Universidade Estadual de Campinas

Análise Combinatória, Probabilidade Noções de Estatística

Tema 2 - Espaços de Probabilidade

Prof. Laura L. R. Rifo
laurarifo at ime.unicamp.br

- Dezembro, 2015 -

Sumário

1 Experimentos aleatórios **1**

1.1 Introdução 1

1.2 Aplicações 3

Amostragem 3

Amostragem como experimento aleatório 3

Dados, moedas, baralhos e urnas 4

Confiabilidade 6

Genética 7

2 Espaço amostral e eventos **9**

2.1 Espaço amostral 9

2.2 Eventos 10

2.3 Criando novos eventos 11

A partir de mais de dois eventos 11

2.4 Classes de eventos 12

3 Variáveis aleatórias **15**

3.1 Eventos induzidos por uma variável aleatória 16

3.2 Aplicações 18

3.3 Modelos geométricos 22

4 Medida de probabilidade **25**

4.1 Probabilidade como grau de informação 25

4.2 Definição 26

4.3 Exemplos de distribuições discretas 28

4.4 Algumas propriedades 30

Lei da probabilidade total 30

Fórmula de inclusão-exclusão 31

4.5 Algumas desigualdades 32

4.6 Distribuição de uma variável aleatória 33

4.7 Exemplos 34

Moedas 34

Dados 35

Distribuição do máximo e do mínimo de variáveis uniformes 36

5 Probabilidade condicional **37**

5.1 Definição 37

5.2 Algumas propriedades 40

Regra do produto 40

Lei da probabilidade total 41

5.3 Regra de Bayes 44

6 Independência **49**

6.1 De dois eventos 49

6.2 De uma coleção de eventos 51

6.3 Independência condicional 52

6.4 De variáveis aleatórias 53

6.5 Ensaios de Bernoulli 53

A Demonstrações **57**

A.1 Desigualdade de Boole 57

Capítulo 1

Experimentos aleatórios

1.1 Introdução

A teoria da probabilidade se baseia na noção de experimento aleatório, definido como um experimento ou observação cujo resultado não é conhecido com certeza.

Esta noção é bastante ampla: tudo o que não conhecemos pode ser considerado um experimento aleatório, um experimento ou observação que será feita, ou que já aconteceu ou que está acontecendo no momento.

A observação sobre se haverá chuva amanhã ou não, ou o resultado do próximo jogo de nosso time pode ser considerado um experimento aleatório. O número de espécies marinhas abaixo de uma certa profundidade ou o nível de poluição em um certo ponto de nossa cidade neste momento também pode ser considerado um experimento aleatório, já que não dispomos de instrumentos de medição suficientemente precisos. A data ou lugar do surgimento de seres humanos no planeta, o número de troncos lingüísticos existentes na América do Sul em 1500, mesmo já tendo ocorrido, podem ser considerados experimentos aleatórios, e de fato, são objeto de inúmeros estudos antropológicos e arqueológicos. O número de nascimentos ocorridos em nossa cidade durante a última hora é um experimento aleatório, enquanto não tivermos acesso a todos os registros, e ainda com este acesso, há uma margem de incerteza referente a erros ou incompletude destes registros.

Esta é a noção de aleatoriedade que será adotada neste curso: um experimento é aleatório sempre que nossa informação a respeito dele for incompleta. Observe que, deste ponto de vista, a aleatoriedade passa a ser uma propriedade do observador, e não do fenômeno. Observadores diferentes, com graus de informação diferentes, têm possivelmente per-

cepções diferentes sobre um mesmo experimento. Um antropólogo ou um profissional da saúde, com mais informação em sua área de trabalho do que eu, têm uma idéia mais precisa do que eu sobre os experimentos exemplificados acima sobre troncos lingüísticos e nascimentos: ou seja, eles têm um grau de incerteza menor do que eu sobre estes assuntos.

Voltaremos a tratar deste assunto quando definirmos o conceito de probabilidade.

Uma descrição correta de um experimento aleatório requer uma determinação precisa do que é que está sendo observado no experimento, ou seja, uma definição do que é de fato um resultado possível.

Em muitos casos, podemos idealizar um experimento dado como uma seqüência de subexperimentos. Assim, o experimento "realizar 5 vezes o lançamento de uma moeda e observar os resultados" pode ser visto como a sequência de 5 subexperimentos "realizar um lançamento de uma moeda e observar o resultado". Neste caso, dizemos que o experimento é um **experimento composto**, e chamamos os subexperimentos de **experimentos simples**.

Um experimento simples com apenas dois possíveis resultados, como, por exemplo, a face observada no lançamento de uma moeda, é chamado experimento ou **ensaio de Bernoulli**, em homenagem ao matemático Jacob Bernoulli (em inglês). Repetições de um experimento deste tipo são chamadas uma sequência de ensaios de Bernoulli.

Se cada experimento simples tiver k possíveis resultados, como, por exemplo, a observação da face obtida no lançamento de um dado de k faces, o experimento é dito **multinomial** repetições deste experimento são chamadas uma sequência de ensaios multinomiais.

Exercícios

1. Considere o experimento de lançar n moedas diferentes e observar o resultado de cada moeda, adotando a notação 1 para cara e 0 para coroa.

   (a) Descreva o experimento como um experimento simples.

   (b) Descreva o experimento como um experimento composto com repetições independentes de um experimento simples, identificando o experimento simples.

   (c) Descreva o experimento como uma amostragem com reposição de uma população, identificando a população e o tamanho da amostra.

   (d) Descreva o experimento como n ensaios de Bernoulli.

2. O applet Coin Sample simula o experimento anterior. Rode o applet algumas vezes, para diversos valores de p, e comente os resultados obtidos.

3. Refaça a questão 1, considerando o experimento de lançar n dados diferentes, cada um com k faces numeradas de 1 a k, observando o resultado de cada dado. No item (d), troque ensaios de Bernoulli por ensaios multinomiais.

4. O applet Dice Sample simula o experimento anterior. Rode o applet algumas vezes, para $n = 5$ e diversos pesos para as faces do dado. Comente os resultados obtidos.

1.2 Aplicações

Amostragem

Na grande maioria dos estudos estatísticos, desejamos estudar uma população de interesse: pessoas com uma certa característica (proveniente de uma certa cidade, ou com uma certa doença ou dentro de uma certa faixa etária, etc.), itens produzidos por uma fábrica, produtos agropecuários de uma certa região, por exemplo.

Em geral, queremos analisar diversas características (numéricas ou não) desta população: sexo, peso e pressão sangüínea de uma pessoa, tempo de vida útil do item produzido, quantidade de fertilizante, salinidade do solo e produtividade de uma plantação de soja, e assim por diante.

Analisar a população inteira pode ser custoso ou mesmo impossível: no exemplo dos itens deveríamos testar TODA a produção para analisar a vida útil, e claramente isto não faz sentido.

Desta forma, recorremos a uma amostra da população, observando as características de interesse em cada elemento da amostra, o qual chamaremos unidade amostral.

Amostragem como experimento aleatório

Uma amostragem pode ser realizada basicamente de duas formas: com ou sem reposição.

Na primeira, cada unidade amostral é devolvida à população antes de extrair a próxima, de modo que um único objeto pode aparecer diversas vezes na amostra. Isto ocorre, por exemplo, quando amostramos exemplares de uma determinada espécie em uma reserva, a cada certo tempo, marcando os indivíduos selecionados.

Na segunda forma, sem reposição, as unidades amostrais não são devolvidas à população durante a amostragem. Isto ocorre tipicamente em alguns experimentos de controle de qualidade em que o item testado é destruído.

Podemos imaginar o processo de amostragem como um experimento composto, baseado na repetição do experimento simples de extrair um único objeto da população e observar as características de interesse.

Em uma amostragem com reposição, as repetições podem ser consideradas independentes entre si, enquanto que em uma amostragem sem reposição, o experimento consiste em etapas dependentes entre si. (A definição formal de independência será vista mais tarde.)

Dados, moedas, baralhos e urnas

Os experimentos clássicos de observar a face obtida no lançamento de uma moeda ou de um dado, ou o resultado da extração de uma carta de um baralho ou a cor de uma bolinha extraída de uma urna, por exemplo, permitem construir modelos matemáticos simples para fenômenos reais mais complexos.

No applet Coin Sample é possível simular uma seqüência de n lançamentos de uma moeda com probabilidade p de obter cara em cada lançamento individual.

No applet Dice Sample temos um experimento análogo com dado de seis faces; clicando no dado, é possível alterar as probabilidades de cada face, de acordo com seis modelos possíveis.

Note que um baralho comum pode ser representado como o espaço produto

$$\Omega = \{\text{As}, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K\} \times \{\clubsuit, \heartsuit, \diamondsuit, \spadesuit\}.$$

O applet Card simula uma extraçao de n cartas deste baralho.

O software Probabilidade com urnas, do projeto Matemática Multimídia [12], permite simular extrações de bolinhas de uma urna, com ou sem reposição, e apresenta o modelo conhecido como urna de Pólya.

Modelos de urnas (extrações de bolinhas de uma urna) podem ser vistos como modelos matemáticos para amostragens de populações finitas, como veremos durante o curso.

1. Considere o experimento de lançar um dado comum de 6 faces e então lançar uma moeda o número de vezes obtido no dado, observando a seqüência de resultados da moeda (1 para cara e 0 para coroa). Descreva o experimento como etapas sucessivas de experimentos simples, identificando estes experimentos simples.

2. O applet Die-Coin Sample simula o experimento anterior para um dado de 6 faces, e uma moeda com probabilidade $p \in [0, 1]$ de obter cara em um lançamento.

   (a) Rode o applet algumas vezes; o que significam os valores que aparecem nas 3 colunas da janela inferior esquerda?

   (b) Para $p = 0.5$, simule o experimento diversas vezes. O que acontece com as frequências de Y?

   (c) Repita o item anterior para $p = 0.6$; $p = 0.7$; $p = 0.8$. O que acontece com as frequências de Y à medida que p cresce? O que deveria acontecer com $p = 1$?

   (d) Repita o item anterior para valores decrescentes de p. O que deveria acontecer com $p = 0$?

3. Considere o experimento em que: uma moeda é lançada; se o resultado for cara, é lançado um dado vermelho observando seu resultado; se for coroa, é lançado um dado verde observando seu resultado. Descreva o experimento como um experimento composto.

4. O applet Coin-Die Sample simula o experimento anterior.

   (a) Rode o applet algumas vezes; o que significam os valores que aparecem nas 3 colunas da janela inferior esquerda?

   (b) Para $p = 0.5$, simule o experimento diversas vezes. O que acontece com as frequências de Y?

   (c) Repita o item anterior para $p = 0.6$; $p = 0.7$; $p = 0.8$. O que acontece com as frequências de Y à medida que p cresce? O que deveria acontecer com $p = 1$?

   (d) Repita o item anterior para valores decrescentes de p. O que deveria acontecer com $p = 0$?

   (e) Repita os itens anteriores para diversas distribuições das faces dos dados.

5. Considere o experimento de extrair um grupo de n cartas de um baralho comum.

   (a) Descreva o experimento como um experimento simples.

   (b) Descreva o experimento como um experimento composto.

   (c) Descreva o experimento como uma amostragem sem reposição de uma população, identificando a população e o tamanho da amostra.

   (d) Usando o applet Card Sample, tome $n = 5$ e determine a frequência de algum evento específico (presença de pelo menos um ás, presença de um certo naipe,

a soma dos valores é maior que 18, o valor mínimo observado está entre 3 e 5, inclusive, etc), em 20 rodadas.

6. No applet Urn Sample, explique o que representam os valores que aparecem na janela inferior esquerda. Rode o aplicativo algumas vezes para diversos valores de m , r, e descreva os resultados que aparecem no gráfico. Repita o anterior para extrações com reposição.

Confiabilidade

No modelo usual de estudos em confiabilidade, um sistema consiste em n componentes, cada um deles ou funcionando bem ou com defeito. Se o status de cada componente for desconhecido, isto define um experimento aleatório.

O funcionamento do sistema como um todo depende do status das componentes e de como elas estão conectadas entre si. Por exemplo, um sistema em série funciona se e somente se todas as componentes estiverem funcionando, enquanto que um sistema em paralelo funciona se e somente se pelo menos uma componente estiver funcionando.

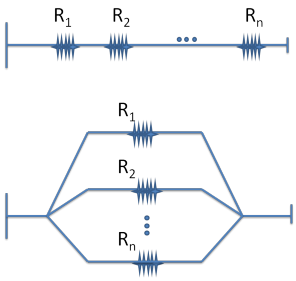

Figura 1.1: Diagrama de dois sistemas com n componentes: o de cima, em série, o de baixo, em paralelo.

Mais geralmente, um sistema k-de-n funciona se ao menos k componentes estiverem funcionando.

Exemplo. Dados naturais $k \leq n$, considere o modelo de confiabilidade k-de-n. Quais valores de k representam um sistema em série? E um sistema em paralelo?

O modelo de confiabilidade definido acima é um modelo estático, ou seja, o status das componentes não varia com o tempo. Podemos estender esta definição para um modelo dinâmico: inicialmente todas as componentes estão funcionando, mas em um instante desconhecido (e portanto aleatório) uma componente qualquer pode falhar. O

sistema como um todo também pode ter um instante de falha aleatório que depende dos tempos de falha das componentes e da estrutura do sistema, exigindo uma modelagem matemática mais elaborada.

Genética

Em sistemas de reprodução sexuada, o material genético de um filho é uma combinação desconhecida (e portanto aleatória) do material genético dos pais. Em particular, o nascimento de um filho pode ser considerado um experimento aleatório com relação a resultados como cor dos olhos, tendência a nível elevado de triglicérides e de outras características possíveis. Em geral, temos interesse por exemplo na transmissão de desordens ou características genéticas.

