II Colóquio de Matemática do Centro Oeste
07-11/11/2011

Introdução à Teoria da Probabilidade

Ralph Costa Teixeira, Augusto César Morgado

Conteúdo

Prefácio

Em sua base, probabilidade é, das teorias “fáceis”, a mais difícil que há.

Você (é, **você** mesmo**!**) usa probabilidades todo dia. Cada vez que você decide se vai levar o guarda-chuva para o trabalho ou não, se pega o ônibus lotado que acabou de chegar ou espera por um mais vazio, se compra o plano A ou B do seu celular, se aquela pessoa na foto é ou não seu(sua) namorado(a), você está usando probabilidades. De fato, exageremos e digamos logo que **cada e toda ação do seu dia-a-dia** envolve algum raciocínio probabilístico. Afinal, a Teoria da Probabilidade é a Teoria da Incerteza, e estamos cercados de incerteza a cada momento de nossas vidas, por mais que queiramos negar ou minimizar este fato.

Não quero dizer com isso que a cada respiração você abre seu caderno e usa a Lei da Multiplicação ou o Teorema de Bayes em suas formas matemáticas. Grande parte dos seus raciocínios probabilísticos é feita no seu subconsciente, e até mesmo a parte feita conscientemente não necessariamente lida explicitamente com números.

Para uma teoria que é utilizada diariamente, é impressionante como resultados básicos da Teoria da Probabilidade são extremamente não-intuitivos para a grande maioria das pessoas (e dos matemáticos!). Parte da razão está no parágrafo acima – como muitos dos raciocínios probabilísticos são feitos a nível inconsciente, não temos noção exata dos processos realizados e dos princípios utilizados. Como os eventos probabilísticos são muitíssimo variados, há ocasiões em que não temos a experiência necessária para avaliá-los corretamente – e acabamos por usar analogias incorretas em suas análises. Então nossa intuição nos leva a erros básicos, muitos deles do tipo que nenhum cidadão poderia cometer.

Muitos deles que nenhum cidadão poderia cometer!

Como corrigir tais erros? A resposta não é exatamente sutil: você (**você mesmo!**) precisa praticar mais probabilidade. ”Praticar”no sentido de trazer os raciocínios diários para o seu consciente e entender os princípios básicos que os regem. A ideia não é trocar a sua intuição por propriedades e teoremas – a ideia é **refinar** a sua intuição ao pensar cuidadosamente em alguns problemas elementares.

Então vamos lá: vamos praticar Teoria da Probabilidade básica por 3 dias. Você vai notar que os pré-requisitos matemáticos são muito simples – não usaremos análise combinatória alguma, mas apenas a matemática das proporções! Traga apenas a sua mente aberta e a vontade de explorar essa teoria – e o objetivo de não levar um bode para casa.

Este texto tem várias origens distintas que têm de ser mencionadas.

Em primeiríssimo lugar, a origem deste texto foi um conjunto de notas de aula preparadas pelo saudosíssimo Prof. Augusto César Morgado e pela Profa. Sheila Zani, que chegaram às minhas mãos quando lecionávamos o curso de Introdução à Probabilidade aos alunos da Graduação da Fundação Getulio Vargas (tanto de Administração como de Economia). Àquelas notas, vários professores adicionaram exemplos e exercícios – como o Prof. Paulo Cezar Carvalho e o Prof. Moacyr Alvim Horta, aos quais também direciono muitos agradecimentos. De fato, desejamos algum dia publicar um livro como referência para uma primeira disciplina de graduação em Probabilidade, do qual este texto em suas mãos seria o primeiro capítulo.

Outra imensa fonte de inspiração e exemplos é o *Mid-Career Summer Program* da *Harvard Ken-*

nedy School de governo, da qual participo todo ano desde 1998. De fato, poder-se-ia dizer que o curso de Probabilidade da FGV e as palestras sobre probabilidade que ministro em Harvard criaram uma sinergia benéfica para ambos – mas no fundo no fundo eu só queria usar uma mesóclise e a palavra "sinergia" no prefácio.

Capítulo 1

O que é probabilidade?

> "As questões mais importantes da vida são, em grande parte, nada mais do que problemas de probabilidade... A Teoria da Probabilidade nada mais é do que o cálculo do bom senso." – *Pierre-Simon Laplace, 1812, Théorie Analytique des Probabilités.*
>
> "O bom senso é bem raro-- *Voltaire, 1764, Dictionaire philosophique portatif.*

O objetivo da Teoria da Probabilidade é modelar matematicamente conceitos como *incerteza*, *risco*, *chance*, *possibilidade, verossimilhança, perspectivas* e, até mesmo, *sorte.* Considere as seguintes frases do nosso dia-a-dia:

- A probabilidade de uma moeda lançada "dar" coroa é de 50%;
- A previsão do tempo é de 40% de probabilidade de chuva amanhã;
- A radiografia indica uma moderada probabilidade de Tromboembolia Pulmonar;
- O Copom afirma que aumentou a probabilidade da convergência da inflação para a trajetória de metas;
- Depois da rodada de ontem, a probabilidade do Flamengo ser rebaixado aumentou muito.

Quase todos nós temos ao menos uma intuição do que estas frasem significam. No entanto, encontre a sua resposta para a seguinte pergunta: o que exatamente significa a palavra **probabilidade**? O que exatamente significam as frases acima? Pense nesta pergunta antes de ler os próximos parágrafos...

Seguem aqui duas interpretações comuns do conceito de probabilidade (ambas levam à mesma formulação matemática – apenas as maneiras de expressar e interpretar os resultados mudam com o ponto-de-vista escolhido):

A interpretação **frequentista** imagina um *grande* número de situações semelhantes à apresentada e tenta descobrir em quantas delas o evento em questão realmente acontece; esta proporção seria a probabilidade do evento. Assim, "dividindo o número de coroas obtidas pelo número de lançamentos, a proporção se aproximará de 50% à medida que o número de lançamentos cresce". Esta interpretação pode precisar de um pouco de imaginação: "chove em 40% dos dias com características climáticas *semelhantes* às de amanhã".

A interpretação **subjetiva** (ou **Bayesiana**, ou **epistemológica**) diz que a probabilidade de um evento é apenas uma medida da fé que temos sobre a sua ocorrência. Assim, a probabilidade de um evento varia de indivíduo para indivíduo, dependendo das informações e crenças que ele tenha. Esta interpretação "maleável" nos permite discutir conceitos como a probabilidade de um evento passado ter ocorrido (ou a probabilidade de uma pessoa *ter cometido* um crime). Citando o matemático (e mágico) Persi Diaconis: "*probabilidades não fazem parte das moedas; probabilidades fazem parte das pessoas*".

A Teoria da Probabilidade é apenas um **modelo**. Modelos não são “A REALIDADE” ou “A VERDADE”. *Modelos são úteis exatamente porque simplificam a realidade para que possamos entendê-la*[1]. Se soubéssemos exatamente as características físicas da moeda, sua posição e velocidade iniciais, e as forças nela aplicadas (pelo seu dedão, pela gravidade da Terra, pela resistência do ar, etc.) seríamos capazes de predizer com exatidão se a moeda daria cara ou coroa[2]. Mas trabalhar com todas estas variáveis é impraticável[3] – é preferível inventar este “misterioso 50% de incerteza”, jogando fora os outros detalhes da realidade. Em outras palavras: esta coisa estranha chamada ”probabilidade”é o preço que você paga para não ter que lidar com toda a física do lançamento de moedas. Como lidamos com nossa própria incerteza e ignorância desde que nascemos, o conceito de probabilidade até que não é tão misterioso assim[4].

1.1 Interpretação Freqüentista: as moedas se compensam?

Uma moeda *justa* deu 10 caras seguidas. Qual resultado é mais provável no próximo lançamento: cara ou coroa?

Dizer que “esta moeda provavelmente é viciada” não é válido no problema proposto – afinal, partimos da hipótese de que a moeda *é justa*. Então, “cara” não é a resposta.

Por outro lado, como a moeda é justa, a proporção $\frac{\#(\text{caras})}{\#(\text{lançamentos})}$ deve se aproximar de 50% *a longo prazo* (esta é a interpretação freqüentista, justificada pela Lei dos Grandes Números que não abordaremos aqui). Note: *a longo prazo*! Assim, não há necessidade alguma da moeda “compensar as 10 caras lançadas” logo no próximo lançamento. Então coroa também não é a resposta!

Mas, se a proporção tem de se aproximar de 50%, mesmo a longo prazo, então no futuro as coroas vão ter que recuperar o terreno perdido para as caras, certo? Errado! Mesmo que nos próximos $2n$ lançamentos tivéssemos n caras e n coroas, a proporção nos $2n+10$ lançamentos se aproximaria de 50% para n grande. Afinal:

$$\lim_{n\to\infty} \frac{n}{2n+10} = \frac{1}{2}$$

O problema é que “a proporção se aproxima de 0.5” não é o mesmo que “o número de caras se aproxima da metade do número de lançamentos”[5]! Considere o experimento de John Kerrich – um matemático sul-africano que, prisioneiro de guerra na Dinamarca durante a Segunda Guerra Mundial, lançou uma moeda 10000 vezes:

[1]Num mapa de metrô, as estações aparecem alinhadas; o mapa não mostra todas as ruas, nem os jardins, nem a topografia da cidade. O mapa está *errado*? Não, o mapa é um *modelo*; ele é perfeito para a sua função (saber se a próxima estação é onde eu tenho que descer ou não), mas, se usado além de suas limitações (para planejar uma caminhada, por exemplo), ele falha miseravelmente.

[2]Ou pelo menos é isso que a Física Clássica diria. Já a Mecânica Quântica (um dos pilares da Física Moderna) diria que as partículas que compõem o universo não *estão* em lugar algum – elas têm *probabilidades* de estar em lugares distintos ao mesmo tempo. Esta incerteza não seria devida à nossa incapacidade de criar instrumentos para medi-las, mas seria uma característica intrínseca da natureza do universo. Assim, é impossível ter conhecimento completo sobre o estado atual do universo – ou seja, há uma parcela de chance em todos os fenômenos físicos. Difícil de engolir? Você não está sozinho: até Einstein tinha dificuldades de aceitar este modelo, dizendo estar “convencido de que Deus não joga dados”. Apesar disto, a Mecânica Quântica explica fenômenos observáveis que contradizem frontalmente a Física Clássica de Newton!

