma loteria e 5.000 em outra, cada uma tendo apenas um ganhador. Um homem tem 100 bilhetes de cada. Qual a probabilidade de que:
    (a) ele ganhe exatamente um prêmio?
    (b) ele ganhe alguma coisa?

31. Uma companhia de seguros vendeu apólices a cinco pessoas, todas da mesma idade e com boa saúde. De acordo com as tábuas atuariais, a probabilidade de que uma pessoa daquela idade esteja viva daqui a 30 anos é de 2/3. Calcular a probabilidade de que daqui a 30 anos:
   (a) exatamente duas pessoas estejam vivas;
   (b) todas as pessoas estejam vivas; e
   (c) pelo menos três pessoas estejam vivas.
   (Indique as suposições necessárias para a resolução do problema.)

32. Num teste com duas marcas que lhe são apresentadas em ordem aleatória, um experimentador de vinhos faz três identificações corretas em três tentativas.
   (a) Qual a probabilidade de isso ocorrer, se na realidade ele não possuir habilidade alguma para distingui-los?
   (b) E se a probabilidade de distinguir corretamente é de 90% em cada tentativa?

33. Um grupo de 12 homens e 8 mulheres concorre a três prêmios através de um sorteio, sem reposição de seus nomes. Qual a probabilidade de:
   (a) nenhum homem ser sorteado?
   (b) um prêmio ser ganho por homem?
   (c) dois homens serem premiados?

34. Um empreiteiro apresentou orçamentos separados para a execução da parte elétrica e da parte de encanamento de um edifício. Ele acha que a probabilidade de ganhar a concorrência da parte elétrica é de 1/2. Caso ele ganhe a parte elétrica, a chance de ganhar a parte de encanamento é de 3/4; caso contrário, essa probabilidade é de 1/3. Qual a probabilidade de ele:
   (a) ganhar os dois contratos?
   (b) ganhar apenas um?
   (c) não ganhar nada?

35. Em média, 5% dos produtos vendidos por uma loja são devolvidos. Qual a probabilidade de que, das quatro próximas unidades vendidas desse produto, duas sejam devolvidas?

36. Três alarmes estão dispostos de tal maneira que qualquer um deles funcionará independentemente quando qualquer coisa indesejável ocorrer. Se cada alarme tem probabilidade 0,9 de trabalhar eficientemente, qual é a probabilidade de se ouvir o alarme quando necessário?

37. Em uma fábrica de parafusos, as máquinas A, B e C produzem 25%, 35% e 40% do total, respectivamente. Da produção de cada máquina 5%, 4% e 2%, respectivamente, são parafusos defeituosos. Escolhe-se ao acaso um parafuso e verifica-se que é defeituoso. Qual a probabilidade de que o parafuso venha da máquina A; da B; e da C?

38. Um fabricante afirma que apenas 5% de todas as válvulas que produz têm duração inferior a 20 horas. Uma indústria compra semanalmente um grande lote de válvulas desse fabricante, mas sob a seguinte condição: ela aceita o lote se, em dez válvulas escolhidas ao acaso, no máximo uma tiver duração inferior a 20 horas; caso contrário, o lote todo é rejeitado.

(a) Se o fabricante de fato tem razão, qual a probabilidade de um lote ser rejeitado?

(b) Suponha agora que o fabricante esteja mentindo, isto é, na verdade a proporção de válvulas com duração inferior a 20 horas é de 10%. Qual a probabilidade de um lote ser aceito, segundo o critério acima?

39. Para estudar o comportamento do mercado automobilístico, as marcas foram divididas em três categorias: marca F, marca W, e as demais reunidas como marca X. Um estudo sobre o hábito de mudança de marca mostrou o seguinte quadro de probabilidade:

Proprietário de carro da marca	Probabilidade de mudança para		
	W	F	X
W	0,50	0,25	0,25
F	0,15	0,70	0,15
X	0,30	0,30	0,40

A compra do primeiro carro é feita segundo as seguintes probabilidades: marca W com 50%, marca F com 30% e marca X com 20%.

(a) Qual a probabilidade de um indivíduo comprar o terceiro carro da marca W?

(b) Se o terceiro carro é da marca W, qual a probabilidade de o primeiro também ter sido W?

40. A empresa $M\,\&\,B$ tem 15.800 empregados, classificados de acordo com a tabela abaixo.

Idade \ Sexo	Homens (M)	Mulheres (F)	Total
< 25 anos (A)	2.000	800	2.800
25 – 40 anos (B)	4.500	2.500	7.000
> 40 anos (C)	1.800	4.200	6.000
Total	8.300	7.500	15.800

Se um empregado é selecionado ao acaso, calcular a probabilidade de ser ele:

(a) um empregado com 40 anos de idade ou menos;

(b) um empregado com 40 anos de idade ou menos, e mulher;

(c) um empregado com mais de 40 anos de idade e que seja homem;

(d) uma mulher, dado que é um empregado com menos de 25 anos.