Consideremos um modelo simplificado de uma característica hereditária com dois possíveis estados (fenótipos), como por exemplo uma planta de ervilha cuja vagem pode ser verde ou amarela. Supondo que uma planta recebe dois alelos que formam um gene em particular para esta característica, v para verde ou a para amarelo, os possíveis genótipos são: vv, dois alelos verdes; va, um alelo verde e outro amarelo, e aa, dois alelos amarelos.

Os genótipos vv e aa são chamados homozigotos, já que os dois alelos são iguais, e o genótipo va, heterozigoto, pois os alelos são diferentes. Em muitos casos, um dos alelos da característica é dominante e o outro recessivo. Se, por exemplo, o verde for um alelo dominante para a cor da vagem, então uma planta com genótipo vv ou va terá vagens verdes, enquanto que uma com genótipo aa terá vagens amarelas. Os genes são passados para os descendentes de forma (que, para nós, pode ser considerada) aleatória, de modo que cada nova planta pode ser vista como um experimento aleatório com respeito à cor da vagem.

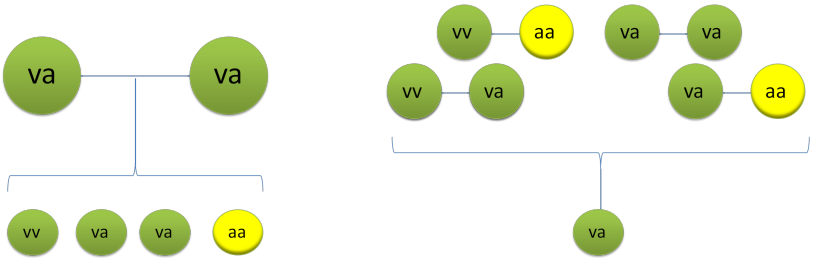

Figura 1.2: Diagrama de duas situações de possíveis genótipos: para os filhos, à esquerda, e para os pais, à direita.

Conhecer os genes dos pais não nos permite afirmar certamente qual será o genótipo do

filho, ou, inversamente, conhecendo o genótipo do filho, existem diversas possibilidades para os genótipos dos pais (e que são analisadas em testes de paternidade). Desta forma, podemos considerar o genótipo desconhecido como um experimento aleatório.

Capítulo 2

Espaço amostral e eventos

2.1 Espaço amostral

O espaço amostral de um experimento aleatório é um conjunto Ω contendo todos os possíveis resultados do experimento. Um elemento $\omega \in \Omega$ é chamado evento elementar.

Para experimentos simples, o espaço amostral pode ser exatamente o conjunto de todos os resultados possíveis, mas em modelos matemáticos mais complexos, o espaço amostral poderia conter mais elementos se for conveniente.

Por exemplo, se o experimento for lançar um dado e observar a face obtida, o espaço amostral pode ser definido como $\Omega = \{1, 2, 3, 4, 5, 6\}$, mas se o experimento for medir o peso de seu gato de estimação, poderíamos definir como espaço amostral o intervalo $\Omega = (0, \infty)$, mesmo que a maioria de seus elementos seja praticamente impossível.

Se o resultado de um experimento entregar informação sobre diversas variáveis, então o espaço amostral contém as seqüências de valores que poderiam ser observadas. Por exemplo, se um experimento consiste em medir o peso, o comprimento do pelo e a cor do seu gato de estimação então o espaço amostral é formado por vetores com três componentes indicando cada uma destas características. Assim, um evento elementar poderia ser o vetor (4kg, pelo médio, laranja e branco com manchas pretas).

Neste caso, se tivermos informação sobre n variáveis entregue pelo experimento, podemos considerar o espaço amostral como o produto cartesiano $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, onde Ω_i é o espaço amostral relacionado à i-ésima variável.

Analogamente, se tivermos n repetições de um mesmo experimento, com espaço amostral Ω, então Ω^n é o espaço amostral natural para o experimento composto, ou seja, para o experimento que consiste em n repetições do experimento original.

Por exemplo, se considerarmos o experimento de lançar uma moeda 7 vezes, então o espaço amostral Ω consiste em todas as seqüências de caras e coroas, com 7 componentes. Por outro lado, podemos ver este conjunto como o produto cartesiano do espaço amostral mais simples, Ω_i, consistindo de apenas dois elementos, cara e coroa. Denotando cara por C e coroa por K, temos

$$\begin{aligned}\Omega &= \{CCCCCCC, CCCCCCK, CCCCCKC, \ldots, KKKKKKK\} \\ &= \{C,K\} \times \{C,K\} \times \cdots \times \{C,K\} = \{C,K\}^7.\end{aligned}$$

Ou seja, este conjunto tem 2^7 elementos.

Vemos neste exemplo que a forma de descrever um espaço amostral pode nos ajudar na contagem de seus elementos.

2.2 Eventos

Chamamos evento qualquer conjunto observável de possíveis resultados do experimento, ou seja, qualquer subconjunto observável do espaço amostral Ω.

Cada vez que o experimento é realizado, diremos que um evento A ocorre se o resultado observado for um elemento de A, e diremos que não ocorre se o resultado observado não for um elemento de A.

Em particular, são eventos o próprio espaço amostral Ω, que por definição é o evento que sempre ocorre, e o conjunto vazio $\emptyset$, que por definição é o evento que nunca ocorre.

No exemplo dos 7 lançamentos de uma moeda, um possível evento é "obter uma única cara", definido pelo conjunto

$$A = \{CKKKKKK, KCKKKKK, KKCKKKK, KKKCKKK,$$

$$KKKKCKK, KKKKKCK, KKKKKKC\}.$$

Denotaremos por $\mathscr{F}$ o conjunto de todos os possíveis eventos associados ao experimento aleatório.

Exercícios

1. Um experimento consiste em lançar um dado comum de 6 faces, até aparecer face 3 ou 5. Seja A o evento em que a última face do experimento é 5 e não 3. Defina o espaço amostral Ω e descreva o evento A como subconjunto de Ω.

2. Um experimento consiste em lançar dois dados comuns de 6 faces, até que a soma obtida seja 5 ou 7. Seja A o evento em que a soma é 5 e não 7 no último lançamento. Suponha que são registrados os pares obtidos em cada lançamento. Defina o espaço amostral Ω e descreva o evento A como subconjunto de Ω.

3. No exercício anterior, suponha que apenas o último par é registrado. Defina o espaço amostral Ω e descreva o evento A como subconjunto de Ω.

2.3 Criando novos eventos

As propriedades e operações entre conjuntos, vistas na primeira parte do curso, permitem descrever e contruir novos eventos a partir de eventos dados.

Lembremos que um evento ocorre em uma realização do experimento se for observado um evento elementar pertencente ao evento.

Assim, por exemplo, dado um evento A, o evento A^C é o evento que ocorre se e somente se A não ocorrer, já que $\omega \in A^C$ se e somente se $\omega \notin A$.

Do mesmo modo, dados os eventos A e B, o evento $A \cup B$ é o evento que ocorre se pelo menos um dos eventos A ou B ocorrer, e $A \cap B$ é o evento que ocorre se ambos os eventos A e B ocorrerem.

Diremos que dois eventos A e B são mutuamente exclusivos se eles não puderem ocorrer conjuntamente (se um deles ocorrer, o outro não pode ocorrer), ou seja, se $A \cap B$ for o evento que nunca ocorre, $\emptyset$.

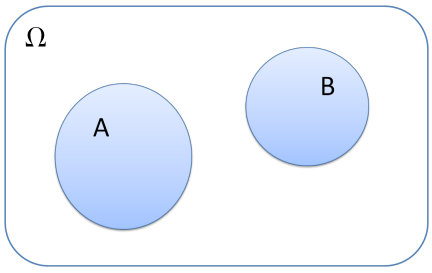

Figura 2.1: Diagrama de dois eventos, A e B, mutuamente exclusivos.

A partir de mais de dois eventos

A definição anterior continua válida para a união e a interseção de mais de dois eventos.

Dados os eventos $A_1, A_2, \ldots, A_n$, $\cup A_i$ é o evento que ocorre se pelo menos um dos eventos A_i ocorrer; $\cap A_i$ é o evento que ocorre se todos os eventos A_i ocorrerem. Formalmente,

$$\omega \in \cup_{i=1}^{n} A_i \text{ se e somente se } \omega \in A_i, \text{ para algum } i \in \{1, 2, \ldots, n\},$$

$$\omega \in \cap_{i=1}^{n} A_i \text{ se e somente se } \omega \in A_i, \text{ para todo } i \in \{1, 2, \ldots, n\}.$$

Consideremos uma coleção enumerável de eventos $\mathscr{C} = \{A_1, A_2, \ldots\}$ de um experimento aleatório.

A união desta coleção $\cup\mathscr{C}$ é o evento que ocorre se e somente se pelo menos um evento da coleção ocorrer.

De fato, consideremos uma realização do experimento, com resultado observado ω. Então, $\cup\mathscr{C}$ ocorre se e somente se $\omega \in \cup\mathscr{C}$. Isto significa que $\omega \in A_n$, para algum $n \geq 1$, que é equivalente a afirmar que A_n ocorre, para algum $n \geq 1$.

Analogamente, a interseção desta coleção $\cap\mathscr{C}$ é o evento que ocorre se e somente se todos os eventos da coleção ocorrerem.

De fato, consideremos o evento complementar $(\cap_n A_n)^C = \cup_n A_n^C$. Pela afirmação anterior, $\cup_n A_n^C$ ocorre se e somente se pelo menos um evento A_n^C ocorrer, ou seja, se pelo menos um evento A_n não ocorrer. Assim $\cap_n A_n = (\cup_n A_n^C)^C$ ocorre se e somente se nenhum dos eventos A_n^C ocorrer, ou seja, se todos os eventos A_n ocorrerem.

Por exemplo, se os A_n's forem os intervalos $[0, 1], [0, 1/2], [0, 1/3], \ldots, [0, 1/n], \ldots$, então, $\cup\mathscr{C} = [0, 1]$ e $\cap\mathscr{C} = \{0\}$.

2.4 Classes de eventos

Consideremos um experimento aleatório $\mathcal{E}$ com espaço amostral Ω. Seja $\mathscr{F}$ uma classe de subconjuntos de Ω.

Dizemos que $\mathscr{F}$ é uma classe de eventos observáveis se forem satisfeitas as seguintes condições:

O1 $\Omega \in \mathscr{F}$;

O2 se $A \in \mathscr{F}$, então $A^C = \Omega \setminus A \in \mathscr{F}$;

O3 se $A, B \in \mathscr{F}$, então $A \cup B \in \mathscr{F}$, e mais geralmente

O3' se $A_1, A_2, \cdots \in \mathscr{F}$, então $\cup A_n \in \mathscr{F}$.

Estas condições nos garantem que os eventos observáveis estão bem e coerentemente definidos. Assim, a condição (O1) nos afirma que, se o experimento fosse realizado, algum dos resultados do espaço amostral deveria ser observado, o que é coerente com o fato do espaço amostral conter todos os resultados possíveis.

A condição (O2) diz que se somos capazes de afirmar se um evento A ocorre, então também somos capazes de afirmar se o evento "não A" ocorre. Finalmente, com a condição (O3), se somos capazes de afirmar se o evento A ocorre e se o evento B ocorre (cada um separadamente), então também somos capazes de afirmar se pelo menos um deles ocorre.

Exemplo. Consideremos o experimento aleatório $\mathcal{E}$: "lançar uma moeda e observar o resultado obtido", com espaço amostral $\Omega = \{C, K\}$. A classe de subconjuntos $\{\{C\}, \{K\}\}$, que observa o resultado C e o resultado K, não é uma classe de eventos, pois não satisfaz a condição O1. Se acrescentarmos Ω, $\{\{C\}, \{K\}, \Omega\}$ ainda não é suficiente, pois agora não satisfaz a condição O2. Completando o que falta, uma classe válida de eventos é $\mathscr{F} = \{\emptyset, \{C\}, \{K\}, \Omega\}$.

Observe que a classe $\mathscr{F} = \{\emptyset, \Omega\}$ é uma classe de eventos válida, qualquer que seja o espaço amostral. Podemos interpretar esta classe como uma classe não informativa sobre o resultado do experimento: o único que sabemos é que ocorre algum dos resultados possíveis, mas não sabemos qual.

A classe de eventos nos indica quais eventos somos capazes de observar ao realizar o experimento. Isto nos permite representar, por exemplo, uma informação parcial sobre o experimento.

Exemplo. Consideremos o experimento aleatório $\mathcal{E}$: "lançar um dado de 6 faces e observar o resultado obtido". Relacionado com o espaço amostral $\Omega = \{1, 2, 3, 4, 5, 6\}$, podemos definir diferentes classes de eventos.
(a) $\mathscr{F} = \{\emptyset, \Omega\}$, que não nos informa nada sobre o resultado obtido, apenas que ele pertence a Ω;
(b) $\mathscr{F} = \mathcal{P}(\Omega)$, o conjunto das partes ou potência de Ω, cujos elementos são todos os subconjuntos de Ω:

$$\mathscr{F} = \{\emptyset, \{1\}, \{2\}, \ldots, \{6\}, \{1, 2\}, \ldots, \{5, 6\}, \{1, 2, 3\}, \ldots, \{1, 2, 3, 4, 5, 6\}\}.$$

Como todos os subconjuntos são eventos observáveis, isto implica que podemos ter uma informação total sobre o resultado do experimento.

(c) Suponha agora que temos uma informação parcial do experimento; por exemplo, suponha que as faces 3, 4, 5 e 6 estão apagadas no dado, e que só podemos identificar as faces 1 e 2. Assim, uma classe de eventos representando esta informação parcial é

$$\mathscr{F} = \{\emptyset, \{1\}, \{2\}, \{1,2\}, \{1\}^C, \{2\}^C, \{1,2\}^C, \Omega\}.$$

Observação Uma classe de subconjuntos satisfazendo as condições (O1-O3) é chamada uma álgebra; se satisfizer também a condição (O3'), é chamada uma σ-álgebra. Para o leitor interessado, uma referência nesta linha é o livro [3], e as referências lá citadas.