[3]*Quase* impraticável: no artigo “Dynamical Bias in the Coin Toss” (2004), Diaconis, Holmes e Montgomery analisam mais cuidadosamente o processo de lançar uma moeda e pegá-la com a mão. Conclusão do artigo: se a moeda mostrava cara no início do lançamento, a probabilidade de mostrar cara ao final é cerca de 50.8%!

[4]Por outro lado, deve haver algum mistério sim. Afinal, os conceitos básicos formais da Teoria da Probabilidade só aparecem no século XVII, quando Pascal e Fermat começaram sua célebre correspondência a respeito de jogos de azar – mas ignorância, incerteza e jogos de azar existem há mais de 5000 anos!

[5]Matematicamente, se $f(n)$ é o número de caras em n lançamentos:

$$\lim_{n\to\infty} \frac{f(n)}{n} = \frac{1}{2} \text{ não significa que } \lim_{n\to\infty} \left(f(n) - \frac{n}{2} \right) = 0$$

Lançamentos	10	40	100	200	400	800	2000	8000	10000
Caras	4	21	44	98	199	413	1013	4034	5067
Acima do Esp.	-1	1	-6	-2	-1	13	13	34	67
Proporção	0.4	0.525	0.44	0.49	0.4975	0.5163	0.5065	0.5043	0.5067

Note como o número de caras acima do "esperado" em termos absolutos **parece oscilar e aumentar com o número de lançamentos!** A Lei dos Grandes números **não diz** que o número de caras se aproxima do número de coroas à medida que o número de lançamentos cresce! Isto não contradiz a Interpretação Freqüentista: a diferença entre caras e coroas diminui **em termos relativos** (ao número total de lançamentos), ou seja, a **proporção** de caras se aproxima de 0.5.

Se você não acredita no experimento de Kerrich, repita-o você mesmo – ou pelo menos faça uma simulação computacional. Os gráficos a seguir foram obtidos exatamente assim, usando Microsoft Excel para simular 10000 lançamentos.

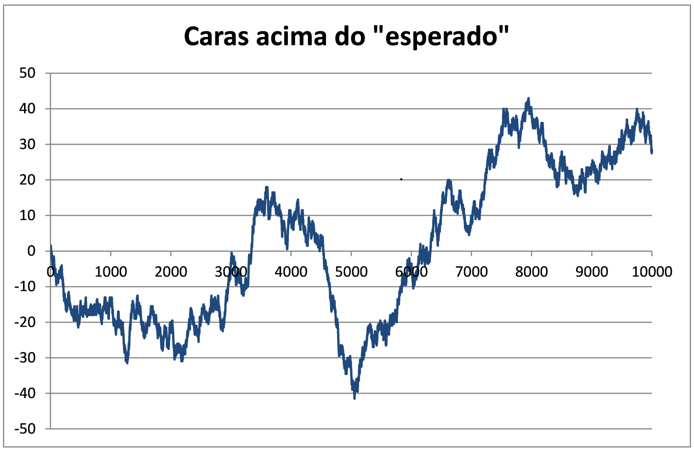


Figura 1.1: Número de caras acima de N/2, versus N (número de lançamentos)

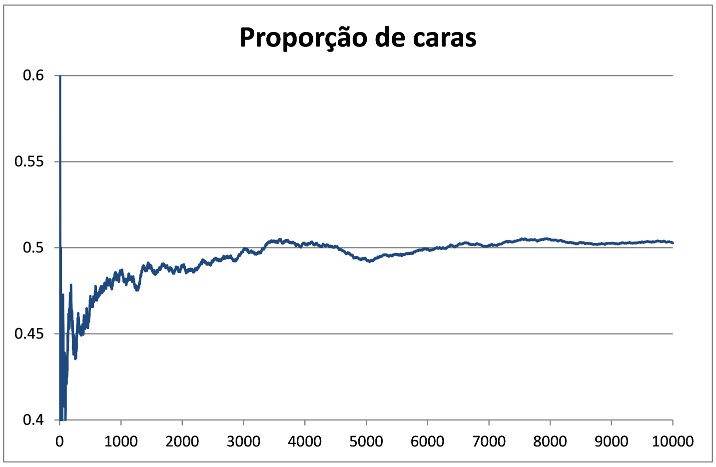


Figura 1.2: Número de caras dividido por N, versus N (número de lançamentos)

Capítulo 2

Modelos de Probabilidade

Considere um experimento qualquer cujo resultado não seja conhecido (ou seja, um experimento **aleatório**). Chamaremos de **espaço amostral** o conjunto de todos os resultados possíveis deste experimento, comumente denotado por S (neste texto, abordaremos apenas o caso em que S é finito). Um **evento** é representado por um subconjunto qualquer de S; diz-se que um evento **ocorre** se algum de seus elementos foi o resultado.

Exemplo 1 *Lança-se um dado e observa-se a face que cai voltada para cima. O espaço amostral é* $S = \{1, 2, 3, 4, 5, 6\}$*. Alguns eventos (que serão utilizados no resto desta seção) são:*
$A =$ *"o número observado é par"*$= \{2, 4, 6\}$
$B =$ *"o número observado é maior do que* 3*"*$=\{4, 5, 6\}$
$C =$ *"o número observado é maior do que* 4*"*$=\{5, 6\}$
Note que, se o resultado for 6*, todos estes três eventos ocorrem.*

A partir de eventos quaisquer, podemos construir novos eventos usando as operações de complemento, união e interseção:

- $\bar{A}$ é o evento "A **NÃO** ocorre";
- $A \cup B$ é o evento "A ocorre **OU** B ocorre";
- $A \cap B$ é o evento "A ocorre **E** B ocorre".

Exemplo 2 *Usando a notação do exemplo anterior:*
$\bar{A} = \{1, 3, 5\} =$ *"o número observado* não *é par"*
$A \cup B = \{2, 4, 5, 6\} =$ *"o número é par* ou *maior do que* 3*"*
$A \cap B = \{4, 6\} =$ *"o número é par* e *é maior do que* 3*".*

Definição 3 *Dois eventos* A *e* B *são chamados de* ***mutuamente excludentes*** *se não podem ocorrer simultaneamente, isto é, se* $A \cap B = \emptyset$*.*

Exemplo 4 X *e* $\bar{X}$ *são sempre mutuamente excludentes; no exemplo anterior,* C *e* $\{1, 2\}$ *são mutuamente excludentes – o número não pode ser maior do que* 4 *e menor do que* 3 *simultaneamente.*

Associaremos a cada evento um número, que chamaremos de **probabilidade do evento** e que traduzirá nossa confiança na capacidade do evento ocorrer.

Definição 5 *Uma* ***probabilidade*** *é uma função que associa a cada evento* A *um número* $\Pr(A)$ *de forma que:*
i) Para todo evento A*,* $0 \leq \Pr(A) \leq 1$*;*

ii) $\Pr(S) = 1$;
iii) Se A e B são eventos mutuamente excludentes então[1]

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

Não é difícil ver que, para atribuir probabilidades a um espaço amostral finito, basta atribuir probabilidades a cada um de seus eventos elementares (representados por conjuntos com um único elemento).

Exemplo 6 *Se acreditarmos que o dado é justo (todas as faces têm a mesma chance), então usaríamos*

$$\Pr(\{1\}) = \Pr(\{2\}) = \Pr(\{3\}) = \Pr(\{4\}) = \Pr(\{5\}) = \Pr(\{6\}) = \frac{1}{6}.$$

Neste caso, teríamos

$$\begin{aligned}
\Pr(A) &= \Pr(\{2\}) + \Pr(\{4\}) + \Pr(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} = 50\% \\
\Pr(B) &= \Pr(\{4\}) + \Pr(\{5\}) + \Pr(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} = 50\% \\
\Pr(C) &= \Pr(\{5\}) + \Pr(\{6\}) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3} = 33.333...\%
\end{aligned}$$

Mas, se você acredita que o dado é viciado, nada impede que você use outros modelos. Por exemplo, talvez eu acredite em

$$\begin{aligned}
\Pr(\{2\}) &= \Pr(\{3\}) = \Pr(\{4\}) = \Pr(\{5\}) = 10\% \\
\Pr(\{1\}) &= 20\%;\ \Pr(\{6\}) = 40\%
\end{aligned}$$

(caso em que o dado não é justo*). Fica a cargo do leitor ver como as probabilidades de A, B e C se alteram neste caso para* 60%, 60% *e* 50%.

As demonstrações das seguintes propriedades do cálculo de probabilidades são simples e deixadas como exercício para o leitor:

Proposição 7 (Lei do Complemento) $\Pr(\bar{A}) = 1 - \Pr(A)$. *Em outras palavras, a probabilidade de um evento ocorrer mais a probabilidade de ele não ocorrer dá* 100%.

Proposição 8 $\Pr(\emptyset) = 0$, *isto é, se um evento é impossível, sua probabilidade deve ser* 0[2].

Proposição 9 (Lei da Adição)

$$\boxed{\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)}$$

isto é, a probabilidade de A ou B ocorrer é a probabilidade de A ocorrer, mais a probabilidade de B ocorrer, menos a probabilidade de A e B ocorrerem (pois esta "havia sido contada duas vezes"!).

[1] Para espaços amostrais infinitos, deveríamos incluir uma condição semelhante com infinitos eventos mutuamente excludentes dois a dois:

$$\Pr(A_1 \cup A_2 \cup ...) = \Pr(A_1) + \Pr(A_2) + ...$$

[2] Note que a recíproca **não** é válida, isto é, $\Pr(A) = 0 \not\Rightarrow A = \emptyset$!

Exemplo 10 *Nos exemplos anteriores, tínhamos* $\bar{A} = \{1,3,5\}$, $A \cup B = \{2,4,5,6\}$ *e* $A \cap B = \{4,6\}$. *Se o dado for justo, teremos:*

$$\begin{aligned} \Pr\left(\bar{A}\right) &= \frac{3}{6} = 50\% = 1 - \Pr(A) \\ \Pr(A \cup B) &= \frac{4}{6} = \frac{3}{6} + \frac{3}{6} - \frac{2}{6} = \Pr(A) + \Pr(B) - \Pr(A \cap B) \end{aligned}$$

Se o dado for viciado como descrito no exemplo anterior, então teríamos

$$\begin{aligned} \Pr\left(\bar{A}\right) &= 20\% + 10\% + 10\% = 40\% = 1 - \Pr(A) \\ \Pr(A \cup B) &= 10\% + 10\% + 10\% + 40\% = 70\% = \\ &= \Pr(A) + \Pr(B) - \Pr(A \cap B) = \\ &= (10\% + 10\% + 40\%) + (10\% + 10\% + 40\%) - (10\% + 40\%) = 70\% \end{aligned}$$

e as leis continuam valendo.