41. Considere o Problema 40 e suponha que escolhamos dois empregados ao acaso, *com reposição*. Qual a probabilidade de que:

(a) ambos sejam do sexo masculino;

(b) o primeiro tenha menos de 25 anos, e o segundo seja do sexo masculino e tenha menos de 25 anos;

(c) nenhum tenha menos de 25 anos.

42. Resolva as questões *(a)* e *(c)* do Problema 41, supondo que a amostragem é feita *sem* reposição.

43. Numa empresa existem operários de determinada categoria, com idades iguais a a, b e c anos (existem pelo menos três com a mesma idade). Escolhem-se três ao acaso para que façam determinado curso. Se indicarmos por x a idade do primeiro, y a do segundo e z a do terceiro, o terno (x, y, z) indica cada possível resultado. Enumere:
(a) o espaço amostral; e *(b)* os eventos $A = \{(x, y, z) | x = y = z\}$, $B = \{(x, y, z) | x = y\}$.

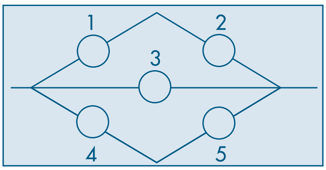
44. Os colégios A, B e C têm as seguintes porcentagens de rapazes, respectivamente: 40%, 20% e 10%. Um desses colégios é selecionado ao acaso e oito alunos são escolhidos, *com* reposição. Se o resultado for $RRRMMMMM$ (R para rapaz e M para moça), qual é a probabilidade de ter sido selecionado o colégio C?

45. Um inspetor da seção de controle de qualidade de uma firma examina os artigos de um lote que tem m peças de primeira qualidade e n peças de segunda qualidade. Uma verificação dos b primeiros artigos selecionados ao acaso do lote mostrou que todos eram de segunda qualidade $(b < n - 1)$. Qual a probabilidade de que entre os dois próximos artigos selecionados, ao acaso, dos restantes, pelo menos um seja de segunda qualidade?

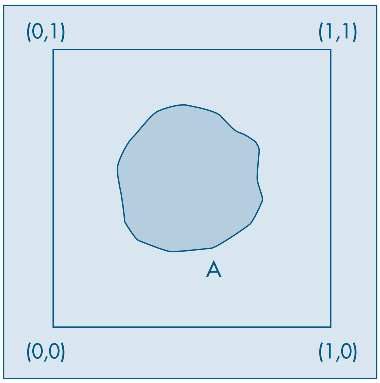
46. Prove que, se A e B são independentes, também o serão A^c e B^c, A e B^c e A^c e B.

47. Obtenha uma fórmula para $P(A \cup B \cup C)$.

48. Na figura abaixo temos um sistema chamado *ponte*. Nas mesmas condições do Problema 22, obtenha a confiabilidade do sistema.



49. Considere o quadrado com vértices (0,0), (1,0), (0,1) e (1,1). Suponha que a probabilidade de uma região A (evento) seja a área dessa região.



(a) Represente graficamente o evento A = conjunto dos pontos cuja distância à origem seja menor ou igual a 1.
(b) Calcule $P(A)$.

(c) Calcule a probabilidade do evento $B = \{(x, y) : x \geqslant b \text{ ou } y \geqslant b\}$, onde b é um número tal que $0 < b < 1$.

(d) Calcule $P(B^c)$, onde B foi definido em (c).

50. Considere Ω como o quadrado da figura do Problema 49. Considere os eventos:
$A = \{(x, y) : 1/3 \leqslant x \leqslant 2/3, 0 \leqslant y \leqslant 1/2\}$
$B = \{(x, y) : 1/2 \leqslant x \leqslant 1, 1/4 \leqslant y \leqslant 3/4\}$.
Calcular $P(A)$, $P(B)$, $P(A \cup B)$, $P(A^c)$, $P(B^c)$ e $P(A^c \cap B^c)$.

51. Considere, agora, a situação do Problema 49, mas suponha que o quadrado não tenha área unitária. Como você definiria a probabilidade de um evento A?

52. Suponha uma *população* de N elementos $a_1, a_2, ..., a_N$. Qualquer arranjo ordenado $a_{i1}, a_{i2}, ..., a_{in}$ de n símbolos é chamado de uma *amostra ordenada de tamanho* n, extraída da população. Considere o símbolo $(N)_n$ como significando $N(N-1) \ldots (N-n+1)$. Suponha $n < N$. Mostre que existem N^n amostras com reposição (um mesmo elemento pode ser retirado mais de uma vez) e $(N)_n$ amostras sem reposição (um elemento, quando escolhido, é removido da população, não havendo, pois, repetição na amostra).

53. Uma amostra ordenada de tamanho n, extraída de uma população com N elementos, produz um plano aleatório simples se todas as possíveis amostras têm a mesma probabilidade de serem escolhidas; essa probabilidade será $1/N^n$ se a amostra for com reposição e $1/(N)_n$ se for sem reposição. Uma amostra casual de tamanho n, com reposição, é extraída de uma população com N elementos. Encontre a probabilidade de não haver repetição na amostra.