Exercícios

Nos exercícios seguintes, A e B são eventos.

1. Mostre que $A \subset B$ se e somente se a ocorrência do evento A implica a ocorrência do evento B.

2. Mostre que $A \setminus B$ é o evento que ocorre se e somente se A ocorre e B não ocorre.

3. Mostre que $(A \cap B^C) \cup (A^C \cap B)$ é o evento que ocorre se e somente se exatamente um entre A e B ocorrer. Este evento é chamado a diferença simétrica entre A e B, e é denotado por $A \triangle B$.

4. Mostre que $(A \cap B) \cup (A \cup B)^C$ é o evento que ocorre se e somente se ou ambos ou nenhum dos eventos A ou B ocorrerem.

5. Mostre em um diagrama de Euler-Venn todos os 16 eventos que podem ser construídos a partir de A e B.

6. Considere o experimento de dois lançamentos de um dado comum de 6 faces, observando ambos os resultados. Sejam Ω o espaço amostral, A o evento de que o resultado do primeiro lançamento é igual a 1, e B o evento de que a soma dos dois resultados obtidos é igual a 7. Descreva todos os elementos de: Ω, A, B, $A \cup B$, $A \cap B$, $A \setminus B$, $A^C \cap B^C$.

7. Nos exemplos vistos até o momento, construa duas classes de eventos válidas (não triviais), e uma classe de subconjuntos que não seja uma classe de eventos.

Capítulo 3

Variáveis aleatórias

Consideremos um experimento aleatório $\mathcal{E}$ com espaço amostral Ω.

Em muitos casos, estamos interessados em características numéricas associadas a um resultado $\omega \in \Omega$.

Uma função real definida em Ω, $X : \Omega \to \mathbb{R}$, é chamada variável aleatória. Denotaremos estas funções usualmente por letras maiúsculas da segunda metade do alfabeto.

Uma variável aleatória em si pode também ser considerada um experimento aleatório, já que seu valor (desconhecido) depende do resultado (desconhecido) do experimento original. Inversamente, se os resultados de um experimento aleatório forem valores numéricos, então o resultado pode ser considerado uma variável aleatória.

Exemplo. Considere o experimento de lançar um dado e observar a face obtida. O espaço amostral é um subconjunto real, $\Omega = \{1, 2, 3, 4, 5, 6\}$. Portanto a função X que indica a face observada é uma variável aleatória.

Exemplo. Considere o experimento de lançar uma moeda 2 vezes e observar a seqüência das faces obtidas. A função real X que indica o número de caras de uma seqüência observada é uma variável aleatória, representada na Figura 3.1.

Quando o experimento é realizado e observamos o resultado ω, a variável aleatória assume o valor $X(\omega) = x$. Denotaremos por χ o conjunto dos possíveis valores assumidos por X.

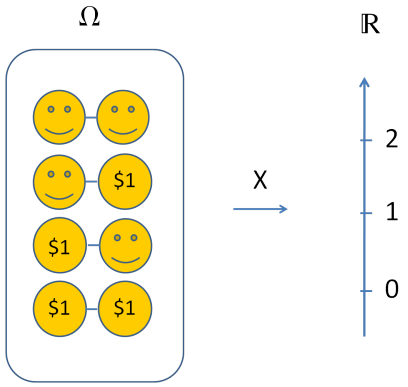

Figura 3.1: Diagrama de uma função (variável aleatória) X entre os conjuntos Ω e $\mathbb{R}$.

3.1 Eventos induzidos por uma variável aleatória

Denotemos por $\mathscr{A}$ o conjunto de eventos em Ω e por $\mathscr{B}$ o conjunto de eventos em $\mathbb{R}$.

Dado um evento $B \in \mathscr{B}$, denotaremos por $(X \in B)$ o conjunto imagem inversa de B, ou seja,

$$(X \in B) = \{\omega \in \Omega : X(\omega) \in B\}$$

é o conjunto de resultados do experimento que tem a característica X com valor em B.

Dois casos particulares importantes desta notação são os eventos em Ω

$$(X = x) = \{\omega \in \Omega : X(\omega) = x\},$$

o conjunto de resultados do experimento com característica X exatamente igual a x, e

$$(X \leq x) = \{\omega \in \Omega : X(\omega) \leq x\},$$

o conjunto de resultados do experimento com característica X menor ou igual a x.

Exemplo. No exemplo dos 2 lançamentos de uma moeda, o evento $(X = 1)$ é o conjunto de seqüências em Ω que apresentam uma única cara,

$$(X = 1) = \{CK, KC\},$$

onde C denota cara e K coroa. O evento $(X \leq 1)$ é o conjunto de seqüências em Ω que apresentam no máximo uma cara,

$$(X \leq 1) = \{KK, CK, KC\}.$$

Podemos generalizar o conceito de variável aleatória para uma função observável $X : \Omega \to \mathcal{X}$, onde $\mathcal{X}$ é um outro conjunto, não necessariamente real. Em particular, se $\mathcal{X} \subset \mathbb{R}^n$, podemos chamar esta função de vetor aleatório. O importante é que ela seja uma função de um espaço amostral (mesmo que não seja mostrado explicitamente).

Suponha que temos um experimento aleatório com espaço amostral Ω, e uma variável aleatória $X : \Omega \to \mathbb{R}$, e seja $f : \mathbb{R} \to \mathbb{R}$ uma função real definida em $\mathbb{R}$. Então $Y = f(X)$ também é uma variável aleatória.

Uma destas funções bastante útil no cálculo de probabilidades é a chamada função indicadora de um evento A dado, denotada por $\mathbf{1}_A$, e definida como

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \text{se } \omega \in A \\ 0 & \text{caso contrário} \end{cases},$$

ou simplesmente

$$\mathbf{1}_A = \begin{cases} 1 & \text{se } A \text{ ocorre} \\ 0 & \text{se não} \end{cases}.$$

Exercícios

Assuma que X é uma variável aleatória e que A e B são eventos em $\mathbb{R}$. As seguintes afirmações trabalham com o conjunto imagem inversa e sua preservação por operações de conjuntos. Prove os resultados.

1. $(X \in A \cup B) = (X \in A) \cup (X \in B)$
2. $(X \in A \cap B) = (X \in A) \cap (X \in B)$
3. $(X \in A \setminus B) = (X \in A) \setminus (X \in B)$
4. Se A e B são disjuntos então $(X \in A)$ e $(X \in B)$ também são.
5. $\mathbf{1}_{A \cap B} = \mathbf{1}_A \mathbf{1}_B = \min\{\mathbf{1}_A, \mathbf{1}_B\}$
6. $\mathbf{1}_{A \cup B} = 1 - (1 - \mathbf{1}_A)(1 - \mathbf{1}_B) = \max\{\mathbf{1}_A, \mathbf{1}_B\}$
7. $\mathbf{1}_{A \setminus B} = \mathbf{1}_A(1 - \mathbf{1}_B)$
8. $\mathbf{1}_{A^C} = 1 - \mathbf{1}_A$
9. $A \subset B$ se e somente se $\mathbf{1}_A \leq \mathbf{1}_B$.

3.2 Aplicações

Os exemplos que veremos geralmente tratarão de problemas com moedas e dados, por sua relativa simplicidade matemática. No entanto, não devemos esquecer que estes modelos podem ser vistos como uma primeira resolução para problemas reais mais complexos.

Lançamentos de uma moeda

Um experimento básico com moedas é o de n lançamentos sucessivos de uma moeda, obtendo como resultado do experimento uma seqüência $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ de zeros e uns, onde 0 denota coroa e 1 denota cara, por exemplo. Esta notação é útil, já que permite obter algumas características do experimento de maneira rápida. Por exemplo, se quisermos o total de caras obtidas nos n lançamentos, digamos S, basta observar que $S = X_1 + X_2 + \cdots + X_n$, e se quisermos o total de coroas, basta obter $n - S$.

O applet Coin Sample realiza este experimento, permitindo ver um padrão nas respostas obtidas. Por exemplo, selecione $n = 6$ lançamentos com $p = 0,5$, o que indica que você lançará 6 vezes uma moeda balanceada (com mesma chance de obter cara ou coroa em um lançamento qualquer). Rode o programa vinte vezes, e veja quantas vezes ocorreu o evento $(S = 2)$. Depois selecione outros valores de p e veja o que ocorre com a freqüência deste evento ao repetir o experimento várias vezes.

Lançamentos de um dado

Uma generalização natural é considerar n lançamentos de um dado de k lados (que pode ser visto como uma moeda com k faces). Este tipo de experimento é chamado uma seqüência de ensaios multinomiais. O caso especial de $k = 6$ corresponde a um dado comum de 6 faces.

O applet Dice Sample realiza este experimento com um dado de 6 faces, permitindo ver algum padrão nas respostas obtidas. Por exemplo, selecione $n = 2$ e rode o programa diversas vezes. O que ocorre com a freqüência do evento $A =$"o resultado do primeiro lançamento é par"?

O experimento Jogo dos Divisores, construído pelo projeto Matemática Multimídia [12], define funções numéricas a partir das faces obtidas no lançamento de um dado comum.

Experimento composto dado-moeda

Consideremos agora o experimento em dois estágios dado-moeda: lançamos um dado e depois lançamos uma moeda o total de vezes que foi obtido no dado. Registramos a seqüência **X** de resultados da moeda. Seja N a variável aleatória que denota o valor obtido no dado e S o total de caras obtidas nos lançamentos da moeda.

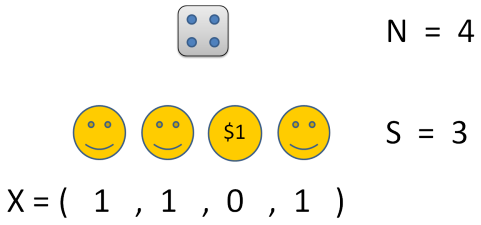

Figura 3.2: Experimento de lançar um dado e uma moeda.

Determine o espaço amostral Ω e $\#\Omega$. Expresse N e S como funções definidas em Ω. Liste os elementos do evento $(S = 5)$.

Resposta:
$\Omega = \{1, 0, 11, 10, 01, 00, 111, 110, 101, 011, 100, 010, 001, 000, 1111, \ldots, 000000\}$ tem

$$\#\Omega = 2 + 2^2 + 2^3 + 2^4 + 2^5 + 2^6 = 126$$

elementos. A variável aleatória N é a função

$$\begin{aligned}
&N(1) = N(0) = 1\\
&N(11) = N(10) = N(01) = N(00) = 2\\
&N(111) = N(110) = N(101) = N(011) = \cdots = N(000) = 3\\
&N(1111) = N(1110) = N(1101) = N(1011) = \cdots = N(0000) = 4\\
&\vdots\\
&N(111111) = N(111110) = \cdots = N(000000) = 6
\end{aligned}$$

e S é

$$\begin{aligned}
&S(0) = S(00) = S(000) = S(0000) = S(00000) = S(000000) = 0\\
&S(1) = S(10) = S(01) = S(100) = S(010) = \cdots = S(000001) = 1\\
&S(11) = S(110) = S(101) = S(011) = \cdots = S(000011) = 2\\
&S(111) = S(1110) = S(1101) = S(1011) = \cdots = S(000111) = 3\\
&\vdots\\
&S(111111) = 6
\end{aligned}$$

O evento $(S = 5)$ é descrito como o conjunto

$$(S = 5) = \{11111, 111110, 111101, 111011, 110111, 101111, 011111\}.$$

Rode o aplicativo Die-Coin Sample 10 vezes. Para cada vez, dê os valores das variáveis aleatórias X, N e S, e conte o total de vezes em que ocorre o evento A: todos os lançamentos são cara.

Exercícios

1. Considere o experimento de lançar uma moeda $n = 4$ vezes, observando a sequencia de resultados, e seja Y o número de caras obtidas.

   (a) Descreva o espaço amostral Ω, listando todos os seus elementos.

   (b) Descreva o evento $(Y = k)$, para todo k possível.

   (c) Quantos elementos tem o evento $(Y = k)$?

2. Considere o experimento anterior no caso geral de n lançamentos. Quantos elementos tem o espaço amostral? Quantos elementos tem o evento $(Y = k)$, para cada $k = 0, 1, \ldots, n$?

3. Considere o experimento de $n = 2$ lançamentos de um dado comum de 6 faces. Seja Ω o espaço amostral ao observar os dois resultados, A o evento de que o primeiro lançamento obteve face 1, e B, o evento de que a soma dos pontos obtidos é 7. Descreva cada um dos eventos abaixo na forma indicada.

   (a) Ω em forma de produto cartesiano.

   (b) A na forma de lista.

   (c) B na forma de lista.

   (d) $A \cup B$ na forma de lista.

   (e) $A \cap B$ na forma de lista.

   (f) $A^C \cap B^C$ em forma de predicado.

4. No applet Dice Sample, selecione $n = 2$ e rode o experimento 100 vezes. Conte o total de vezes que cada evento do exercício anterior ocorre.

5. No contexto do exercício anterior, sejam Y a variável aleatória que indica a soma obtida nos dois lançamentos, U a variável aleatória que indica o menor resultado e V o maior resultado obtidos nos dois lançamentos. Expresse cada uma destas

variáveis aleatórias como uma função do espaço amostral Ω e determine o conjunto de possíveis valores. Determine o conjunto de possíveis valores de (U, V) na forma de predicado.