Da Lei da Adição, note que

$$\boxed{\Pr(A \cup B) = \Pr(A) + \Pr(B) \Leftrightarrow \Pr(A \cap B) = 0}$$

ou seja, você pode somar probabilidades **apenas** no caso em que os eventos sejam mutuamente excludentes[3] (bom, e quando você quiser calcular $\Pr(A \cup B)$, a probabilidade de pelo menos um deles ocorrer).

Exemplo 11 *Numa rotina clássica dos trapalhões, Didi argumenta que, sendo sua jornada apenas de 8 horas diárias, ele não precisa trabalhar nos outros* $\frac{2}{3}$ *do tempo do ano. Mas ele também não precisa trabalhar durante* $\frac{2}{7}$ *do ano (finais de semana), e a lei lhe garante um mês de férias – outros* $\frac{1}{12}$ *do ano em que não se trabalha. Somando tudo, a probabilidade do Didi não trabalhar num dia escolhido a esmo seria* $\frac{2}{3} + \frac{2}{7} + \frac{1}{12} = \frac{29}{28}$, *o que já deu mais de* 100% *(sem contar feriados, hora do almoço, Copa do Mundo, etc.)! Assim, o patrão do Didi tem que deixá-lo em casa o ano todo e ainda lhe pagar hora extra... Onde está o erro? Ora, não se podem simplesmente somar estas proporções pois os eventos não são mutuamente excludentes! Por exemplo, Didi contou horas de dormir, em finais de semana, durante as férias, três vezes!*

Um **modelo equiprobabilístico** num espaço amostral S com n elementos associa a cada evento elementar a probabilidade $\frac{1}{n}$. Se o modelo é equiprobabilístico, então a probabilidade de um evento é simplesmente[4]

$$\boxed{\Pr(A) = \frac{\#(A)}{\#(S)} = \frac{\text{“número de casos favoráveis”}}{\text{“número de casos totais”}}}$$

Nota 12 *Cuidado! Um erro muito muito muito comum é usar esta fórmula (ou este tipo de raciocínio) para modelos que não são equiprobabilísticos! Só porque o seu espaço amostral é* $S = \{$*ganho na loteria, não ganho na loteria*$\}$ *não significa que você tem* 50% *de chance de ganhar na loteria! Mais à frente veremos problemas (como o de Monty Hall) onde nossa intuição tem uma vontade terrível de fazer este tipo de raciocínio – e nossa intuição erra redondamente.*

[3]Tecnicamente, isto não é bem verdade – há eventos de probabilidade 0 que *podem* acontecer, e assim $\Pr(A \cap B) = 0$ não significa necessariamente “mutuamente excludentes”... Tais eventos não aparecerão neste texto.

[4]Dado um conjunto X, a notação $\#(X)$ representa o número de elementos de X.

2.1 Exercícios

Ex. 1 *Defina espaços amostrais razoáveis para os seguintes experimentos:*
a) Jogue uma moeda três vezes e anote a seqüência de caras (K) e coroas (C).
b) Jogue dois dados e anote a soma de seus pontos.
c) Jogue dois dados e anote a diferença de seus pontos.
d) Jogue um dado até que o número 6 *apareça e anote quantas vezes ele foi jogado.*
e) Jogue uma moeda 100 *vezes e anote quantas caras foram obtidas.*
f) Tire 6 *bolas de uma urna com* 100 *bolas azuis e* 200 *bolas brancas e anote quantas bolas brancas foram retiradas.*
g) Anote o lanterna do próximo campeonato brasileiro.
h) Anote o instante em que você recebe a primeira ligação telefônica do dia.
i) Anote a temperatura máxima do dia no seu quarto.
Em quais dos exemplos a-g acima é razoável usar um modelo eqüiprovável?

Ex. 2 *A partir dos três axiomas básicos da Probabilidade:*

$$\begin{aligned} \textit{Para todo evento } A &: \quad 0 \leq \Pr(A) \leq 1; \\ \textit{Para o espaço amostral } S &: \quad \Pr(S) = 1; \\ \textit{Para quaisquer eventos mutuamente excludentes } A \textit{ e } B &: \quad \Pr(A \cup B) = \Pr(A) + \Pr(B) \end{aligned}$$

Demonstre as seguintes propriedades:
a) A Lei do Complemento: $\Pr\left(\bar{A}\right) = 1 - \Pr(A)$ *[Dica: A e $\bar{A}$ são mutuamente excludentes.]*
b) $\Pr(\emptyset) = 0$ *[Dica: use o item anterior.]*
c) A Lei da Adição: $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$ *[Dica: $B - A$ e $A \cap B$ são mutuamente excludentes, assim como $B - A$ e A.]*
d) Se $A \subseteq B$ então $\Pr(A) \leq \Pr(B)$ *[Dica: $B - A$ e A são...]*

Ex. 3 *Mostre que*

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

Ex. 4 *Dados* $\Pr(A) = 0.4$, $\Pr(B) = 0.5$, $\Pr(C) = 0.3$, $\Pr(A \cap B) = 0.3$, $\Pr(A \cap C) = 0$ *e* $\Pr(B \cap C) = 0.1$, *determine:*
a) $\Pr(A \cap B \cap C)$
b) $\Pr(A \cup (B \cap C))$
c) $\Pr(A \cap (B - C))$
d) $\Pr(A - (B \cup C))$
e) $\Pr(A \cup B \cup C)$

Ex. 5 *Em certa escola a probabilidade de um aluno ser torcedor do Flamengo é $0,6$, de assistir novela é $0,7$ e de gostar de praia é $0,8$. Entre que valores está compreendida a probabilidade de um aluno dessa escola, simultaneamente, torcer pelo Flamengo, assistir novela e gostar de praia?*

Ex. 6 *Lança-se uma moeda* justa *três vezes e anota-se a seqüência de Caras (K) e Coroas (C) obtidas.*
a) Que modelo de probabilidade lhe parece razoável em S?
Sejam A o evento "dois primeiros resultados são iguais", B o evento "o primeiro lançamento é uma cara" e C o evento "pelo menos um lançamento é uma cara".
b) Escreva A, B e C como subconjuntos de S e calcule as probabilidades de cada um.
c) Interprete os seguintes eventos em linguagem comum e calcule as suas probabilidades:
i) $\bar{A}$ *ii)* $\bar{C}$ *iii)* $A \cap B$ *iv)* $B \cap C$ *v)* $B \cup C$ *vi)* $A \cup B$.

Ex. 7 *Lança-se uma moeda* justa *até obter-se duas caras ou duas coroas, não necessariamente consecutivas (ou seja, Kuerten e Coria disputam uma partida de tênis em três sets e têm chances iguais de vencer cada set). Anota-se a seqüência obtida (os vencedores de cada set). Repita os itens a-c do exercício anterior. Que respostas mudaram?*

Ex. 8 *Dois dados são lançados – um vermelho e um verde. Escreva um espaço amostral para este experimento, e calcule a probabilidade de a soma dos dois dados ser* 9*. O problema se altera se os dados forem da mesma cor?*

Ex. 9 *Os* 12 *times do campeonato do Rio são sorteados de forma completamente aleatória em dois grupos de* 6 *times cada. Qual a probabilidade de o Flamengo e o Fluminense acabarem no mesmo grupo?*

Ex. 10 *Em uma roda são colocadas n pessoas. Qual é a probabilidade de duas dessas pessoas ficarem juntas?*

Ex. 11 *Em uma fila são colocadas n pessoas. Qual é a probabilidade de duas dessas pessoas ficarem juntas?*

Ex. 12 *Laura e Telma retiram cada uma um bilhete numerado de uma urna que contém bilhetes numerados de* 1 *a* 100*. Determine a probabilidade do número de Laura ser maior que o de Telma, supondo a extração:*
a) sem reposição.
b) com reposição.

Ex. 13 *Três jogadores, A, B e C, disputam um torneio. Os três têm probabilidades iguais de ganhar o torneio; têm também probabilidades iguais de tirarem o segundo lugar e têm probabilidades iguais de tirarem o último lugar. É necessariamente verdadeiro que cada uma das seis ordens possíveis de classificação dos três jogadores tem probabilidade* $\frac{1}{6}$ *de ocorrer?*

Ex. 14 *Dois dados são lançados. Os eventos* $A =$ *"número do primeiro dado foi a" e* $B =$ *" a soma dos dados é b" são mutuamente excludentes (onde* $1 \leq a \leq 6$ *e* $2 \leq b \leq 12$*). Que outras conclusões você pode tirar sobre a e b?*

2.2 Respostas dos Exercícios

Resp. 1 *Note, estamos pedindo apenas espaços amostrais, não estamos pedindo probabilidades:*
a) $S = \{KKK, KKC, KCK, CKK, CCK, CKC, KCC, CCC\}$
b) $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
c) $S = \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$
d) $S = \{1, 2, 3, 4, 5, ...\} = \mathbb{N}^*$
e) $S = \{1, 2, 3, 4, 5, ..., 100\}$
f) $S = \{0, 1, 2, 3, 4, 5, 6\}$
g) $S = \{$*Flamengo*$\}$ *:) :) :) Tá bom,* $S = \{$*Flamengo, Fluminense, Botafogo, Vasco, ..., Americana*$\}$
h) $S = [0, 24]$ *(onde marquei o tempo em horas)*
i) $S = [0, 45]$ *(em Graus Celsius)*
Num mundo de moedas e dados justos, lançamentos independentes e times que não fazem pré-temporada, apenas (a) é eqüiprovável.