54. Considere $\binom{N}{n} = \dfrac{(N)_n}{n!} = \dfrac{N!}{n!(N-n)!}$. Observe a situação do Problema 52, na qual *não levamos em consideração a ordem* do conjunto $a_{i1}, a_{i2}, ..., a_{in}$. Mostre que existem $\binom{N}{n}$ amostras *sem reposição*.

55. (a) Se A, B e C são independentes, prove que A e $B \cap C$ são independentes.
(b) Nas mesmas condições, prove que $A \cup B$ e C são independentes.

56. Dizemos que $A \subset B$ (A é subconjunto de B) se todo elemento de A também pertence a B. Por exemplo, $\{1, 2\} \subset \{1, 2, 3\}$. Se $P(A) = 1/3$, $P(B^c) = 1/4$, A e B podem ser disjuntos (ou mutuamente exclusivos)? (Sugestão: $P(A) = P(A \cap B) + P(A \cap B^c)$ e $A \cap B^c \subset B^c$. Use o fato de que, se $A \subset B$, $P(A) \leqslant P(B)$.)

57. Um sistema é composto de três componentes 1, 2 e 3, com confiabilidade 0,9, 0,8 e 0,7, respectivamente. O componente 1 é indispensável ao funcionamento do sistema; se 2 ou 3 não funcionam, o sistema funciona, mas com um rendimento inferior. A falha simultânea de 2 e 3 implica o não-funcionamento do sistema. Supondo que os componentes funcionem independentemente, calcular a confiabilidade do sistema.

58. Prove (5.4). (Sugestão: Escreva $U \cup V$ e V como reuniões de eventos mutuamente exclusivos.)

59. Há quatro bolas numa urna, numeradas 000, 011, 101, 110. Selecione uma bola ao acaso da urna. Considere os eventos
A_i: na bola selecionada, o número 1 aparece na posição i, $i = 1, 2, 3$.
Seja $A = A_1 \cap A_2 \cap A_3$.
(a) Calcule $P(A_i)$, $i = 1, 2, 3$ e $P(A)$.
(b) Mostre que A_1, A_2 e A_3 são mutuamente independentes, mas não são independentes.

60. Como fica a relação (5.10) para n eventos quaisquer $A_1, A_2, ..., A_n$?

61. Definir independência para n eventos quaisquer $A_1, ..., A_n$.

62. O problema do aniversário. Considere k pessoas numa sala. Qual a probabilidade de que pelo menos duas pessoas façam aniversário no mesmo dia e mês? A partir de qual valor de k essa probabilidade é maior que 0,5?
(Sugestão: seja A o evento "pelo menos duas pessoas fazem aniversário no mesmo dia". O evento complementar é A^c: "todas as k pessoas fazem aniversário em dias diferentes". Calcule primeiro a $P(A^c)$. Para isso, use o resultado do Problema 53. Aqui, temos $N = 365$ dias e $k = n$ pessoas. Se $P(A) = p$, então mostre que

$$1 - p = P(A^c) = \frac{(365)_k}{365^k} = \frac{365 \cdot 364 \cdot 363 \ldots (365 - k + 1)}{365^k}.$$

Note que há k fatores no numerador e no denominador dessa expressão.)

63. Mostre que a probabilidade $1 - p$ do Problema 62 pode ser escrita como

$$1 - p \approx 1 - \frac{1 + 2 + \ldots + k - 1}{365} = \frac{1 - (k-1)k}{730},$$

para k pequeno. Como ficará $P(A)$ neste caso?

64. Num mercado, três corretoras A, B e C são responsáveis por 20%, 50% e 30% do volume total de contratos negociados, respectivamente. Do volume de cada corretora, 20%, 5% e 2%, respectivamente, são contratos futuros em dólares. Um contrato é escolhido ao acaso e este é futuro em dólares. Qual é a probabilidade de ter sido negociado pela corretora A? E pela corretora C?

65. Lance uma moeda duas vezes e sejam os eventos: A: cara no primeiro lançamento, B: cara no segundo lançamento e C: as duas moedas mostram faces diferentes.
Mostre que A, B e C são dois a dois independentes, mas não totalmente independentes.

66. O Problema de Monty Hall. Num programa de TV o objetivo é ganhar um carro como prêmio. O apresentador do programa mostra a você três portas, P_1, P_2 e P_3: atrás de uma há um carro e, das outras, duas cabras. Ele pede a você para escolher uma porta, você escolhe P_1, mas esta não é aberta. Então, ele abre uma das outras duas portas e mostra uma cabra (ele sabe o que há atrás de cada porta). Então ele pergunta se você quer mudar sua escolha de porta. O que você faria?

[Sugestão: Solução informal: Faça a árvore de possibilidades. Solução formal: seja G o evento: ganhar o carro, mudando sua escolha. Seja C_i o evento: carro está atrás da porta P_i, $i = 1, 2, 3$ e seja H_i o evento: apresentador abriu a porta P_i, $i = 1, 2, 3$. Escreva G como uma reunião disjunta de dois eventos e use (5.8).]