6. No contexto do exercício anterior, denote por X_1 o resultado do primeiro lançamento e por X_2, o resultado do segundo. Descreva os elementos dos seguintes eventos como subconjuntos do espaço amostral Ω:

   (a) $(X_1 < 3, X_2 > 4)$;

   (b) $(Y = 7)$;

   (c) $(U = 2)$;

   (d) $(V = 5)$;

   (e) $(U = V - 1)$.

7. No applet Dice Sample, selecione $n = 2$ e rode o experimento 100 vezes. Conte o total de vezes que cada evento do exercício anterior ocorre.

8. Suponha que 3 dados comuns de 6 faces são lançados e que o resultado de cada um (X_1, X_2, X_3) é registrado. Uma pessoa paga \$1 para lançar os dados e recebe \$1 por cada 6 que aparecer no lançamento. Seja W o lucro dessa pessoa em uma realização do experimento. Descreva o espaço amostral Ω do experimento e expresse W como função definida em Ω.

9. Rode o aplicativo Chuck-a-luck algumas vezes, e descreva os resultados obtidos: espaço amostral, variável aleatória, evento e respectivas cardinalidades.

10. No caso geral de n lançamentos de um dado de k faces, seja Y a soma dos pontos, U o mínimo e V, o máximo dos pontos.

    (a) Descreva o espaço amostral do experimento e determine sua cardinalidade.

    (b) Expresse Y como uma função no espaço amostral, e liste seus possíveis valores.

    (c) Expresse U como uma função no espaço amostral, e liste seus possíveis valores.

    (d) Expresse V como uma função no espaço amostral, e liste seus possíveis valores.

    (e) Determine o conjunto de possíveis valores de (U, V) em forma de predicado.

11. Um experimento consiste em lançar uma moeda até obter uma cara. Seja X o total de lançamentos realizados. Determine o espaço amostral Ω do experimento, se forem observados os resultados de todos os lançamentos, e Ω_X, indicando a cardinalidade de cada conjunto.

12. Um experimento consiste em lançar um par de dados repetidas vezes até que a soma seja 5 ou 7. Seja A o evento de que a soma é 5 no último lançamento.

    (a) Suponha que o par de resultados em cada lançamento é observado. Defina o espaço amostral deste experimento e descreva A como subconjunto deste espaço amostral, indicando suas cardinalidades.

    (b) Suponha que o par de resultados do último lançamento é observado. Defina o espaço amostral deste experimento e descreva A como subconjunto deste espaço amostral, indicando suas cardinalidades.

13. Três bolas são selecionadas sem reposição de uma urna contendo 20 bolas numeradas de 1 a 20. Defina o evento A de que pelo menos uma das bolas sorteadas é maior ou igual a 17. Se cada um dos três valores for observado, determine a cardinalidade do espaço amostral e do evento A.

14. Três bolas são sorteadas de uma urna contendo 3 bolas brancas, 3 bolas vermelhas e 5 bolas pretas. Suponha que ganhemos \$1 por cada bola branca sorteada e percamos \$1 para cada bola vermelha sorteada. Seja X o saldo ao fim do sorteio. Determine o espaço amostral se forem observadas as cores das três extrações e sua cardinalidade. Determine os valores de X e a cardinalidade dos conjuntos $(X = k)$ para cada valor de k.

3.3 Modelos geométricos

Nos exemplos anteriores, nos restringimos a modelos probabilísticos discretos, ou seja, com espaço amostral finito ou infinito enumerável. O seguinte experimento, chamado moeda de Buffon, consegue dar uma boa ideia de modelos mais gerais, envolvendo espaços amostrais não enumeráveis: tipicamente, subconjuntos de $\mathbb{R}^n$.

Consideremos um quadrado de lado 1, centrado na origem, como na Figura 3.3. O experimento consiste em lançar uma moeda de raio $r \leq 1/2$, observando o centro (X, Y) da moeda.

Exercícios

1. Neste experimento, seja A o evento de que a moeda não toca os lados do quadrado, e seja Z a variável aleatória definida como a distância do centro da moeda ao centro do quadrado.

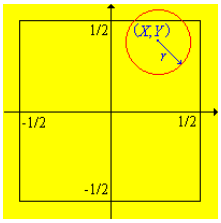

Figura 3.3: Diagrama do experimento "moeda de Buffon" (extraído de [14]).

(a) Descreva o espaço amostral Ω matematicamente.

(b) Descreva A como um subconjunto de Ω.

(c) Descreva A^C como um subconjunto de Ω.

(d) Expresse Z como função definida em Ω.

(e) Expresse o evento $(X < Y)$ como um subconjunto de Ω.

(f) Expresse o evento $(Z \leq 1/2)$ como subconjunto de Ω.

2. Rode o applet Moeda de Buffon 100 vezes, para $r = 0.2$. Para cada rodada, registre se o evento A ocorreu e o valor de Z. Quantas vezes A ocorreu?

3. Considere o experimento aleatório de escolher um ponto (X, Y) na região circular de raio 1 centrada na origem, em $\mathbb{R}^2$. Seja A o evento que o ponto selecionado está no quadrado inscrito centrado na origem, com lados paralelos aos eixos coordenados. Seja B o evento que o ponto selecionado está no quadrado inscrito com vértices em $(\pm 1, 0)$, $(0, \pm 1)$.

4. Considere o experimento aleatório de escolher um ponto X em $[-1, 1]$. Seja A o evento que o ponto escolhido é menor que $1/2$ da origem e seja Z a variável aleatória distância de X até a origem.

(a) Descreva o espaço amostral Ω matematicamente.

(b) Descreva A como um subconjunto de Ω.

(c) Descreva A^C como um subconjunto de Ω.

(d) Expresse Z como função definida em Ω.

(e) Expresse o evento $(Z \leq 1/2)$ como subconjunto de Ω.

5. Considere o experimento aleatório de escolher um ponto X em $[-1, 1]$, e seja A o evento que $X^3 + X^2 - 2X > 0$.

(a) Descreva o espaço amostral Ω matematicamente.

(b) Descreva A como um subconjunto de Ω.

(c) Descreva A^C como um subconjunto de Ω.

6. Considere o experimento aleatório de escolher um ponto $P = (X, Y)$ no segmento com extremos $(0, 2)$ e $(2, 0)$. Defina A como o evento que P está a uma distância maior que $6\sqrt{2}/5$. Seja Z a variável aleatória que indica a área do triângulo com vértices $(0, 0)$, P, $(2, 0)$.

   (a) Descreva o espaço amostral Ω matematicamente.

   (b) Descreva A como um subconjunto de Ω.

   (c) Descreva A^C como um subconjunto de Ω.

   (d) Expresse Z como função definida em Ω.

   (e) Expresse o evento $(Z \leq 1/4)$ como subconjunto de Ω.

7. Considere o experimento aleatório de escolher um ponto $P = (X, Y)$ na região circular de raio r centrada na origem, em $\mathbb{R}^2$. Considere a menor corda da circunferência com ponto médio em P, e denote por Z o seu comprimento. Grafique e expresse o evento $(Z \leq r)$ como subconjunto de Ω.

8. Considere o experimento aleatório de escolher um ponto na região circular de raio r centrada na origem, em $\mathbb{R}^2$, de acordo com suas coordenadas polares, $P = (R, \Theta)$ em $[0, r] \times [0, 2\pi]$. Considere a menor corda da circunferência com ponto médio em P, e denote por Z o seu comprimento. Grafique e expresse o evento $(Z \leq r)$ como subconjunto de Ω.

9. Considere o experimento aleatório de escolher dois pontos P e Q na circunferência de raio r centrada na origem, em $\mathbb{R}^2$. Seja Z o comprimento da corda definida por P e Q. Grafique e expresse o evento $(Z \leq r)$ como subconjunto de Ω.

Capítulo 4

Medida de probabilidade

4.1 Probabilidade como grau de informação

Dependendo do grau de informação do observador, é possível ter diversos graus de precisão sobre os possíveis resultados de um experimento aleatório. Um antropólogo, mesmo não sabendo exatamente, deve ter uma idéia mais precisa a respeito do número de troncos lingüísticos na América do Sul em 1500 do que alguém que não tem informação especializada a respeito.

Este grau de informação pode ser quantificado no que definiremos como função de probabilidade. Da discussão anterior, na maioria dos casos reais, observadores diferentes terão informações diferentes a respeito do fenômeno estudado, e portanto funções de probabilidade diferentes. Em alguns casos teóricos, no entanto, é possível que haja consenso entre diversos observadores, levando assim a uma mesma função de probabilidade para o problema estudado.

Qualquer que seja o caso, a probabilidade de um resultado reflete um grau de certeza a respeito da ocorrência desse resultado.

Diversas interpretações

Historicamente, encontramos basicamente duas interpretações para o conceito de probabilidade.

A mais antiga é a chamada interpretação freqüentista, baseada na suposição de que o experimento aleatório em questão pode ser repetido indefinidamente sob as mesmas condições. Neste caso, a probabilidade de um evento é proporcional ao limite da freqüência observada do evento nas repetições.

A segunda é a chamada interpretação subjetivista, baseada no conhecimento ou grau de informação do observador a respeito dos possíveis resultados do experimento. Se o experimento não for repetível (como é o caso da maioria das situações na prática), a interpretação frequentista fica sem sentido, e utilizamos naturalmente toda nossa informação para atribuir probabilidade a um evento de interesse.

A interpretação freqüentista pode ser vista como um caso particular da subjetivista, já que um observador poderia achar razoável atribuir para um evento uma probabilidade igual ao limite da freqüência se o experimento pudesse ser repetido.

Independentemente da interpretação, uma definição completa de uma probabilidade requer uma definição precisa do espaço amostral e do conjunto de eventos observáveis.

O processo de atribuir uma função de probabilidade aos resultados de um experimento aleatório é o que chamamos de modelagem probabilística ou estocástica.

O vídeo BrasilxArgentina mostra uma aplicação da teoria subjetivista no processo de tomada de decisão.

4.2 Definição

Uma probabilidade é uma função real definida em conjuntos. Mais precisamente, consideremos um experimento aleatório com espaço amostral Ω e conjunto de eventos observáveis $\mathscr{F}$.

Uma medida de probabilidade P em Ω é uma função real com domínio $\mathscr{F}$, $P : \mathscr{F} \to \mathbb{R}$, satisfazendo as seguintes propriedades:

P1. $P(A) \geq 0$ para todo evento $A \in \mathscr{F}$.

P2. $P(\Omega) = 1$.

P3. Dada uma coleção contável de eventos $\{A_1, A_2, \dots\}$, disjuntos dois a dois, então

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

Estas propriedades são chamadas axiomas de Kolmogorov, em homenagem ao matemático russo Andrei Kolmogorov. A terceira propriedade é conhecida como a propriedade de aditividade contável, e afirma que a probabilidade de uma coleção finita ou enumerável de eventos mutuamente exclusivos é igual à soma de suas probabilidades.

As propriedades P1 e P2 são uma convenção na qual decidimos medir a probabilidade de um evento como um número entre 0 e 1; mas a propriedade P3 é fundamental, e análoga às demais formas de medir o "tamanho" de um conjunto: cardinalidade de conjuntos finitos, comprimento de intervalos reais, área de subconjuntos em $\mathbb{R}^2$, e volume de subconjuntos em $\mathbb{R}^3$, por exemplo.

Com isto, temos os três ingredientes necessários para modelar matematicamente um experimento aleatório:

- um espaço amostral, Ω;
- uma coleção de eventos observáveis, $\mathscr{F}$;
- uma função de probabilidade que atribui um grau de certeza para cada um destes eventos observáveis, P.

Esta terna, $(\Omega, \mathscr{F}, P)$, é o que chamamos um espaço de probabilidade. A função P será chamada indistintamente de medida, distribuição ou lei de probabilidade.

Exercícios

Suponha que temos um experimento aleatório com espaço amostral Ω e uma medida de probabilidade P. Nos seguintes exercícios, A e B são eventos. Prove os seguintes resultados usando os axiomas de Kolmogorov.

1. Regra do complementar. $P(A^C) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. Regra da diferença. $P(B \setminus A) = P(B) - P(A \cap B)$.
4. Se $A \subset B$ então $P(B \setminus A) = P(B) - P(A)$.
5. A probabilidade é uma função crescente relativa à ordem parcial dos conjuntos, ou seja, se $A \subset B$ então $P(A) \leq P(B)$. Em particular, $P(A) \leq 1$ para todo evento A.
6. Suponha que $A \subset B$.
   (a) Se $P(B) = 0$ então $P(A) = 0$.
   (b) Se $P(A) = 1$ então $P(B) = 1$.

7. Se $P(A) = 0$ então $P(A \cup B) = P(B)$.
   (Observe que **não** estamos dizendo que $A \cup B = B$, mas apenas que suas probabilidades são iguais. Também **não** estamos dizendo que $A = \emptyset$, apenas que sua probabilidade é zero. Você pode visualizar este resultado e o próximo com um exemplo de modelos geométricos, vistos na Seção 3.3.)

8. Se $P(A) = 1$ então $P(A \cap B) = P(B)$.

4.3 Exemplos de distribuições discretas

Dizemos que uma medida de probabilidade é **discreta** se o espaço amostral associado, Ω, for finito ou infinito enumerável.

Distribuição uniforme discreta

Suponhamos que Ω é um conjunto finito e não-vazio. Sob algumas condições, pode ser razoável considerar o modelo matemático de que todos os resultados elementares de Ω têm a mesma chance de ocorrer: por exemplo, em um lançamento de um dado simétrico, é razoável supor que todas as faces têm a mesma chance; em uma extração de cartas de um baralho ou de bolinhas de uma urna, é razoável supor que todas as cartas (ou todas as bolinhas) têm a mesma chance de serem extraídas.

Chamamos este tipo de modelo probabilístico de equiprovável ou uniforme.

Assim, se Ω tiver n elementos, neste modelo a probabilidade de cada elemento $\omega \in \Omega$ é $P(\{\omega\}) = 1/n$. (Denotaremos $P(\{\omega\})$ simplesmente por $P(\omega)$.)