Resp. 2 *a) Como A e $\bar{A}$ são disjuntos:*

$$\Pr\left(A \cup \bar{A}\right) = \Pr(A) + \Pr\left(\bar{A}\right)$$

Mas $A \cup \bar{A} = S$ e $\Pr(S) = 1$.
b) Como $\emptyset = \bar{S}$ e $\Pr(S) = 1$ usando o item anterior, $\Pr(\emptyset) = 1 - \Pr(S) = 0$.
c) Faça um diagrama. Como $B - A$ e A são mutuamente excludentes:

$$\Pr((B - A) \cup A) = \Pr(B - A) + \Pr(A)$$

Mas a união do lado esquerdo é $A \cup B$, isto é:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B - A)$$

Agora, como $B - A$ e $A \cap B$ são mutuamente excludentes,

$$\Pr((B - A) \cup (A \cap B)) = \Pr(B - A) + \Pr(A \cap B)$$

e a união do lado esquerdo é B, Então:

$$\Pr(B) = \Pr(B - A) + \Pr(A \cap B)$$

Tire $\Pr(B - A)$ daqui e substitua na outra para acabar o problema.
d) Vimos ali em cima que

$$\Pr(B - A) = \Pr(B) - \Pr(A \cap B)$$

Como, neste caso, $A \subseteq B$, temos $A \cap B = A$, isto é

$$\Pr(B) - \Pr(A) = \Pr(B - A) \geq 0$$

pois toda probabilidade é maior ou igual a 0. Acabou.

Resp. 3 *Desenhe um diagrama de Venn – há 7 pedaços excludentes para $A \cup B \cup C$. Escreva cada termo da expressão do lado direito em função destes 7 pedaços, some tudo e veja que, depois de cortar muita coisa, cada pedaço aparece representado apenas uma vez, dando $A \cup B \cup C$.*

Resp. 4 *A princípio, temos $\Pr(A \cap C) = 0$, $\Pr(C) = 0.3$, $\Pr(A \cap B) = 0.3$, $\Pr(B \cap C) = 0.1$:*

	AB	$A\bar{B}$	$\bar{A}B$	$\bar{A}\bar{B}$	*Total*
C	0	0	0.1		0.3
$\bar{C}$					
Total	0.3				

Como $\Pr(A) = 0.4$ e $\Pr(B) = 0.5$, conseguimos completar dois novos totais:

	AB	$A\bar{B}$	$\bar{A}B$	$\bar{A}\bar{B}$	*Total*
C	0	0	0.1		0.3
$\bar{C}$					
Total	0.3	$0.4 - 0.3 = 0.1$	$0.5 - 0.3 = 0.2$		

Agora virou Sudoku:

	AB	$A\bar{B}$	$\bar{A}B$	$\bar{A}\bar{B}$	*Total*
C	0	0	0.1	0.2	0.3
$\bar{C}$	0.3	0.1	0.1	0.2	0.7
Total	0.3	0.1	0.2	0.4	1

a) 0 *b)* 0.5 *c)* 0.3 *d)* 0.1 *e)* 0.8

Resp. 5 *Entre* 0.1 *e* 0.6.

Resp. 6 *Fazendo* $S = \{CCC, CCK, CKC, KCC, KKC, KCK, CKK, KKK\}$ *é razoáve usar um modelo eqüiprovável. Como* $A = \{CCC, CCK, KKC, KKK\}$, $B = \{KCC, KCK, KKC, KKK\}$ *e* $C = \{KKK, CCK, CKC, KCC, KKC, KCK, CKK\}$ *, temos:*
$\Pr(A) = \frac{4}{8}$*;* $\Pr(B) = \frac{4}{8}$*.* $\Pr(C) = \frac{7}{8}$
$\bar{A}$ *:dois primeiros resultados diferentes, probabilidade* $\frac{4}{8}$
$\bar{C}$ *:nenhuma cara, todas são coroas, probabilidade* $\frac{1}{8}$
$A \cap B$*: duas caras nos dois primeiros lançamentos,* $\frac{2}{8} = \frac{1}{4}$
$B \cap C$*: o primeiro é cara, que é* B *de novo, com* $\frac{4}{8}$ *de chance.*
$B \cup C$*: basta uma cara, que é* C *de novo, com* $\frac{7}{8}$ *de chance.*
$A \cup B$*: cara de primeira ou duas coroas nas duas primeiras,* $\frac{6}{8} = \frac{3}{4}$ *de chance.*

Resp. 7 *É quase igual ao anterior, mas* CCC *e* CCK *viram simplesmente* CC*, enquanto* KKC *e* KKK *viram simplesmente* KK *(pois o jogo acaba dois a zero). Agora* $A = \{CC, KK\}$, $B = \{KK, KCK, KCC\}$ *e* $C = \{KK, CCK, CKC, KCC, KCK\}$ *(note como* CCK *some daqui, pois esta última coroa não exisitirá). As probabilidades que envolvem* A *e* B *não mudam, mas* C *mudou:* $\Pr(C) = \frac{6}{8} = \frac{3}{4}$, $\Pr\left(\bar{C}\right) = \frac{1}{4}$, $\Pr(B \cup C) = \Pr(C) = \frac{6}{8}$ *e* $\Pr(B \cap C) = \Pr(B) = \frac{4}{8}$.

Resp. 8 *Probabilidade* $\frac{4}{36} = \frac{1}{9}$*, que não se altera se os dados forem da mesma cor.*

Resp. 9 $\frac{5}{11}$

Resp. 10 $\frac{2}{n-1}$

Resp. 11 $\frac{2}{n}$

Resp. 12 *a)* $\frac{1}{2}$ *b)* $\frac{4950}{10000} = 49.5\%$

Resp. 13 *Não. Podia ser* $\Pr(ABC) = \Pr(BCA) = \Pr(CAB) = \frac{1}{3}$ *e as outras três ordens impossíveis, por exemplo.*

Resp. 14 *Serão mutuamente excludentes quando* $b - a < 1$ *ou* $b - a > 6$.

Capítulo 3

Probabilidade Condicional

Se tivermos informação adicional sobre um experimento, podemos ser forçados a reavaliar as probabilidades dos eventos a ele associados.

Exemplo 1 *Como no capítulo anterior, jogue um dado e anote o valor de sua face superior. Então* $S = \{1, 2, 3, 4, 5, 6\}$*. Sejam* $A = \{2, 4, 6\}$*,* $B = \{4, 5, 6\}$ *e* $C = \{5, 6\}$*. Se o dado é justo, teremos:*

$$\Pr(A) = \frac{3}{6};\ \Pr(B) = \frac{3}{6};\ \Pr(C) = \frac{2}{6}$$

Agora, suponha que você sabe de alguma forma que o número rolado é par. Então seu novo universo é $A = \{2, 4, 6\}$. Sabendo-se que o número é par, *qual a probabilidade de ele ser maior do que* 3*? Ou seja, qual a chance de* B *ocorrer* na ceretza de que A ocorreu*? Esta é a chamada* ***probabilidade condicional de*** B ***dado*** A*; neste caso*

$$\Pr(B|A) = \frac{2}{3}$$

pois há apenas 2 *casos "favoráveis a* B*" dentre os* 3 *casos "possíveis em* A*". Analogamente, convença-se de que:*

$$\Pr(A|B) = \frac{2}{3};\ \Pr(A|C) = \frac{1}{2};\ \Pr(C|A) = \frac{1}{3};\ \Pr(B|C) = 1;\ \Pr(C|B) = \frac{2}{3}$$

Escreva estas probabilidades em linguagem comum: $\Pr(A|B) = \frac{2}{3}$ *significa que "sabendo-se que o número é maior que três, há* $\frac{2}{3}$ *de chance de ele ser par". Numa interpretação freqüentista, diríamos "se rolarmos o dado várias vezes, dará um número par cerca de* $\frac{2}{3}$ *das vezes em que o número foi maior do que* 3*".*
Note que $\Pr(B|C) = 100\%$*, isto é, "na certeza de que deu mais do que quatro, é óbvio que deu mais do que três", ou seja, "*B *acontece sempre que* C *acontece".*
Por outro lado, $\Pr(C|B) = \frac{2}{3}$ *apenas. Assim, "de cada* 3 *vezes em que* B *ocorre,* C *ocorre em apenas* 2*".*

O exemplo acima inspira a seguinte fórmula:

Definição 2 *Sejam* A *e* B *dois eventos com* $\Pr(A) \neq 0$*. A probabilidade condicional de* B *dado* A *é*

$$\boxed{\Pr(B|A) = \tfrac{\Pr(A \cap B)}{\Pr(A)}}$$

Exemplo 3 *Usando esta fórmula no exemplo anterior, temos*

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{2/6}{3/6} = \frac{2}{3}$$

Note como o número de elementos do espaço amostral S *(no caso,* 6*) desaparece e ficamos ao final apenas com a proporção dos elementos de* B *(que estão também em* A*) com relação aos elementos de* A*.*

Exemplo 4 *A tabela abaixo dá a distribuição dos alunos de uma turma, por sexo e por carreira pretendida:*

	M	*F*	*total*
ADM	*15*	*45*	*60*
ECO	*21*	*9*	*30*
total	*36*	*54*	*90*

Escolhe-se ao acaso um aluno. Sejam M, F, A e E os eventos o aluno selecionado é do sexo masculino, é do sexo feminino, cursa ADM e cursa ECO, respectivamente. Temos:
i) $\Pr(A) = \frac{60}{90}$, *isto é,* 66.67% *dos alunos cursam ADM; os outros* $\frac{30}{90} = 33.33\%$ *cursam ECO. Se você escolher um aluno ao acaso, há* 66.67% *de chance de ele ser de ADM.*
ii) $\Pr(A|M) = \frac{15}{36}$, *isto é,* 41.67% dos alunos homens *cursam ADM; os outros* 58.33% dos homens *estão em ECO. Se você escolher um aluno* homem *ao acaso, há* 41.67% *de chance de ele estudar ADM.*
iii) $\Pr(M|A) = \frac{15}{60}$, *isto é,* 25% dos alunos de ADM *são homens; se você escolher um aluno* de ADM *ao acaso, há* 25% *de chance deste aluno ser homem.*

A fórmula da probabilidade condicional é freqüentemente utilizada para **descobrir** $\Pr(A \cap B)$:

Proposição 5 (Lei da Multiplicação)

$$\boxed{\Pr(A \cap B) = \Pr(B|A) . \Pr(A) = \Pr(A|B) . \Pr(B)}$$

Exemplo 6 *Uma urna contém* 4 *bolas brancas e* 6 *bolas pretas. Sacam-se, sucessivamente e sem reposição, duas bolas dessa urna. Determine a probabilidade de ambas serem brancas.*
Solução: *Sejam* $B_1 = \{\text{primeira bola é branca}\}$ *e* $B_2 = \{\text{a segunda bola é branca}\}$. *Então*

$$\Pr(B_1 \cap B_2) = \Pr(B_1) . \Pr(B_2|B_1) = \frac{4}{10}\frac{3}{9} = \frac{12}{90} = \frac{2}{15}$$

Note que foi bastante simples o cálculo de $\Pr(B_2|B_1)$. *Realmente, na certeza de que a primeira bola foi branca, é fácil calcular a probabilidade da segunda bola ser branca, pois, para a segunda extração, a urna está com 3 bolas brancas e 6 pretas. De modo mais geral, é fácil calcular probabilidades condicionais quando as coisas estão na ordem certa, isto é, é fácil calcular probabilidades de coisas futuras na certeza de coisas passadas.*

Exemplo 7 *Você tem duas moedas, uma com duas caras e a outra justa. Escolha uma delas e a lance. O resultado é cara. Qual a chance de ela ser a moeda "viciada"?*
Solução: *seja V o evento "escolhemos a moeda viciada" e K o evento "deu cara". Então:*

$$\Pr(V|K) = \frac{\Pr(V \cap K)}{\Pr(K)}$$

Mas

$$\Pr(V \cap K) = \Pr(V) . \Pr(K|V) = \frac{1}{2}\frac{1}{1} = \frac{1}{2}$$

e

$$\Pr(K) = \Pr(V \cap K) + \Pr\left(\bar{V} \cap K\right)$$

onde

$$\Pr\left(\bar{V} \cap K\right) = \Pr\left(\bar{V}\right) . \Pr\left(K|\bar{V}\right) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$$

Juntando tudo

$$\begin{aligned} \Pr(K) &= \frac{1}{2} + \frac{1}{4} \Rightarrow \\ &\Rightarrow \Pr(V|K) = \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{4}} = \frac{2}{3} \end{aligned}$$

3.1 Visualização: Tabelas e Árvores

Qualquer problema básico de probabilidade pode ser feito de maneita puramente algébrica; no entanto, ferramentas visuais como tabelas e árvores facilitam bastante a resolução e discussão de tais problemas. Vejamos alguns exemplos.

Exemplo 1 *Uma urna contém* 4 *bolas brancas e* 6 *bolas pretas. Sacam-se, sucessivamente e sem reposição, duas bolas dessa urna. Determine a probabilidade da primeira bola ser branca sabendo que a segunda bola é branca.*
Solução: *Sejam* B_1 *e* B_2 *como no problema anterior. Queremos* $\Pr(B_1|B_2)$. *Note que essa é uma probabilidade do passado na certeza do futuro. Aqui usamos a fórmula da definição de probabilidade condicional:*

$$\Pr(B_1|B_2) = \frac{\Pr(B_1 \cap B_2)}{\Pr(B_2)}$$

Calculamos $\Pr(B_1 \cap B_2)$ *no exemplo anterior. Para calcular* $\Pr(B_2)$, *basta notar a simetria do problema – não há motivo para imaginar que a segunda bola seja branca mais ou menos freqüentemente do que a primeira! Se este argumento não lhe parece convincente, faça o seguinte: considere separadamente os casos em que a primeira bola é branca e os casos onde a primeira bola não é branca:*

$$\Pr(B_2) = \Pr(B_1 \cap B_2) + \Pr\left(\overline{B_1} \cap B_2\right)$$

A primeira parcela já foi calculada. Quanto à segunda:

$$\Pr\left(\overline{B_1} \cap B_2\right) = \Pr\left(\overline{B_1}\right).\Pr\left(B_2|\overline{B_1}\right) = \frac{6}{10}.\frac{4}{9} = \frac{4}{15}$$

já que, após retirar uma bola preta, ficam 4 *brancas dentre* 9 *bolas. Juntando tudo,*

$$\Pr(B_2) = \frac{2}{15} + \frac{4}{15} = \frac{6}{15} = \frac{4}{10}$$

como havíamos afirmado anteriormente. Enfim:

$$\Pr(B_1|B_2) = \frac{\Pr(B_1 \cap B_2)}{\Pr(B_2)} = \frac{2/15}{6/15} = \frac{1}{3}$$

Exemplo 2 (Árvore)

Ilustremos a solução anterior por uma árvore de probabilidades; para tanto:
– Desenhe as retiradas da esquerda para a direita, criando ramificações sempre que houver um evento aleatório:

A seguir, escreva as probabilidades **condicionais** acima de cada ramo, isto é, a probabilidade de aquele ramo ocorrer **dada toda a estória passada até aquele ramo.** Por exemplo, note que escrevemos $\Pr(P_2|B_1) = 6/9$ sobre o ramo superior direito.
– Agora, para calcular a probabilidade de um caminho inteiro da raiz até a folha, basta multiplicar probabilidades. Por exemplo, $\Pr(B_1 \cap B_2) = \frac{4}{10}\frac{3}{9} = \frac{2}{15}$.
– Enfim, use as probabilidades encontradas para resolver o seu problema. Por exemplo, neste caso note que B_2 ocorre em dois caminhos distintos, cujas probabilidades são $\frac{4}{10}\frac{3}{9} = \frac{2}{15}$ e $\frac{6}{10}\frac{4}{9} = \frac{4}{15}$ Assim, $\Pr(B_2) = \frac{6}{15}$, e $\Pr(B_1|B_2) = \frac{2/15}{6/15} = \frac{1}{3}$.

Exemplo 3 (Tabela) *Menos intuitiva (mas mais fácil de digitar) seria uma tabela. Para criá-la, comece de uma "população fictícia de* 900 *retiradas". A partir dali, escreva os totais esperados para* B_1 *(40% do total geral) e* $\overline{B_1}$ *(60% do total geral). Usando as probabilidades condicionais, preencha então as quatro células no interior da tabela. Enfim escreva os totais para* B_2 *e* $\overline{B_2}$*, e resolva o problema que quiser:*

	B_1	$\overline{B_1}$	*Totais*
B_2	120	240	360
$\overline{B_2}$	240	300	540
Totais	360	540	900

Note que ***não estamos dizendo*** *que de cada* 900 *experimentos, exatamente* 360 *terão a primeira bola branca – estamos apenas dizendo que as proporções representadas pelos números da tabela são exatamente as probabilidades do problema! No problema em questão, vem*

$$\Pr(B_1|B_2) = \frac{120}{360} = \frac{1}{3}$$

Enfim, note que o número 900 *não tem importância alguma e desaparecerá quando qualquer probabilidade for calculada. Você poderia até colocar* 1 *no lugar de* 900 *se desejasse.*

No problema anterior, é interessante notar que as posições das bolas (primeira e segunda) são intercambiáveis, isto é, $\Pr(B_1) = \Pr(B_2)$, e mais, $\Pr(B_1|B_2) = \Pr(B_2|B_1)$, e assim por diante.

3.2 Probabilidade Total e Teorema de Bayes

Os problemas anteriores ilustram duas técnicas comuns para obtenção de probabilidades:

Proposição 1 (Lei da Probabilidade Total) *Suponha que* B_1, B_2, ..., B_n *formam uma partição*[1] *de* S*. Então*

$$\begin{aligned} \Pr(A) &= \Pr(A \cap B_1) + \Pr(A \cap B_2) + ... + \Pr(A \cap B_n) \\ &= \Pr(A|B_1).\Pr(B_1) + \Pr(A|B_2).\Pr(B_2) + ... + \Pr(A|B_n).\Pr(B_n) \end{aligned}$$

Em particular, a partição $S = B \cup \bar{B}$ *nos dá*

$$\boxed{\Pr(A) = \Pr(A|B).\Pr(B) + \Pr\left(A|\bar{B}\right).\Pr\left(\bar{B}\right)}$$

De fato, A é a união dos conjuntos (sem interseção dois a dois!) da forma $A \cap B_i$, justificando a primeira igualdade. A segunda igualdade vem simplesmente de aplicar a Lei da Multiplicação várias vezes. Compare esta "lei" com os exemplos da subseção anterior.

Proposição 2 (Teorema de Bayes) *Suponha que* B_1, B_2, ..., B_n *formam uma partição de* S*. Então:*

$$\Pr(B_1|A) = \frac{\Pr(A|B_1).\Pr(B_1)}{\Pr(A|B_1).\Pr(B_1) + \Pr(A|B_2).\Pr(B_2) + ... + \Pr(A|B_n).\Pr(B_n)}$$

Em particular

$$\boxed{\Pr(B|A) = \frac{\Pr(A|B).\Pr(B)}{\Pr(A|B).\Pr(B) + \Pr\left(A|\bar{B}\right).\Pr\left(\bar{B}\right)}}$$

O Teorema de Bayes nos dá a fórmula exata para calcular uma condicional quando temos as condicionais "na outra ordem". Apesar de muito útil, em geral ele é mais fácil de ser entendido com o auxílio de tabelas ou árvores – novamente, perceba como ele foi utilizado nos exemplos anteriores.

[1] Isto é, eles são mutuamente excludentes dois a dois e $B_1 \cup B_2 \cup ... \cup B_n = S$.

3.3 Independência

Em algumas ocasiões, o conhecimento sobre a ocorrência de um evento não muda a probabilidade de um outro – este é o conceito de independência estatística:

Definição 1 *Dois eventos (de probabilidades não nulas) A e B são ditos* ***independentes*** *se o conhecimento de um deles não afeta a probabilidade do outro ocorrer, isto é, se*

$$\Pr(B|A) = \Pr(B)$$

Intuitivamente, se tudo o que você quer saber é se B acontece ou não, informações sobre o evento A não vão lhe ajudar em nada (e, portanto, você não pagaria dinheiro algum pela informação de A ter acontecido ou não – mesmo que você tenha certeza de que A acontece, a probabilidade de B não muda).

Exemplo 2 *No caso dos dados do início desta seção, note que A e C são independentes, pois*

$$\Pr(A|C) = \frac{1}{2} = \Pr(A); \quad \Pr(C|A) = \frac{1}{3} = \Pr(C)$$

No entanto, A e B não são independentes, muito menos B e C.