Observe que, neste caso, a probabilidade de um evento A qualquer é proporcional à quantidade de elementos que ele contém: se o evento A tiver o dobro de elementos que o evento B, então sua probabilidade também deve ser o dobro da de B. Daqui a importância de construir formas eficientes de contagem.

Definimos, desta forma, a distribuição uniforme em Ω como

$$P(A) = \frac{\#A}{\#\Omega}, \text{ para todo evento } A \subset \Omega\,.$$

Esta função é particularmente importante em experimentos amostrais e combinatórios, como os exemplificados anteriormente.

Exemplo. Considere o experimento de lançar uma moeda simétrica, ou seja, nenhuma das faces tem preferência sobre a outra. Neste caso,

$$P(C) = P(K) = 1/2\,.$$

Exemplo. Considere o experimento de lançar um dado simétrico com 6 faces. Neste caso, nosso modelo é equiprovável sobre o conjunto $\{1, 2, 3, 4, 5, 6\}$. Desta forma, a probabilidade de obter face par é

$$P(\{2, 4, 6\}) = 3/6 = 1/2\,.$$

Observe que, pelo axioma P3, poderíamos ter calculado esta probabilidade como

$$P(\{2, 4, 6\}) = P(2) + P(4) + P(6) = 1/6 + 1/6 + 1/6 = 1/2\,.$$

Exercícios

Para todos os exemplo e exercícios dos capítulos anteriores, determine as probabilidades dos eventos considerados, supondo distribuição uniforme no espaço amostral correspondente.

Distribuição discreta geral

Novamente pelo axioma P3, se Ω for um conjunto discreto e não-vazio, podemos construir uma função de probabilidade em $\mathscr{A}$ conhecendo a probabilidade de todos os eventos elementares $\omega \in \Omega$, $P(\omega)$.

Neste caso, temos que a probabilidade de um evento é a soma das probabilidades de seus elementos:

$$P(A) = \sum_{\omega \in A} P(\omega)\,.$$

Exemplo. Considere o experimento de lançar um dado com 6 faces numeradas de 1 a 6, não-simétrico, de modo que cada face tenha probabilidade proporcional a seu valor. Neste caso, o modelo equiprovável sobre o conjunto $\{1, 2, 3, 4, 5, 6\}$ já não é apropriado. Em particular, para determinar a probabilidade de obter face par não basta apenas contar os elementos deste evento:

$$P(\{2, 4, 6\}) \neq 3/6\,.$$

Precisamos considerar a probabilidade de cada um de seus elementos. Pelo axioma P3 (que vale para qualquer modelo probabilístico), temos

$$P(\{2,4,6\}) = P(2) + P(4) + P(6)\,.$$

Por outro lado, pela informação dada de que cada face tem probabilidade proporcional ao seu valor, podemos deduzir que, para cada $k \in \{1,2,3,4,5,6\}$,

$$P(k) = k/(1+2+3+4+5+6)\,,$$

pois lembre que a soma de todas as probabilidades deve ser igual a 1.

Assim,

$$P(\{2,4,6\}) = (2+4+6)/(1+2+3+4+5+6) = 12/21\,.$$

Em geral, se a probabilidade da face k for proporcional a um valor w_k, teremos que

$$P(k) = w_k/(w_1+w_2+w_3+w_4+w_5+w_6)\,.$$

4.4 Algumas propriedades

Lei da probabilidade total

Generalizemos a idéia anterior de escrever um conjunto A como a união disjunta de seus elementos,

$$A = \bigcup_{a\in A} \{a\}\,.$$

Para isso, consideremos uma partição finita de Ω, $\{A_1, A_2, \ldots, A_n\}$, ou seja, Ω pode ser escrito como a união disjunta

$$\Omega = \bigcup_{i=1}^{n} A_i\,.$$

Observe que, para qualquer evento B, podemos então escrever

$$B = \bigcup_{i=1}^{n} (B \cap A_i)\,.$$

Como esta é uma união dos eventos disjuntos $B \cap A_i$, pelo axioma P3,

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i).$$

Esta igualdade é conhecida como **lei da probabilidade total**, e é útil quando as probabilidades das interseções são conhecidas. Esta lei pode ser ainda generalizada para uma partição inifinita enumerável de Ω.

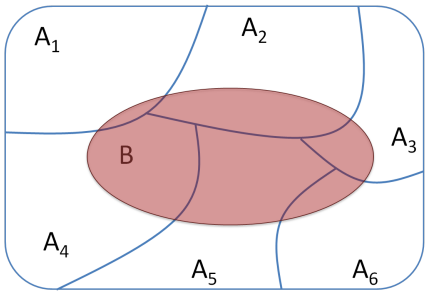

Figura 4.1: Lei da probabilidade total.

Fórmula de inclusão-exclusão

A fórmula de inclusão-exclusão, vista para a medida de contagem, se aplica também a medidas de probabilidade, e a demonstração é muito similar.

Dados três eventos A, B, C, temos que

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

e

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

Em geral, dados $A_1, A_2, \ldots, A_n$, temos que

$$P(\cup A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n).$$

Exercícios

Nos seguintes exercícios, considere A, B, C eventos de um espaço amostral Ω.

1. Prove a fórmula de inclusão-exclusão.

2. Suponha que $P(A) = 1/3$, $P(B) = 1/4$, $P(A \cap B) = 1/10$. Expresse cada um dos seguintes eventos em linguagem de experimentos e determine sua probabilidade: $A \setminus B$, $A \cup B$, $A^C \cup B^C$, $A^C \cap B^C$, $A \cup B^C$.

3. Suponha que $P(A) = 0.3$, $P(B) = 0.2$, $P(C) = 0.4$, $P(A \cap B) = 0.04$, $P(A \cap C) = 0.1$, $P(B \cap C) = 0.1$, $P(A \cap B \cap C) = 0.01$. Expresse cada um dos seguintes eventos em notação de conjuntos e determine sua probabilidade:

   (a) pelo menos um dos três eventos ocorre;

(b) nenhum dos três eventos ocorre;

(c) exatamente um dos três eventos ocorre;

(d) exatamente dois dos três eventos ocorrem.

4.5 Algumas desigualdades

Para os seguintes resultados, suponha que $\{A_n : n \in I\}$ é uma coleção enumerável de eventos em Ω.

Desigualdade de Boole

$$P(\bigcup_{n\in I} A_n) \leq \sum_{n\in I} P(A_n).$$

Veja a prova A.1.

Desigualdade de Bonferroni

$$P(\bigcap_{n\in I} A_n) \geq 1 - \sum_{n\in I}(1 - P(A_n)).$$

A prova é feita aplicando a desigualdade de Boole à coleção $\{A_n^C : n \in I\}$.

Exercícios

1. Suponha que $\{A_n : n \in I\}$ é uma coleção enumerável de eventos com $P(A_n) = 0$, para $n \in I$. Use a desigualdade de Boole para mostrar que $P(\cup_n A_n) = 0$.

   Um evento A com $P(A) = 0$ é dito um evento nulo. Desta forma, a união enumerável de eventos nulos é um evento nulo.

2. Suponha que $\{A_n : n \in I\}$ é uma coleção enumerável de eventos com $P(A_n) = 1$, para todo $n \in I$. Use a desigualdade de Bonferroni para mostrar que $P(\cap_n A_n) = 0$.

   Um evento A com $P(A) = 1$ é dito um evento quase certo. Desta forma, a interseção enumerável de eventos quase certos é um evento quase certo.

4.6 Distribuição de uma variável aleatória

Seja $(\Omega, \mathscr{F}, P)$ um espaço de probabilidade, e seja X uma variável aleatória (real) definida em Ω.

A estrutura de probabilidade definida em Ω, por $\mathscr{F}$ e P, induz uma estrutura de probabilidade na imagem da v.a. X, que denotaremos por P_X.

Esta probabilidade, chamada a função de probabilidade induzida pela v.a. X, é definida para todo evento B real como

$$P_X(B) = P(X \in B) = P(\{\omega \in \Omega : X(\omega) \in B\})$$

ou seja, é $P_X(B)$ é a probabilidade da imagem inversa de B.

Desta forma, observe que uma variável aleatória X induz um novo espaço de probabilidade em $\mathbb{R}$, $(\mathbb{R}, \mathscr{B}, P_X)$. Usualmente, chamamos P_X de distribuição ou lei de probabilidade de X.

Para uma variável aleatória X discreta, a lei da probabilidade total pode ser bastante útil, já que X define uma partição natural em Ω com os eventos da forma $(X = k)$,

$$\Omega = \bigcup_{k \in \Omega_X} (X = k)\,.$$

Neste caso, para qualquer evento A de Ω, podemos escrever a união disjunta

$$A = \bigcup_{k \in \Omega_X} (A \cap (X = k))\,,$$

e, portanto,

$$P(A) = \sum_{k \in \Omega_X} P(A \cap (X = k))\,.$$

Exemplo. Considere novamente o exemplo da página 19. Para a variável N, resultado do lançamento do dado, considere a partição definida pelos eventos $(N = n)$,

$$(N = n) = \{\omega \in \Omega : N(\omega) = n\}\,.$$

Assim, por exemplo,

$$(N = 2) = \{\omega \in \Omega : N(\omega) = 2\} = \{00, 01, 10, 11\}\,.$$

Seja A o evento "obter uma única cara" ao realizar o experimento. Pelo anterior, podemos escrever A como a união disjunta

$$A = \bigcup_{n=1}^{6} (A \cap (N = n))\ ,$$

e desta forma podemos determinar a probabilidade de A pela soma

$$P(A) = \sum_{n=1}^{6} P\left(A \cap (N = n)\right) .$$

O raciocínio implícito nesta igualdade é que podemos obter $P(A)$ considerando os elementos de A para cada valor de N, separadamente.

4.7 Exemplos

Moedas

Consideremos o experimento do lançamento de uma moeda n vezes, observando a seqüência de resultados obtidos $X = (X_1, \ldots X_n)$, onde 1 denota cara e 0 denota coroa.

Observemos que o espaço amostral do experimento é $\Omega_X = \{0, 1\}^n$. Se supusermos que a probabilidade de obter cara em cada lançamento é a mesma de obter coroa, então cada resultado elementar tem a mesma probabilidade de ocorrer, ou seja, X tem distribuição uniforme em Ω_X. Como temos 2^n resultados, cada um tem probabilidade $1/2^n = (1/2)^n$.

X = (1 , 1 , 1 , 0 , 1 , 0)

Figura 4.2: Resultado X do experimento "resultados em $n = 6$ lançamentos de uma moeda".

Definamos a variável aleatória Y como o total de caras obtidas em n lançamentos de uma moeda.

O evento $(Y = k)$ consiste em todos os valores de X com exatamente k caras. Pelo já visto, temos um total de $\binom{n}{k}$ possibilidades de ordenar as k caras em n lançamentos.

Portanto,

$$P(Y = k) = \binom{n}{k} \left(\frac{1}{2}\right)^n ,$$

para todo $k \in \{0, 1, \ldots, n\}$.

O vídeo Noite de forró mostra uma aplicação destas distribuições.

Exercícios

1. Considere o experimento de lançar uma moeda balanceada 3 vezes. Seja A o evento "o primeiro lançamento é cara" e B, o evento "exatamente dois lançamentos

resultam em cara". Para cada um dos eventos seguintes, liste seus elementos e determine sua probabilidade: A, B, $A \cap B$, $A \cup B$, $A^C \cup B^C$, $A^C \cap B^C$, $A \cup B^C$.

2. Considere o experimento de lançar uma moeda balanceada 4 vezes, e denote por Y o total de caras observadas. Liste os elementos do evento $(Y = k)$, para cada k possível, e determine a probabilidade do evento.

3. No experimento Coin, selecione $n = 2$ moedas e rode o experimento 50 vezes, atualizando a tabela depois de cada rodada. Diretamente dos resultados, determine a freqüência dos eventos $A =$"o primeiro lançamento é cara" e $B =$"os dois lançamentos são cara", $A \cap B$, $A \cup B$. Relaciones estes valores com as respectivas probabilidades e com as relações vistas anteriormentes.

Dados

Considere o experimento de lançar n vezes um dado de k faces, com faces numeradas de 1 a k, registrando a seqüência de resultados $X = (X_1, X_2, \ldots, X_n)$. O caso $k = 6$ corresponde ao dado comum.

Figura 4.3: Resultado X do experimento "resultados em $n = 6$ lançamentos de um dado".

Se supusermos que cada face tem a mesma probabilidade de ser observada em cada lançamento, então todos os k^n valores possíveis de X têm a mesma probabilidade, $1/k^n$.

Exercícios

1. No experimento Dice, selecione $n = 2$ dados e rode o experimento 50 vezes, atualizando a tabela depois de cada rodada. Determine a freqüência dos eventos $A =$"o primeiro lançamento é menor que 3" e $B =$"a soma dos dois lançamentos é 6", $A \cap B$, $A \cup B$. Relaciones estes valores com as respectivas probabilidades e com as relações vistas anteriormentes.

Distribuição do máximo e do mínimo de variáveis uniformes

Considere o experimento de lançar n vezes um dado de k faces igualmente prováveis, e definamos as variáveis aleatórias U igual ao mínimo valor obtido nos n lançamentos e V igual ao máximo valor.

Claramente, U e V podem assumir qualquer valor entre 1 e k. Obteremos a distribuição de U para $n = 2$ e $k = 6$. As provas do caso geral e da distribuição de V são análogas.

Assim, U pode assumir os valores de 1 a 6. Observemos que $(U = 6)$ ocorre somente se em ambos os lançamentos for obtido 6. Como temos um total de $6^2 = 36$ possibilidades para os resultados dos dois lançamentos, então $P(U = 6) = 1/36$. A Tabela 4.1 mostra todos os possíveis resultados dos dois lançamentos e o valor de U em cada caso.

(D_1, D_2)	1	2	3	4	5	6
1	1	1	1	1	1	1
2	1	2	2	2	2	2
3	1	2	3	3	3	3
4	1	2	3	4	4	4
5	1	2	3	4	5	5
6	1	2	3	4	5	6

Tabela 4.1: Possíveis resultados do mínimo obtido em dois lançamentos de um dado.

Sendo assim, para determinar a probabilidade do evento $(U = k)$ basta contar o total de resultados do experimento cujo mínimo é igual a k.

O software Explorando o Jogo do Máximo trabalha com a simulação de V para dois dados.

Capítulo 5

Probabilidade condicional

5.1 Definição

Como antes, consideremos o esquema básico de um experimento aleatório, um espaço amostral Ω, um conjunto de eventos $\mathscr{F}$ e uma medida de probabilidade P.

Suponhamos que um evento B tenha ocorrido. Eventualmente, esta informação pode alterar a probabilidade atribuída a outros eventos. De fato, tendo esta informação sobre B, um outro evento A poderá ter ocorrido se e somente se $A \cap B$ puder ter ocorrido (ou seja, se for diferente de vazio). Daqui, a probabilidade de A, supondo que B ocorreu, deve ser proporcional a $P(A \cap B)$.

Em particular, $P(\Omega)$ deve ser proporcional a $P(\Omega \cap B) = P(B)$.

Definição 1 *Seja B um evento com $P(B) > 0$. Definimos a* probabilidade condicional dado B *como a lei de probabilidade $P(\cdot \mid B) : \mathscr{F} \to \mathbb{R}$ que a cada evento $A \in \mathscr{F}$ atribui o valor $P(A \mid B)$ igual a*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Intuitivamente, podemos interpretar $P(A \mid B)$ da seguinte maneira: sabendo ou supondo que B ocorreu, qual é a "nova" probabilidade de que A ocorra?

Exemplo. Considere o experimento de observar os resultados de dois lançamentos de um dado, e os eventos A: "o primeiro resultado é par", e B: "a soma dos resultados é 6". Supondo que os resultados são equiprováveis, sabemos que $P(A) = 1/2$ e que $P(B) = 5/36$. Agora, suponhamos que B ocorreu; isto significa que ocorreu um dos

resultados: $(1,5),(2,4),(3,3),(4,2),(5,1)$, que são equiprováveis entre si. Portanto, com esta informação, a probabilidade de que A tenha ocorrido é $2/5$. De fato, pela definição anterior, temos

$$P(A \mid B) = \frac{P(\{(2,4),(4,2)\})}{P(\{(1,5),(2,4),(3,3),(4,2),(5,1)\})} = \frac{2/36}{5/36} = \frac{2}{5}\,.$$

Se supusermos que A ocorreu, então isto quer dizer que ocorreu um dos 18 resultados: $(2,1),(2,2),\ldots,(2,6),(4,1),\ldots,(4,6),(6,1),\ldots,(6,6)$. Com esta informação, o evento B só terá ocorrido se tiverem ocorrido $(2,4)$ ou $(4,2)$. Assim, a probabilidade condicional de B dado A deveria ser $2/18 = 1/9$. De fato, pela definição,

$$P(B \mid A) = \frac{2/36}{18/36} = \frac{2}{18}\,.$$

Exemplo. Uma caixa contém 25 lâmpadas, 5 das quais estão em boas condições e durarão pelo menos 30 dias, 10 estão parcialmente defeituosas e falharão no segundo dia e 10 estão totalmente defeituosas e não acenderão. Escolhendo uma lâmpada da caixa que inicialmente acende, qual é a probabilidade de que ela ainda funcione após uma semana de uso? Definamos os eventos A:"a lâmpada escolhida está em boas condições", e B: "a lâmpada escolhida está parcialmente defeituosa". O problema diz que ocorreu o evento $A \cup B$, cuja probabilidade inicialmente era $15/25$. Tendo essa informação, a probabilidade de que A tenha ocorrido é, pela definição,

$$P(A \mid A \cup B) = \frac{5/25}{15/25} = \frac{5}{15}\,.$$

Outra forma de visualizar este resultado é mediante uma árvore de probabilidade, como na Figura 5.1. A diferença de uma árvore de contagem, nesta colocamos os resultados possíveis nos galhos, com suas respectivas probabilidades.

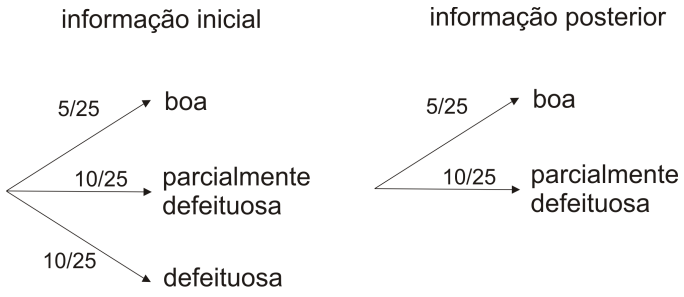

Figura 5.1: Árvore de probabilidade para o exemplo das lâmpadas.

Com a informação de que a lâmpada acendeu, excluímos uma das possibilidades. O que a função probabilidade condicional faz é reescalar as probabilidades restantes, para que sua soma seja um, depois da nova informação, mantendo a proporcionalidade entre si.

Exemplo. Considere o experimento de observar o resultado de dois lançamentos de uma moeda. Supondo que o espaço amostral é equiprovável, determine a probabilidade condicional de obter cara em ambos os lançamentos, dado que: (a) foi obtido cara no primeiro lançamento; (b) foi obtido cara em pelo menos um dos lançamentos.

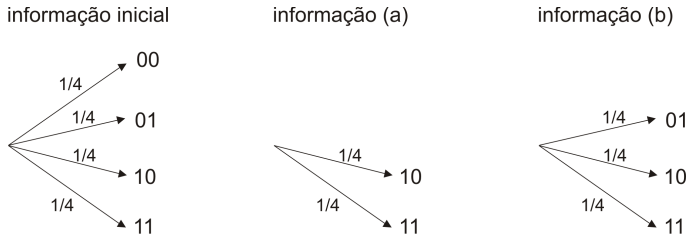

Figura 5.2: Árvore de probabilidade para o exemplo dos dois lançamentos de uma moeda, com 0 indicando coroa e 1, cara.

Resolveremos este problema usando as árvores de probabilidade da Figura 5.2. Fica para o leitor obter a solução analítica. Da figura, a solução é quase imediata: para o item (a), a probabilidade condicional de obter cara em ambos os lançamentos é $1/2$, enquanto que para o item (b) é $1/3$.

Note que a função $P(\cdot \mid B)$ é uma medida de probabilidade e tem, portanto, todas as propriedades vistas no capítulo anterior.

Os experimentos Jogo da trilha e Jogo das amebas mostram uma aplicação de probabilidade condicional.

Exercícios

Prove as seguintes afirmações, onde A, B são eventos com $P(B) > 0$.

1. A função $P(\cdot \mid B)$ é uma medida de probabilidade em $\mathscr{F}$.
2. Se $B \subset A$ então $P(A \mid B) = 1$.
3. Se $A \subset B$ então $P(A \mid B) = P(A)/P(B)$.
4. Se A e B forem disjuntos então $P(A \mid B) = 0$.
5. Suponha que A também tem probabilidade positiva. Prove as seguintes afirmações.
   (a) $P(A \mid B) > P(A)$ se e só se $P(B \mid A) > P(B)$ se e só se $P(A \cap B) > P(A)P(B)$. Neste caso, dizemos que A e B são eventos positivamente correlacionados.

(b) $P(A \mid B) < P(A)$ se e só se $P(B \mid A) < P(B)$ se e só se $P(A \cap B) < P(A)P(B)$. Neste caso, dizemos que A e B são eventos negativamente correlacionados.

(c) $P(A \mid B) = P(A)$ se e só se $P(B \mid A) = P(B)$ se e só se $P(A \cap B) = P(A)P(B)$. Neste caso, dizemos que A e B são eventos não correlacionados ou independentes: intuitivamente, a ocorrência de um dos eventos não altera a probabilidade do outro evento.

6. A e B têm a mesma correlação que A^C e B^C.

5.2 Algumas propriedades

Regra do produto

Em alguns problemas, é possível quantificar probabilidades condicionais de maneira simples e usá-las para determinar a probabilidade de outros eventos.

Observe que da definição de probabilidade condicional, dados os eventos A e B, podemos escrever

$$P(A \cap B) = P(B)P(A \mid B) = P(A)P(B \mid A)\,, \tag{5.1}$$

se $P(B) \neq 0 \neq P(A)$. Observe também que a igualdade permanece válida se $P(A)$ ou $P(B)$ for zero, se a probabilidade condicional neste caso fosse qualquer valor real arbitrário.

Para dois eventos quaisquer, A e B, a igualdade (5.1) é chamada **regra do produto**.

Em palavras, a probabilidade de que dois eventos ocorram é igual à probabilidade de um deles ocorrer vezes a probabilidade do outro ocorrer, condicional na ocorrência do primeiro.

Esta regra permite determinar de maneira natural a probabilidade da interseção de mais de dois eventos. Dados os eventos $A_1, A_2, \ldots, A_n$,

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2 \mid A_1) \ldots P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1})\,,$$

com a respectiva interpretação das probabilidades condicionais envolvidas.

A igualdade anterior é particularmente útil para experimentos que consistem de etapas dependentes, com A_i um evento relacionado à etapa i.

Exemplo. Uma urna contém sete bolas pretas e cinco bolas brancas. Considere o experimento de extrair duas bolas da urna, sem reposição. Supondo que todas as bolas têm mesma chance de serem extraídas, determine a probabilidade de que ambas sejam pretas.

Denotemos por P_n o evento de obter uma bola preta na n-ésima extração, $n \in \{1, 2\}$. Então

$$P(P_1 \cap P_2) = P(P_1)P(P_2 \mid P_1) = \frac{7}{12}\,\frac{6}{11}\,.$$

Em palavras, a probabilidade de obter bola preta na primeira e na segunda extração é igual à probabilidade de obter bola preta na primeira extração vezes a probabilidade de obter bola preta na segunda extração, sabendo que uma bola preta foi extraída na primeira extração.

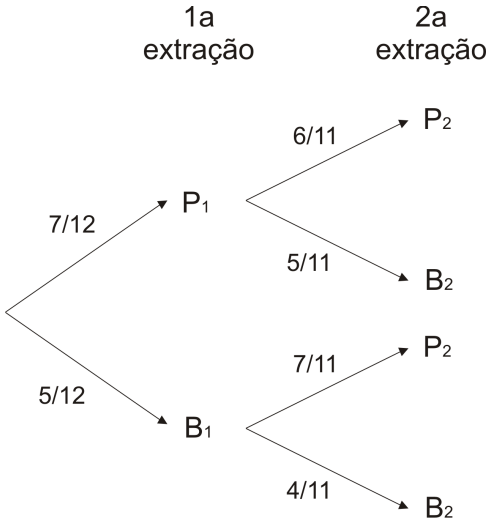

Figura 5.3: Árvore de probabilidade para a regra do produto, onde P indica bola preta e B, bola branca, em cada extração.

Usando a representação em árvore de probabilidade, como na Figura 5.3, utilizamos duas sequências de galhos, correspondentes às duas etapas do experimento: primeira e segunda extrações. A probabilidade de qualquer sequência de galhos (da esquerda para a direita) é o produto das probabilidades de cada galho.

Lei da probabilidade total

Com a regra do produto, podemos reescrever a lei da probabilidade total como

$$P(B) = \sum_{k \in I} P(A_k)P(B \mid A_k),$$

onde $\{A_k : k \in I\}$ é uma partição finita ou enumerável de eventos de Ω.

Este resultado é útil quando conhecemos as probabilidades dos eventos da partição, $P(A_k)$, e as probabilidades condicionais, $P(B \mid A_k)$, e com isso podemos determinar $P(B)$ por partes.

Podemos representar probabilidades condicionais e a lei da probabilidade total por árvores de probabilidades, como na Figura 5.4. Para simplicidade, consideremos uma partição com 3 elementos, A_1, A_2, A_3.

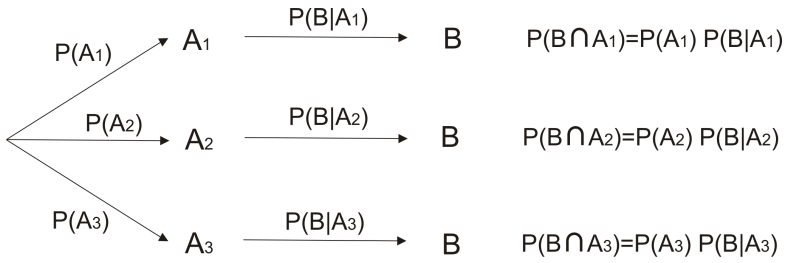

Figura 5.4: Árvore de probabilidade para a lei da probabilidade total.

Os primeiros galhos (mais à esquerda) representam as probabilidades iniciais de cada evento na partição. Os galhos seguintes representam as probabilidades condicionais sobre os galhos anteriores. Para cada sequência de galhos (da esquerda para a direita), o produto das probabilidades é a probabilidade da interseção dos eventos considerados.

Exemplo. Considere as urnas U_1, U_2, U_3, nas quais a proporção de bolas brancas é, respectivamente, p_1, p_2, p_3. Considere o experimento de extrair uma bola de uma das urnas, e seja B, o evento de obter uma bola branca.