Note que, da definição de probabilidade condicional, concluímos que dois eventos são independentes se, e somente se, $\Pr(B|A) = \frac{\Pr(A\cap B)}{\Pr(A)} = \Pr(B)$, isto é

$$\boxed{A \text{ e } B \text{ são independentes} \Leftrightarrow \Pr(A \cap B) = \Pr(A).\Pr(B)}$$

De quebra, passando $\Pr(B)$ para o lado esquerdo, acabamos de mostrar que

$$\Pr(B|A) = \Pr(B) \Leftrightarrow \Pr(A \cap B) = \Pr(A) . \Pr(B) \Leftrightarrow \Pr(A|B) = \Pr(A)$$

Pode-se mostrar também que, se A e B são independentes, então

$$\Pr(B|A) = \Pr(B) = \Pr\left(B|\bar{A}\right)$$

Exemplo 3 *Suponha que, numa família com duas crianças, a probabilidade do filho estar gripado é 40% ($\Pr(H) = 0.4$) e a probabilidade da filha estar gripada é 60% ($\Pr(M) = 0.6$). É possível calcular a probabilidade de ambos estarem gripados?*
Se supusermos que estes dois eventos são independents, então é simples: $\Pr(H \cap M) = (0.4)(0.6) =$ 24%. Mas será que esta suposição é razoável? Afinal, se um deles estiver gripado, imagina-se que a probabilidade do outro estar gripado aumenta. Matematicamente falando, acreditamos que $\Pr(H|M) \geq \Pr(H) = 40\%$, e a probabilidade condicional é que teria de ser usada:

$$\Pr(H \cap M) = \Pr(H|M) . \Pr(M)$$

Sem mais dados, não é possível resolver o problema.

Exemplo 4 *Por outro lado, se no problema anterior forem dados $\Pr(H) = 0.4$, $\Pr(M) = 0.6$ e $\Pr(H \cap M) = 0.3$, é possível* verificar *se os eventos H e M são independentes! De fato, como $\Pr(H \cap M) \neq \Pr(H) . \Pr(M)$, os eventos não seriam independentes. Outras maneiras de chegar à mesma conclusão:*

$$\begin{aligned} \Pr(H|M) &= \frac{0.3}{0.6} = 0.5 > 0.4 = \Pr(H) \\ \Pr(M|H) &= \frac{0.3}{0.4} = 0.75 > 0.6 = \Pr(M) \end{aligned}$$

Neste caso, diz-se que o evento H ***atrai*** *o evento M ou que os eventos são* ***positivamente associados***.

Exemplo 5 *Algumas pesquisas estatísticas podem causar constrangimentos aos entrevistados com perguntas do tipo "você usa drogas?" e correm o risco de não obter respostas sinceras ou não obter respostas de espécie alguma. Para estimar a proporção p de usuários de drogas em certa comunidade, pede-se ao entrevistado que, longe das vistas do entrevistador, jogue uma moeda: se o resultado for coroa, responda a "você usa drogas?" e, se o resultado for cara, responda "sim". Assim, caso o entrevistado diga sim, o entrevistador não saberá se ele é um usuário de drogas ou se a moeda deu cara.*
Se s é a probabilidade de um entrevistado responder sim, s é facilmente estimado pela proporção de respostas sim obtidas nas entrevistas. Estime p a partir de s.
Solução: *seja D o evento "usuário disse que usa drogas" e K o evento "moeda deu cara". Colocando tudo numa tabela, temos.*

	D	$\overline{D}$	
K			0.5
$\overline{K}$			0.5
	p	$1-p$	1

Como D e K são independentes, podemos completar a tabela simplesmente multiplicando as probabilidades correspondentes:

	D	$\overline{D}$	
K	$0.5p$	$0.5\,(1-p)$	0.5
$\overline{K}$	$0.5p$	$0.5\,(1-p)$	0.5
	p	$1-p$	1

Note que os entrevistados que dizem "sim" estão em 3 lugares da tabela acima – ambas da linha K e os usuários de drogas da célula $D\bar{K}$. Assim

$$s = 0.5p + 0.5\,(1-p) + 0.5p = 0.5\,(1+p) \Rightarrow p = 2s - 1$$

Por exemplo, se 60% dos entrevistados respondem sim, você pode estimar em 20% a proporção de usuários de drogas.

3.4 Estudo de Caso: Teste Elisa e AIDS

Problema: Uma pessoa deseja saber se está contaminada com o vírus da AIDS ou não. Ao fazer o teste, este pode indicar POSITIVO (+) ou NEGATIVO (-). No entanto, nenhum teste é 100% correto – em algumas ocasiões, o teste pode ser + mesmo que esta pessoa não tenha a doença (o chamado *falso positivo*); em outras, apesar do paciente estar doente, o teste apresenta resultado - (um *falso negativo*). Digamos que você tem em mãos os seguintes dados a respeito do "Teste Elisa" para AIDS:

- Apenas 0.5% das pessoas no seu país têm AIDS (diz-se que a *prevalência* da doença é 0.5%)
- "Elisa" identifica corretamente (como +) 98% das pessoas que têm o vírus (diz-se que a *sensitividade* do teste é de 98%);
- "Elisa" identifica corretamente (-) 93% das pessoas que não têm o vírus (diz-se que a *especificidade* do teste é de 93%).

Um paciente escolhido aleatoriamente neste país é testado e o resultado é +. Qual a probabilidade de ele ter o vírus?

Resposta: Este tipo de problema pode ser facilmente resolvido usando uma tabela com uma população "fictícia". Comece a tabela supondo que haja 10000 pessoas neste país, destas, 50 teriam

AIDS e as outras 9950 não teriam:

	$AIDS$	$\overline{AIDS}$	Totais
$+$			
$-$			
Totais	50	9950	10000

Mas, daquelas 50, 98% (49 pessoas) testarão positivo; daquelas 9950, 7% (696.5 pessoas) serão falsos positivos. Use estes números para completar a tabela (não se preocupe com as populações fictícias e fracionárias – afinal, o que interessam são as proporções):

	$AIDS$	$\overline{AIDS}$	Totais
$+$	49	696.5	745.5
$-$	1	9253.5	9254.5
Totais	50	9950	10000

Agora podemos proceder a quaisquer respostas como no exemplo anterior. Por exemplo, se sabemos que o paciente testou +, qual a chance de ele ter $AIDS$? Seria:

$$\Pr(AIDS|+) = \frac{49}{745.5} = 6.573\%$$

Dizemos que o *poder preditivo positivo do teste* é de 6.573%.

Note como o Teorema de Bayes resolve o problema com uma fórmula só, mas escondendo um bocado a intuição do exemplo. Afinal, os dados são

$$\Pr(+|AIDS) = 98\%; \quad \Pr\left(-|\overline{AIDS}\right) = 93\%; \quad \Pr(AIDS) = 0.5\%$$

dos quais tiramos via Lei do Complemento:

$$\Pr\left(+|\overline{AIDS}\right) = 7\% \text{ e } \Pr\left(\overline{AIDS}\right) = 99.5\%$$

Enfiando tudo na fórmula de Bayes:

$$\begin{aligned}
\Pr(AIDS|+) &= \frac{\Pr(+|AIDS) \,.\, \Pr(AIDS)}{\Pr(+|AIDS) \,.\, \Pr(AIDS) + \Pr\left(+|\overline{AIDS}\right) . \Pr\left(\overline{AIDS}\right)} = \\
&= \frac{(0.98)\,(0.005)}{(0.98)\,(0.005) + (0.07)\,(0.995)} = 6.573\%
\end{aligned}$$

Análise: Os dados apresentados acima são compatíveis com os valores de sensitividade e especificidade do Teste Elisa para AIDS. A prevalência da AIDS varia muito de país para país e de ano para ano – em 1990, era de cerca de 1% nos Estados Unidos, mas apenas 0.2% na Austrália. Como um teste que parecia tão preciso pode errar tanto? O problema é que esta doença é muito pouco comum (apenas 0.5% da população a tem). É mais provável que esta pessoa seja um dos "proporcionalmente poucos falsos positivos" dentre a grande massa de pessoas sadias do que um dos "proporcionalmente muitos corretos positivos" dentre as poucas pessoas doentes!

Este tipo de probabilidade tem de ser divulgada às pessoas que são testadas! É por este motivo que, ao testar + para uma doença, você deve realizar um segundo teste!

Por outro lado, não é que o primeiro teste foi "inútil" não. Um resultado positivo no primeiro teste aumenta a probabilidade de doença de 0.5% (**a priori**) para uns 7% (**a posteriori**). O paciente deve sim se preocupar muito mais do que antes do teste. O poder do teste não está em *determinar* a probabilidade de se estar doente, mas em *aumentá-la* a partir duma probabilidade a priori. A propósito, note como estes cálculos podem variar terrivelmente dependendo do que se sabe sobre este indivíduo *antes* de ele se testar. Por exemplo, se o indivíduo pertence a um grupo de risco (digamos, no caso da AIDS, se é hemofílico) não é razoável usar o 0.5% como probablidade a priori – usar-se-ia um número maior que refletisse a percentagem de hemofílicos que contraiu a doença. Aliás, é por este motivo que é impossível divulgar os tais 7% exatamente – este número depende da probabilidade a priori de cada indivíduo.

Para administradores, este tipo de raciocínio é algo que deve ser levado em conta antes de se decidir por testar ou não membros de uma organização (associado às reações psicológicas dos testados que não sabem a diferença entre $\Pr(+|AIDS)$ e $\Pr(AIDS|+)$; preconceitos que possam estar associados aos resultados dos testes; etc.).

Postcript: O seguinte texto foi retirado do site do Superior Tribunal de Justiça[2]. Textos semelhantes foram publicados em vários jornais do país em Setembro de 2002.

Quinta-feira, 26 de setembro de 2002

09:33 - Fundação Pró-Sangue terá de pagar indenização por erro em exame de HIV

Por causa de diagnóstico errado para HIV positivo, a Fundação Pró-Sangue Hemocentro de São Paulo terá de pagar uma indenização no valor de R$ 40 mil ao torneiro P.G.S.. No entendimento unânime da Terceira Turma do Superior Tribunal de Justiça (STJ), instituição que emite laudo sobre o vírus da Aids sem ressalva quanto à falibilidade do diagnóstico, tem de se responsabilizar se houver uma falha no resultado.