Com a informação anterior, o que temos são as probabilidades condicionais de obter uma bola branca, para cada urna:

$$P(B \mid U_n) = p_n\,, \text{ para cada } n \in \{1, 2, 3\}\,.$$

Suponha que neste experimento, a urna U_n será sorteada com probabilidade π_n, para cada n,

$$P(U_n) = \pi_n\,.$$

Assim, a probabilidade de obter bola branca ao realizar o experimento é

$$P(B) = \sum_{n=1}^{3} P(U_n)\, P(B \mid U_n) = \sum_{n=1}^{3} \pi_n\, p_n\,.$$

Observe que esta igualdade representa uma ponderação das proporções p_n, com respeito às respectivas probabilidades π_n.

Exercícios

1. Suponha que A, B são eventos com $P(A) = 1/3$, $P(B) = 1/4$, $P(A \cap B) = 1/10$. Determine: $P(A \mid B)$, $P(B \mid A)$, $P(A^C \mid B)$, $P(B^C \mid A)$, $P(A^C \mid B^C)$.

2. Suponha que A, B, C são eventos com $P(A \mid C) = 1/2$, $P(B \mid C) = 1/3$, $P(A \cap B \mid C) = 1/4$. Determine: $P(A \setminus B \mid C)$, $P(A \cup B \mid C)$, $P(A^C \cap B^C \mid C)$.

3. Suponha que A, B são eventos com $P(A) = 1/2$, $P(B) = 1/3$, $P(A \cap B) = 3/4$. Determine: $P(A \cap B)$, $P(A \cup B)$, $P(B \mid A)$, $P(A^C \cup B)$; A e B são positiva, negativamente correlacionados ou não correlacionados?

4. Uma empresa tem 200 funcionários: 120 mulheres e 80 homens. Das 120 funcionárias, 30 são gerentes, enquanto que 20 dos 80 funcionários são gerentes. Selecionando um funcionário, determine a probabilidade de que:

   (a) seja mulher;

   (b) seja gerente;

   (c) seja gerente, dado que é mulher;

   (d) seja mulher, dado que é gerente.

   As características mulher e gerente são correlacionadas? como?

5. Considere o experimento de lançar 2 dados e observar o resultado obtido $X = (X_1, X_2)$ em cada dado. Assuma que os dados são equilibrados e que os lançamentos não favorecem nenhuma face. Defina Y como a soma dos resultados. Para cada par de eventos a seguir, determine a probabilidade de cada evento, a probabilidade condicional de um evento dado o outro, e que tipo de correlação eles apresentam.

   (a) $\{X_1 = 3\}$, $\{Y = 5\}$;

   (b) $\{X_1 = 3\}$, $\{Y = 7\}$;

   (c) $\{X_1 = 2\}$, $\{Y = 5\}$;

   (d) $\{X_1 = 3\}$, $\{X_1 = 2\}$.

6. Simule o exercício anterior no applet Dice, selecionando $n = 2$.

7. Considere novamente o exercício anterior, e defina U como o resultado mínimo e V como o resultado máximo. Determine:

   (a) $P(U = u \mid V = 4)$, para os valores possíveis de u;

(b) $P(Y = y \mid V = 4)$, para os valores possíveis de y;

(c) $P(V = v \mid Y = 8)$, para os valores possíveis de v;

(d) $P(U = u \mid Y = 8)$, para os valores possíveis de u;

(e) $P(X_1 = x_1, X_2 = x_2 \mid Y = 8)$, para os valores possíveis de (x_1, x_2).

8. Um baralho comum de 52 cartas é dividido em 4 pilhas de 13 cartas. Determine a probabilidade de que cada pilha contenha exatamente um ás, supondo que todos os empilhamentos possíveis são equiprováveis.

5.3 Regra de Bayes

Seja $\{A_i : i \in I\}$ uma partição finita ou enumerável de eventos de Ω e seja B um evento. Da regra do produto, dado $j \in I$, podemos obter $P(A_j \cap B)$ mediante a igualdade

$$P(A_j \cap B) = P(A_j)P(B \mid A_j)\,,$$

que indica a probabilidade de ambos, A_j e B, ocorrerem, conhecendo as probabilidades dos eventos da partição, $P(A_j)$, e as probabilidades condicionais de B, $P(B \mid A_j)$.

Suponha que você recebe a informação de que B ocorreu. A pergunta natural é qual dos eventos da partição ocorreu. A lei da probabilidade total nos permite determinar as probabilidades condicionais destes eventos.

Da definição de probabilidade condicional, temos, para cada j,

$$P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j)P(B \mid A_j)}{P(B)}\,,$$

onde $P(B)$ é a probabilidade de B ocorrer antes de realizar o experimento. Se não conhecermos esta probabilidade, podemos usar a lei da probabilidade total no denominador para obter

$$P(A_j \mid B) = \frac{P(A_j)P(B \mid A_j)}{\sum_{i \in I} P(A_i)P(B \mid A_i)}\,.$$

Esta igualdade é conhecida como a regra de Bayes.

Exemplo. Continuando com o exemplo das três urnas da seção anterior, suponha que a probabilidade de cada urna ser escolhida para a extração é 1/3. Suponha também que as proporções de bolas brancas são $p_1 = 0.1$, $p_2 = 0.5$, $p_3 = 0.9$.

A probabilidade de extrair uma bola branca é

$$P(B) = \frac{1}{3}0.1 + \frac{1}{3}0.5 + \frac{1}{3}0.9 = 0.5\,.$$

Você recebe a informação de que ao realizar o experimento, foi observada bola branca. Com isto, a probabilidade condicional de cada urna é

$$P(U_1) = \frac{P(U_1)P(B \mid U_1)}{P(B)} = \frac{\frac{1}{3}0.1}{\frac{1}{3}0.1 + \frac{1}{3}0.5 + \frac{1}{3}0.9} = 1/15\,,$$

$$P(U_2) = \frac{P(U_2)P(B \mid U_2)}{P(B)} = \frac{\frac{1}{3}0.5}{\frac{1}{3}0.1 + \frac{1}{3}0.5 + \frac{1}{3}0.9} = 5/15\,,$$

$$P(U_3) = \frac{P(U_3)P(B \mid U_3)}{P(B)} = \frac{\frac{1}{3}0.9}{\frac{1}{3}0.1 + \frac{1}{3}0.5 + \frac{1}{3}0.9} = 9/15\,.$$

Perceba que, depois de realizar o experimento e observar bola branca, a probabilidade das urnas muda: aquela que tinha maior proporção de bolas brancas passa a ser a mais provável.

Intuitivamente, a regra de Bayes nos permite atualizar as probabilidades dos eventos U_i, após saber ou supor que B ocorreu. É comumente utilizada para atualizar a probabilidade dos diversos modelos probabilísticos (no exemplo, as urnas) considerados para uma população após obter informação de uma amostra da mesma (no exemplo, uma amostra de tamanho 1 de uma população de bolas brancas e pretas).

Os vídeos Teste de gravidez e Crime da rua do Gasômetro apresentam duas situações em que a regra de Bayes pode ser aplicada.

Razão de chances

Considere um evento E, com $0 < P(E) < 1$. Definimos a **razão de chances** do evento E (ou a favor do evento E) como

$$\frac{P(E)}{P(E^C)} = \frac{P(E)}{1 - P(E)}\,.$$

Por exemplo, se $P(E) = 2/3$, então $P(E^C) = 1/3$ e a razão de chances de E é igual a 2. Em linguagem mais usual dizemos que a razão de chances a favor de E é de 2:1 (lê-se: de 2 para 1).

No contexto de um modelo probabilístico, H, a ser testado e uma evidência E observada, a razão de chances a favor do modelo H após observar E é, pela regra de Bayes,

$$\frac{P(H \mid E)}{P(H^C \mid E)} = \frac{P(H)}{P(H^C)}\,\frac{P(E \mid H)}{P(E \mid H^C)}\,.$$

Aqui, $P(H)/P(H^C)$ é a razão de chances a favor de H, antes de observar a evidência E. A razão $P(E \mid H)/P(E \mid H^C)$ é chamada razão de verossimilhanças a favor de H, a partir da evidência E.

Exemplo. Suponha que quando a moeda A é lançada, a probabilidade de obter cara é $1/4$, enquanto que para a moeda B, esta probabilidade é $3/4$. Em um experimento, uma destas moedas é lançada duas vezes. Se ambas as faces obtidas resultarem ser cara, determine a probabilidade de que a moeda lançada tenha sido a B.

Podemos resolver este problema, obtendo a razão de chances a favor da moeda B. Denotemos por E (a evidência) o evento de obter duas caras. Como não temos informação nenhuma sobre como foi feita a escolha da moeda, podemos supor que ambas são igualmente prováveis; em outras palavras, a razão de chances a favor de B, antes de realizar o experimento, é igual a 1 (de 1:1).

Assim,

$$\frac{P(B \mid E)}{P(A \mid E)} = \frac{P(B)}{P(A)} \frac{P(E \mid B)}{P(E \mid A)} = 1 \frac{3/4 \times 3/4}{1/4 \times 1/4} = \frac{9}{1}\,.$$

Portanto, a razão de chances a favor de B é de 9:1, ou, equivalentemente, a probabilidade de ter sido B a moeda escolhida é $9/(9+1) = 0.9$. Note que, com apenas dois lançamentos obtendo duas caras, a probabilidade de B passou de 50% para 90%.

Exemplo. Considere três cartas idênticas na forma, mas tais que: a primeira tem dois lados verdes; a segunda, os dois lados brancos; a terceira, um lado verde e o outro branco. As três cartas são misturadas em um chapéu e uma delas é escolhida e colocada sobre uma mesa. Se o lado superior da carta escolhida for verde, qual é a probabilidade de que a outra face seja branca? Dica: use regra de Bayes ou árvore de probabilidade. O resultado obtido era o esperado?

Exercícios

1. Considere o experimento dado-moeda de lançar um dado e depois uma moeda o número de vezes que aparece no dado. Seja N o resultado do dado e C o evento de que todos os lançamentos da moeda resultam em cara. Supondo que os resultados são equiprováveis, determine: $P(C)$, $P(N = n \mid C)$ para $n \in \{1, 2, 3, 4, 5, 6\}$, compare estes resultados com $P(N = n)$ para $n \in \{1, 2, 3, 4, 5, 6\}$, dizendo como os eventos C e $(N = n)$ estão correlacionados.

2. Simule o exercício anterior no applet Die-Coin, comparando as freqüências observadas com as probabilidades calculadas no item anterior.

3. Uma bolsa contém 12 moedas indistinguíveis: 5 moedas equilibradas, 4 moedas viesadas com probabilidade de cara igual a $1/3$, e 3 moedas com duas caras. Uma

moeda é selecionada e lançada. Qual é a probabilidade de obter cara? Se o resultado fosse cara, qual é a probabilidade condicional de cada tipo de moeda?

4. Considere o experimento moeda-dado, no qual uma moeda é lançada. Se o resultado for coroa, lançamos um dado balanceado; se for cara, lançamos um dado ás-seis (as faces 1 e 6 têm probabilidade 1/4 cada uma, e as demais, 1/8 cada). Seja H o evento de obter cara, e seja Y o resultado obtido no dado. Supondo que a moeda é balanceada, determine:

   (a) $P(Y = y)$ para $y \in \{1, 2, 3, 4, 5, 6\}$;

   (b) $P(H \mid Y = y)$ para $y \in \{1, 2, 3, 4, 5, 6\}$;

   (c) compare as probabilidades do item anterior com $P(H)$, indicando o tipo de correlação entre os eventos H e $(Y = y)$.

   (d) Simule o exercício anterior no applet Coin-Die, comparando as freqüências observadas com as probabilidades calculadas no item anterior.

5. Uma fábrica tem 3 linhas de montagem para produzir chips de memória. A linha 1 produz 50% dos chips, com uma taxa de defeituosos de 4%; a linha 2 produz 30% dos chips, com uma taxa de defeituosos de 5%; a linha 3 produz 20% dos chips, com uma taxa de defeituosos de 1%. Ao selecionar um chip, determine a probabilidade de que o chip seja defeituoso. Determine a probabilidade condicional de cada linha se o chip for defeituoso.

6. Em uma população, composta igualmente por homens e mulheres, 10% dos homens são daltônicos, enquanto que apenas 1% das mulheres o são. Determine a proporção de daltônicos na população, e a proporção de daltônicos que são homens.

7. Uma urna contém 15 bolas brancas. Uma segunda urna contém 15 bolas pretas. Cinco bolas são extraídas de uma das urnas e colocadas na outra. Duas bolas são extraídas sem reposição de uma das duas urnas e resultam ser da mesma cor. Qual é a probabilidade de que a outra urna contenha 10 bolas?

Capítulo 6

Independência

6.1 De dois eventos

Como antes, consideremos um experimento aleatório com espaço amostral Ω, um conjunto de eventos $\mathscr{F}$ e uma medida de probabilidade P.

Independência é um dos conceitos importantes em teoria de probabilidade, e utilizado como suposição para uma ampla gama de modelos.

Definição 2 *Dados dois eventos A, B, dizemos que eles são* independentes *se*

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

Observe que se ambos os eventos tiverem probabilidade positiva, então independência é equivalente a

$$P(A \mid B) = P(A) \quad \text{e} \quad P(B \mid A) = P(B).$$

Desta forma, fica evidenciado que dois eventos são independentes se ao supor que um deles ocorre, a probabilidade do outro ocorrer não é alterada. Ou seja, a (suposição de) ocorrência de um dos eventos não entrega informação relevante (que altere a probabilidade) sobre o outro evento.

Dois eventos que não são independentes são ditos dependentes ou correlacionados.

Exemplo. Uma carta é selecionada de um baralho comum de 52 cartas. Considere os eventos E: "a carta selecionada é ás", F: "a carta selecionada é de ouros". Os eventos são independentes?

Exemplo. Uma moeda é lançada duas vezes. Suponha que os quatro resultados possíveis para a sequência são igualmente prováveis. Considere os eventos E: "o primeiro lançamento é cara", e F: "os dois lançamentos são cara".