Com isso, os ministros do STJ mantiveram a decisão do Tribunal de Justiça de São Paulo (TJ-SP) que condenou a Fundação Pró-Sangue a pagar a indenização. De acordo com o acórdão do TJ-SP, o laudo feito pelo maior hemocentro da América Latina não trouxe nenhuma ressalva, observação ou advertência de que o resultado deveria ser confirmado para que houvesse certeza do diagnóstico. Para o Tribunal, a falibilidade do teste é de conhecimento notório de pessoas bem informadas. Não seria o caso, entretanto, de P.G.S., "um modesto operário".

Ele propôs ação de indenização por ato ilícito contra a Fundação Pró-Sangue para reparação dos danos causados pela notícia equivocada. Ao doar sangue ao hemocentro em junho de 1996, o torneiro teve de submeter-se ao teste para detectar se era soropositivo. O resultado apontou que ele era portador do vírus da Aids. Inconformado, Paulo Gomes fez outro exame que constou um diagnóstico "indeterminado".

Durante dois meses, o operário viveu um inferno, segundo relatou no processo. Em conseqüência do choque, passou a faltar várias vezes ao trabalho, sobreviveu à base de calmantes e adquiriu gastrite nervosa. O erro do laboratório também teria lhe causado insônia, depressão e ansiedade. Com as reiteradas faltas ao trabalho, o torneiro foi advertido pelo chefe. Ao saber do que se passava, o chefe o aconselhou a procurar um laboratório particular para fazer um novo exame. Foi, então, que ele descobriu a verdade: não era portador do vírus.

A Fundação Pró-Sangue diz que não agiu com imperícia ou negligência. A instituição somente teria efetuado os testes sorológicos, mas não transmitido os resultados. O argumento é de que a coleta do sangue, triagem e os demais contatos teriam sido feitos com o doador no Núcleo de Hematologia de São Caetano do Sul, onde os funcionários deveriam ter orientado o operário sobre a falibilidade do laudo.

Segundo a Fundação, no exame em que constou o resultado positivo foi realizado o "Teste Elisa". Depois, quando o diagnóstico foi "indeterminado" foi aplicado o "Teste Western Blot". A Pró-Sangue explica que todos os métodos sorológicos possuem uma

[2]Em 30/9/2011, o endereço era: HTTP://WWW.STJ.GOV.BR/PORTAL_STJ/PUBLICACAO/ENGINE.WSP?TMP.AREA=368&TMP.TEXTO=70849.

faixa de resultados falsos-positivos. Por isso, são realizados testes para confirmação ou não do resultado inicial.

O pedido da ação de indenização proposta pelo operário foi julgado improcedente na primeira instância. O TJ-SP, entretanto, reverteu a decisão. Tampouco foi acolhido recurso apresentado pela Fundação àquele Tribunal. Foi então que a instituição propôs agravo regimental que foi julgado improcedente pela ministra Nancy Andrighi, em despacho monocrático.

A Fundação Pró-Sangue decidiu apresentar novo recurso (Agravo Regimental em Agravo de Instrumento), que foi apreciado pela Terceira Turma do STJ. Por unanimidade, os ministros negaram provimento sob o argumento de que para mudar a conclusão do TJ-SP seria necessário rever as provas existentes nos autos, o que é vedado ao Superior Tribunal de Justiça, de acordo com o que estabelece a súmula n^o 7/STJ.

Da mesma forma, os ministros negaram pedido de redução do valor da indenização arbitrado pelo TJ-SP. Segundo a relatora, ministra Nancy Andrighi, a quantia estabelecida é razoável. Na inicial, o autor da ação pediu uma indenização de mil salários mínimos, ou seja, R$ 200 mil. Este valor foi considerado excessivo pelos desembargadores de São Paulo.

3.5 Exercícios

Ex. 1 *Joga-se um dado não-viciado duas vezes. Determine a probabilidade condicional de obter* 3 *na primeira jogada sabendo que a soma dos resultados foi* 7.

Ex. 2 *Um estudante resolve um teste de múltipla escolha de* 10 *questões, com* 5 *alternativas por questão. Ele sabe* 60% *da matéria do teste. Quando ele sabe uma questão, ele acerta, e, quando não sabe, escolhe a resposta ao acaso. Se ele acerta uma questão, qual é a probabilidade de que tenha sido por acaso?*

Ex. 3 *Na sua capa de 21/3/2001, a revisa VEJA afirma que "*47% *dos brasileiros não sentem vontade de fazer sexo". Dentro da revista, encontramos que "*35% *das mulheres não sentem nenhuma vontade de ter relações" e "entre os homens, apenas* 12% *se queixam de falta de desejo". Explique porque a informação da capa é incompatível com a do texto da reportagem; que dados você precisaria para estimar corretamente o número da capa?*

Ex. 4 *Na seção anterior, Kuerten e Coria jogavam uma partida de tênis em* 3 *sets. Cada um deles tem* 50% *de chance de vencer cada set (e supõe-se os sets independentes entre si). Considere os eventos $A =$ "Kuerten vence a partida" e $B =$ "o jogo termina em* 2 *sets". Calcule as probabilidades de cada um deles. Eles são independentes? Mutuamente excludentes?*

Ex. 5 *Repita o problema anterior onde Ralph joga contra Kuerten – agora, a probabilidade de Kuerten vencer um set é* 70%, *mas os sets ainda são independentes entre si.*

Ex. 6 *Lança-se um dado* 3 *vezes. Cada vez você tirar* 5 *ou* 6, *você ganha* \$1, *caso contrário, você paga* \$1. *Seja $A =$ "você teve algum lucro ao final do jogo" e $B =$ "você perdeu* \$1 *no primeiro lançamento". Calcule* $\Pr(A)$, $\Pr(B)$, $\Pr(A \text{ e } B)$ *e* $\Pr(A|B)$. *Os eventos A e B são independentes? Mutuamente excludentes?*

Ex. 7 (Excel) *Há n alunos em uma sala de aula. Qual a probabilidade de haver pelo menos um par que faça aniversário no mesmo dia (e mês)? Monte uma planilha mostrando os valores de n e as probabilidades correspondentes para $1 \leq n \leq 366$. Qual valor de n nos dá uma probabilidade de aproximadamente* 50%*?*

Ex. 8 *a) Uma loteria semanal tem* 100 *bilhetes. Quem tem a maior chance de ganhar algum prêmio: quem compra* 10 *bilhetes numa semana ou quem compra* 1 *bilhete por semana durante* 10 *semanas?*
b) Seja $n > 1$*. Encontre o mínimo da função* $f(x) = (1+x)^n - (1+nx)$ *para* $x \in (-1, \infty)$*. Conclua que* $f(x) \geq 0$ *para* $x > -1$*.*
c) Generalize o item (a): suponha $1 < n < N$*. Numa loteria com* N *bilhetes, é melhor comprar* n *numa semana ou* 1 *por dia durante* n *semanas?*

Ex. 9 *Ralph está na FGV* 70% *do horário comercial, enquanto Morgado está na FGV* 20% *do horário comercial. Sabe-se também que, em* 20% *do horário comercial, nenhum dos dois está presente à FGV. Os eventos "Ralph está na FGV" e "Morgado está na FGV" são independentes?*

Ex. 10 *Um dos problemas analisados no século XVII por Pascal e Fermat e que deu origem à Teoria da Probabilidade é o chamado "Problema do Cavalheiro de Méré". Num dos jogos em questão, jogavam-se* 4 *dados e apostava-se que ao menos um* 1 *ocorreria. O cavalheiro (Antoine Gombaud, que propôs o problema a Pascal) argumentava que a probabilidade disto ocorrer seria* $\frac{1}{6}$ *para cada dado, somando um total de* $\frac{4}{6} = \frac{2}{3}$ *nos* 4 *dados. O que está errado com este argumento? Qual é a probabilidade correta?*

Ex. 11 *A outra parte do "Problema do Cavalheiro de Méré" era calcular a probabilidade de conseguir um duplo* 6 *em* 24 *lançamentos de um par de dados. Novamente, o "cavalheiro" propunha que a probabilidade era de* $\frac{1}{36}$ *para um lançamento, então deveria ser* $\frac{24}{36} = \frac{2}{3}$ *para* 24 *lançamentos. Corrija este argumento.*

Ex. 12 *Quantas vezes, no mínimo, se deve lançar um dado para que a probabilidade de obter algum seis seja superior a* 90%*?*

Ex. 13 (Bertrand´s Box) *Você tem à sua frente três caixas; uma delas tem duas bolas brancas, uma outra tem duas bolas pretas e a terceira tem uma bola de cada cor. Você escolhe uma caixa ao acaso e, dela, retira uma bola ao acaso, verificando que ela é branca. Qual a chance de ela ter vindo da caixa com duas bolas brancas?*

Ex. 14 (Monty Hall) *a) Em um programa da televisão, o candidato devem escolher uma dentre três portas. Atrás de uma dessas portas há um prêmio e atrás de cada uma das outras duas portas há um bode. Escolhida uma porta pelo candidato, o apresentador abre uma das outras portas (nota: o apresentador nunca abre a porta do candidato e nunca abre a porta com o prêmio), e pergunta ao candidato se ele quer ficar com a porta que escolheu ou se prefere trocá-la pela outra porta que ainda está fechada. Você acha que o candidato deve trocar, não deve trocar ou que tanto faz?*
b)Agora suponha que os prêmios são um carro, um bode ou uma caixa de sabão em pó ESPUMOSO. O candidato escolhe uma porta ao acaso. O apresentador nunca abre a porta do carro nem a do candidato; no entanto, se as regras acima ainda permitirem, ele abre a do sabão em pó ESPUMOSO, que limpa mais branco, faz mais bolhinhas, e você lava lava lava esfrega esfrega esfrega e hmmmm! que cheirinho de limão!
i) O candidato escolhe uma porta e o apresentador abre a porta do bode. Qual a chance do carro estar na outra porta?
ii) E se a porta aberta pelo apresentador tiver o sabão ESPUMOSO?