Exemplo. Considere o experimento de dois lançamentos de um dado balanceado, e os eventos E: "a soma dos resultados é 6", F: "o primeiro resultado é igual a 4", G: "a soma dos resultados é 7". Os eventos E e F são independentes? e F e G?

Pergunta. Se o evento E for independente de F e também for independente de G, então podemos afirmar que E é independente de $F \cap G$?

Exemplo. Considere o experimento de dois lançamentos de um dado balanceado, e os eventos E: "o segundo resultado é igual a 3", F: "o primeiro resultado é igual a 4", G: "a soma dos resultados é 7".

Definição 3 *Três eventos A, B, C são ditos* independentes *se*

$$P(A\cap B) = P(A)P(B)\,,\ P(A\cap C) = P(A)P(C)\,,\ P(B\cap C) = P(B)P(C)\,,\ P(A\cap B\cap C) = P(A)P(B)P($$

Desta definição, podemos concluir que se A, B, C são independentes então A é independente de qualquer evento formado por B e C.

Exercícios

1. Suponha que A, B são eventos disjuntos, ambos com probabilidade positiva. Mostre que $P(A \cap B) = 0$, mas que $P(A)P(B) > 0$. Portanto, A e B não são independentes. Mais ainda, eles são negativamente correlacionados (já que se um dele ocorrer, o outro não pode ter ocorrido).

2. Suponha que A, B são eventos independentes. Mostre que também são independentes: A^C e B, B e A^C, A^C e B^C.

3. Se A, B, C forem independentes, mostre que A é independente de $B \cup C$.

4. Uma urna contém uma bola branca e uma bola preta. Um prêmio é oferecido a três jogadores, A, B, C, que extrairão uma bola da urna, repondo-a, um de cada vez. O primeiro jogador que extrair a bola branca leva o prêmio. Se nehum jogador a extrair, o prêmio será recolhido. Quais são as chances de cada jogador de ganhar o prêmio? E do prêmio ser recolhido?

5. Considere o experimento de extrair uma bola de uma urna com quatro bolas numeradas de 1 a 4, e os eventos $E = \{1, 2\}$, $F = \{1, 3\}$, $G = \{1, 4\}$. Supondo que os quatro resultados são igualmente prováveis, verifique se os eventos E, F, G são independentes.

6. Um sistema formado por n componentes é dito ser em paralelo se ele funcionar quando pelo menos uma das componentes funcionar. Para tal sistema, suponha que a i-ésima componente funciona com probabilidade p_i, independentemente das outras, para $i \in \{1, \ldots, n\}$. Determine a probabilidade de que o sistema funcione.

6.2 De uma coleção de eventos

A definição anterior pode ser estendida para mais de três eventos.

Consideremos uma coleção finita $\mathscr{C} = \{A_1, \ldots A_n\}$ de eventos. Dizemos que eles são independentes se, para qualquer subcoleção $\{A_{i_1}, A_{i_2}, \ldots, A_{i_k}\} \subset \mathscr{C}$, tivermos

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k}).$$

Finalmente, uma coleção infinita enumerável de eventos é independente se todas as subcoleções finitas forem independentes.

Esta definição de independência é muito mais geral do que uma simples independência dois a dois; todas as coleções finitas de A_i's devem ser independentes, dois a dois, três a três, etc. Os exercícios dão exemplos desta diferença.

Lembremos que os experimentos de uma sequência de experimentos idênticos, ou seja, com mesmo espaço amostral e mesma função de probabilidade, podem ser chamados de ensaios. A definição anterior nos permite falar de uma sequência (finita ou infinita enumerável) de ensaios independentes: infinitos lançamentos independentes de uma moeda; infinitas extrações de bolas de uma urna, com reposição etc. Neste caso, estes experimentos são chamados ensaios.

Exercícios

1. Descreva todas as condições para que A, B, C sejam eventos independentes.

2. Suponha que A, B, C são eventos independentes. Mostre que os eventos $A \cup B^C$ e C também são independentes.

3. Suponha que $\{A_1, A_2, \ldots, A_n\}$ é uma coleção finita de eventos independentes. Mostre que
$$P(\bigcup_{i=1}^{n} A_i) = 1 - \prod_{i=1}^{n}(1 - p(A_i)).$$

4. Considere uma sequência infinita de ensaios independentes, de modo que cada ensaio resulta em sucesso com probabilidade p e fracasso, com probabilidade $1-p$. Qual é a probabilidade de que ocorra pelo menos um sucesso nos n primeiros ensaios? Determine a probabilidade de que ocorram exatamente k sucessos nos n primeiros ensaios, para $k \in \{0\, dots, n\}$.

5. Suponha que A, B, C são eventos independentes com $P(A) = 0.3$, $P(B) = 0.5$, $P(C) = 0.8$. Expresse cada um dos seguintes eventos em notação de conjuntos e determine sua probabilidade:

   (a) todos os três eventos ocorrem;

   (b) nenhum dos três eventos ocorre;

   (c) ao menos um dos três eventos ocorre;

   (d) exatamente um dos três eventos ocorre;

   (e) exatamente dois dos três eventos ocorrem.

6. Suponha que A, B, C são eventos independentes com $P(A) = 1/2$, $P(B) = 1/3$, $P(C) = 1/4$. Determine a probabilidade dos seguintes eventos: $(A \cap B) \cup C$, $A \cup B^C \cup C$, $(A^C \cap B^C) \cup C^C$.

6.3 Independência condicional

Definição 4 *Dizemos que os eventos A_1 e A_2 são condicionalmente independentes dado B se*
$$P(A_1 \mid A_2, B) = P(A_1 \mid B)$$
ou
$$P(A_1 \cap A_2 \mid B) = P(A_1 \mid B)\, P(A_2 \mid B)\,.$$

Intuitivamente, a definição anterior nos diz que, supondo que B ocorra, a probabilidade condicional de que A_1 ocorra permanece inalterada com a informação da ocorrência ou não de A_2.

Esta definição se estende naturalmente para mais de dois eventos.

Exemplo. Uma moeda é lançada duas vezes. Suponha que os quatro resultados possíveis para a sequência são igualmente prováveis. Considere os eventos A: "o primeiro lançamento é cara", B: "o segundo lançamento é cara", e C: "os dois lançamentos apresentam a mesma face". Determine se A e B são independentes, e se são condicionalmente independentes dado C.

Exemplo. Considere uma urna com 3 bolas brancas e 1 preta, e o experimento: lançar uma moeda balanceada e extrair 2 bolas da urna: com reposição, se sair cara, e sem reposição se sair coroa. Considere os eventos F: "obter cara", E_1: "a primeira extração é bola preta", E_2: "a segunda extração é bola preta". Determine se E_1 e E_2 são independentes, e se são condicionalmente independentes dado F.

6.4 De variáveis aleatórias

Suponha que X_1 e X_2 são variáveis aleatórias. Intuitivamente, duas variáveis aleatórias são independentes se o conhecimento do valor de uma delas não altera a distribuição de probabilidade da outra variável.

Formalmente, X_1 e X_2 são variáveis aleatórias independentes se as coleções de eventos

$$\{(X_1 \in B) : B \in \mathscr{B}\} \text{ e } \{(X_2 \in B) : B \in \mathscr{B}\}$$

forem independentes, ou equivalentemente, se para cada escolha $B_1, B_2 \in \mathscr{B}$, tivermos que

$$P(X_1 \in B_1, X_2 \in B_2) = P(X_1 \in B_1)P(X_2 \in B_2).$$

Generalizando a definição anterior, dizemos que uma seqüência de variáveis aleatórias $X_1, X_2, \dots$ é independente se qualquer subcoleção finita delas for independente.

Suponhamos que temos um experimento básico no qual observamos a variável X_0. Por definição, o resultado do experimento que consiste em repetições independentes do experimento básico é uma seqüência de variáveis aleatórias independentes $X = (X_1, X_2, \dots)$, cada uma com a mesma distribuição de probabilidade que X_0.

6.5 Ensaios de Bernoulli

Um seqüência de ensaios de Bernoulli é uma seqüência $X = (X_1, X_2, \dots)$ de variáveis independentes identicamente distribuídas, onde cada variável pode assumir apenas os

valores 0 ou 1. Da terminologia de teoria da confiabilidade, usualmente chamamos o resultado 1 de sucesso e o 0 de fracasso.

Um exemplo usual é o de sucessivos lançamentos de uma moeda não necessariamente balanceada, ou de repetições de um experimento básico no qual temos interesse em saber se um evento A ocorre ou não.

Este processo tem um único parâmetro $p = P(X_i = 1)$ que determina completamente o modelo probabilístico.

Para este modelo, temos

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = p^{x_1+x_2+\cdots+x_n}(1-p)^{n-(x_1+x_2+\cdots+x_n)},$$

para $x_i \in \{0, 1\}$, $i \in \{1, 2, \ldots, n\}$.

Observemos que esta seqüência de variáveis é permutável, ou seja, se permutarmos a seqüência $(x_1, x_2, \ldots, x_n)$, a probabilidade não muda.

Exercícios

1. Seja $X = (X_1, X_2, \ldots)$ uma seqüência de ensaios de Bernoulli. Mostre que

   $$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = p^{x_1+x_2+\cdots+x_n}(1-p)^{n-(x_1+x_2+\cdots+x_n)},$$

   para $x_i \in \{0, 1\}$, $i \in \{1, 2, \ldots, n\}$.

2. Seja Y o total de sucessos nas n primeiras tentativas. Mostre que

   $$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

   para $k \in \{0, 1, \ldots, n\}$. A distribuição de Y é chamada distribuição binomial com parâmetros n e p.

3. Considere o experimento de lançar 2 dados balanceados de seis faces e observar a seqüência obtida. Seja A o evento de obter 3 no primeiro dado, B o evento de obter 4 no segundo dado e C o de que a soma seja 7.

   (a) Mostre que os eventos A, B, C são independentes dois a dois.

   (b) Mostre que $A \cap B$ implica C e que portanto eles são dependentes.

4. No applet Dice Sample, selecione $n = 2$ e rode o experimento 500 vezes. Para cada par de eventos no exercício anterior, determine o produto das freqüências observadas e a freqüência observada da interseção. Compare os resultados com o valor teórico.

5. Considere o experimento de lançar um dado balanceado e observar a face obtida, e os eventos $A = \{1, 2, 3, 4\}$ e $B = C = \{4, 5, 6\}$. Mostre que $P(A \cap B \cap C) = P(A)P(B)P(C)$, mas que B e C são dependentes.

6. Um dado balanceado é lançado 4 vezes. Determine a probabilidade de que: 6 não ocorra; 6 ocorra pelo menos uma vez; a soma dos dois primeiros resultados seja 5 e a soma dos dois últimos resultados seja 7.

7. Uma moeda com probabilidade de cara igual a $1/3$ é lançada 5 vezes. Seja X o resultado dos lançamentos (em 0's e 1's) e Y o total de caras. Determine:

   (a) $P(X = x)$ para cada $x \in \{0, 1\}^5$;

   (b) $P(Y = k)$ para cada $k \in \{0, 1, 2, 3, 4, 5\}$;

   (c) $P(1 \leq Y \leq 3)$.

Apêndice A

Demonstrações

A.1 Desigualdade de Boole

Suponha que $\{A_i : i \in I\}$ é uma coleção enumerável de eventos em Ω.

Consideremos inicialmente o caso em que I é um conjunto finito, ou seja, podemos considerar $I = \{1, 2, \ldots n\}$.

Definamos os eventos $B_1 = A_1$ e $B_i = A_i \setminus (A_1 \cup \cdots \cup A_{i-1})$ para $i \in \{2, \ldots, n\}$, como feito no Tema 1.

Portanto, os conjuntos $B_1, B_2, \ldots, B_n$ são disjuntos dois a dois e tem a mesma união que A. Desta forma $P(\cup A_i) = P(\cup B_i)$.

Pelo axioma de aditividade, $P(\cup B_i) = \sum P B_i$.

Finalmente, como $B_i \subset A_i$ e P é uma função crescente, temos que

$$P\left(\cup_i A_i\right) = P\left(\cup_i B_i\right) = \sum_i P(B_i) \leq \sum_i P(A_i),$$

como queríamos provar.

O caso não finito é demonstrado usando o Princípio de Indução Finita e o argumento anterior para os conjuntos $\cup_{i=1}^{n} A_i$ e A_{n+1}.

Voltar 4.5.

Referências Bibliográficas

[1] Carvalho, P.C.P., de Carvalho, J.B.P., Fernandez, P., Morgado, A.C.O. (2004) *Análise Combinatória e Probabilidade.* Editora SBM.

[2] Feller, P. (1976) *Introdução à Teoria de Probabilidades e suas Aplicações.* Editora Edgard Blücher.

[3] Fernández, P. (2007) *Medida e Integração.* Coleção Projeto Euclides.

[4] Gnedenko, B. (2008) *A Teoria da Probabilidade.* Editora Ciência Moderna.

[5] Halmos, P. (2001) *Teoria Ingênua dos Conjuntos.* Editora Ciência Moderna.

[6] James, B. (1996) *Probabilidade: um curso em nível intermediário.* Coleção Projeto Euclides.

[7] Meyer, P. (2000) *Probabilidade: aplicações à estatística.* Editora LTC.

[8] *Revista do Professor de Matemática*. SBM.

[9] Ross, S. (2010) *Probabilidade: um curso moderno com aplicações.* Editora Bookman Co.

Páginas da internet

Em português

[10] ALEA, Acção Local de Estatística Aplicada, Portugal.

[11] Mais - Recursos Educacionais em Matemática, Brasil.

[12] Matemática Multimídia, Unicamp.

Em inglês

[13] História da Matemática, Universidade de St. Andrews, Escócia.

[14] Laboratório de Probabilidade e Estatística, Universidade do Alabama, EUA.