Ex. 15 *Um juiz de futebol meio trapalhão tem no bolso um cartão amarelo, um cartão vermelho e um cartão com uma face amarela e uma face vermelha. Depois de uma jogada violenta, o juiz mostra um cartão, retirado do bolso ao acaso, para um atleta. Se a face que o jogador vê é amarela, qual é a probabilidade da face voltada para o juiz ser vermelha?*

Ex. 16 *O Departamento de Justiça dos Estados Unidos reportou que o número de adultos no Estados Unidos sob algum tipo de supervisão judiciária (prisões, casas de detenção ou prisão condicional) chegava a* 6.5 *milhões em* 2000*, dos quais* 3.4 *milhões eram brancos e* 2.15 *milhões eram negros. Outro relatório um pouco anterior dizia que* 9% *da população negra adulta dos Estados Unidos estavam sob algum tipo de supervisão judiciária, comparados com* 2% *da população adulta branca e* 1.3% *das outras raças. Usando estas informações, calcule:*
a) x *tal que "1 em cada* x *adultos nos Estados Unidos esteja sob supervisão judiciária".*
b) A probabilidade de um negro adulto estar sob supervisão. E um branco?
c) A probabilidade de um adulto sob supervisão ser negro. E ser branco?
d) A probabilidade de um adulto que não esteja sob supervisão *ser negro. E branco?*
e) Os eventos "ser branco" e "estar sob supervisão" são independentes?

Ex. 17 *No estudo de caso acima, suponha que um paciente que já testou + para AIDS faz um segundo teste independente do primeiro mas com os mesmos valores de sensitividade e especificidade. Se ele testa + de novo, qual a chance de ter AIDS? [Dica: é como se este indivíduo pertencesse a um grupo onde a prevalência da doença fosse* 6.573%*]*

Ex. 18 *Suponha que disputa de pênaltis é loteria: cada pênalti tem probabilidade* p *de ser gol e eles são independentes entre si. Dois times* S *e* U *disputam uma vaga nas quartas de final de uma copa nos pênaltis. Eles batem* 5 *pênaltis cada, alternadamente, e a disputa termina assim que um dos times não tiver mais como alcançar o outro (por exemplo, se* U *começa batendo e está* 2×0 *para* U*, então* S *nem bate seu último pênalti).*
a) Supondo que S *começa batendo, qual a probabilidade de a disputa terminar* 3×0 *para* U*?*
b) Supondo que U *começa batendo, qual a probabilidade de a disputa terminar* 3×0 *para* U*?*
c) Supondo que uma moeda justa decide quem bate primeiro, qual a chance de terminar 3×0 *para* U*? Com o auxílio de uma calculadora, mostre que, sob estas hipóteses, esta chance não ultrapassa* 2.8%*.*

Ex. 19 *Três eventos* A*,* B *e* C *são ditos independentes quando são independentes dois a dois e, além disto,* $\Pr(A \cap B \cap C) = \Pr(A)\Pr(B)\Pr(C)$*. Jogue um dado duas vezes. Sejam* $A = \{$*primeiro número é par*$\}$*,* $B = \{$*segundo número é par*$\}$ *e* $C = \{$*a soma dos números é par*$\}$*. Pergunta-se:*
a) A *e* B *são independentes?*
b) A *e* C *são independentes?*
c) B *e* C *são independentes?*
d) A*,* B *e* C *são independentes?*

Ex. 20 *Mostre que*

$$\Pr(A|B) = \Pr(A) \Rightarrow \Pr\left(A|\bar{B}\right) = \Pr(A)$$

isto é, se A *e* B *são independentes, então* A *e* $\bar{B}$ *também são independentes.*

Ex. 21 *Se* A *joga uma moeda honesta* $n+1$ *vezes e* B *joga* n *vezes, determine a probabilidade de* A *obter mais caras do que* B*.*

Ex. 22 *Sejam* A *e* B *eventos não-impossíveis. Vimos que* A *é* ***independente*** *de* B *quando* $\Pr(B|A) = \Pr(B)$*. Dizemos que* A ***atrai*** B *(denotado* $A \uparrow B$*) quando* $\Pr(B|A) > \Pr(B)$ *e que* A ***repele*** B *(denotado* $A \downarrow B$*) quando* $\Pr(B|A) < \Pr(B)$*.*
a) Mostre que, se $0 < \Pr(A) < 1$*, então* $A \uparrow A$*.*
b) Mostre que

$$\begin{aligned} (A \uparrow B) &\Leftrightarrow (B \uparrow A) \\ (A \downarrow B) &\Leftrightarrow (B \downarrow A) \end{aligned}$$

ou seja, podemos dizer que A e B ***se atraem*** *(ou* ***repelem****)* ***mutuamente.***
c) Intuitivamente, quais pares de eventos abaixo se atraem e quais se repelem?
i) Time A ser campeão e time B (diferente de A) ser campeão.
ii) Time A ser rebaixado e time B (diferente de A) ser rebaixado.
iii) Irmão ter gripe e irmã, na mesma casa, ter gripe.
iv) Irmão ter olhos azuis e irmã ter olhos azuis.
v) Muitos sorvetes serem vendidos num dia e haver muitos afogamentos no mesmo dia.
d) Mostre que as seguintes afirmações ***não são necessariamente verdadeiras****:*

$$\begin{array}{lcl}(A\uparrow B)\ e\ (B\uparrow C) & \Rightarrow & (A\uparrow C)\\ (A\uparrow B)\ e\ (B\uparrow C) & \Rightarrow & ((A\cap C)\uparrow B)\\ (A\downarrow B)\ e\ (B\downarrow C) & \Rightarrow & (A\downarrow C)\\ (A\downarrow B)\ e\ (B\downarrow C) & \Rightarrow & ((A\cap C)\downarrow B)\end{array}$$

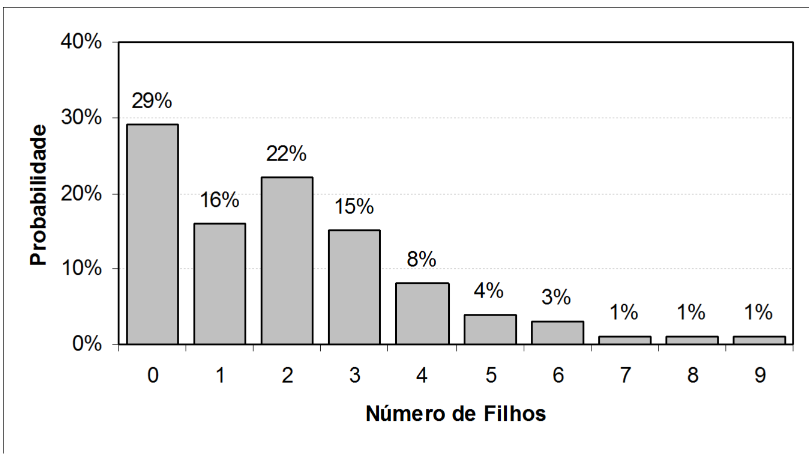
(Intuição do último item: eu não gosto da minha irmã nem no namorado dela. Eu raramente saio com ela, e quase nunca com o namorado dela – é mais provável eu sair sozinho! Mas, nas raríssimas ocasiões em que os dois saem, meus pais me forçam a ir com eles, então ***certamente*** *eu vou. Você consegue formalizar esta idéia com probabilidades?)*

Ex. 23 (A1 2004.2) *Um dado honesto tem duas de suas faces pintadas de vermelho e as demais de azul. O dado é lançado três vezes, anotando-se a cor da face obtida.*
a) Qual é a probabilidade de que a cor obtida no 1o. lançamento seja igual à obtida no 3o?
b) Dado que a mesma cor foi obtida no 1o e 2o lançamentos, qual é a probabilidade de que no 3o lançamento saia esta mesma cor?

Ex. 24 (A1 2004.2)

A figura abaixo mostra a probabilidade de que uma pessoa, escolhida ao acaso em um grupo de mulheres com idades de 25 a 35 anos, tenha um certo número de filhos, ou seja,

$$\text{Probabilidade} = \frac{\text{Número de mulheres com um certo número de filhos}}{\text{Número total de mulheres na amostra}}$$



a) Se uma mulher é escolhida ao acaso neste grupo, é mais provável que ela tenha quantos filhos? Qual é a probabilidade correspondente?
b) Se uma mãe é escolhida ao acaso neste grupo, é mais provável que ela tenha quantos filhos? Qual é a probabilidade correspondente?
c) Suponhamos que, dentre todos os filhos das mulheres da amostra, um seja escolhido ao acaso. Qual é a probabilidade de que ele seja filho único?

Ex. 25 (AS 2004.2) *Os alunos de um certo período de uma faculdade fazem 5 matérias. As provas finais serão marcadas para uma única semana (de segunda a sábado). Admitindo que cada professor escolha ao acaso e independentemente dos demais a data de sua prova, qual é a probabilidade:*
a) de que não haja provas no sábado?
b) de que os alunos não façam mais de uma prova por dia?
c) de que haja algum dia com 4 ou mais provas?

Ex. 26 (AS 2005.2) *A probabilidade do tenista Berrando Gemigemi vencer um set contra Maria Xaropova é de* 70%*. Eles disputam uma partida de* 3 *sets (suponha que os sets são independentes uns dos outros).*
a) Qual a probabilidade do Berrando vencer a partida?
b) Qual a probabilidade de a partida terminar em 2 *sets?*
c) Qual a probabilidade de Berrando vencer em 2 *sets?*
d) Qual a probabilidade de uma partida vencida pelo Berrando ter terminado em 2 *sets?*

Ex. 27 (T1 2006.1) *Segundo uma pesquisa de opinião do IBOPE*[3] *de 2003, 2% dos brasileiros torciam para o Botafogo, 15% torciam para o Flamengo, 2% para o Fluminense e 5% para o Vasco. Restringindo a população a apenas brasileiros que tivessem grau superior (que eram apenas 8% do total), as porcentagens mudavam para 4%, 10%, 2% e 8%, respectivamente. Escolha um brasileiro da amostra total do IBOPE ao acaso.*
a) Qual a chance de ele não torcer para nenhum dos quatro grandes clubes cariocas?
b) Qual a chance de ele ser um flamenguista com grau superior?
c) Que porcentagem dos torcedores do Fluminense tem grau superior? E do Botafogo?
d) Se um brasileiro não tem grau superior, qual a chance de ele ser flamenguista?
e) De acordo com estes dados, os eventos "torcer para o Fluminense" e "ter grau superior" são independentes? Mutuamente excludentes?

Ex. 28 (AS 2006.1) *O assassino é um dos* 100 *mil habitantes adultos de uma cidade, então, na falta de qualquer outra evidência, é razoável supor que o senhor Simpson, morador desta cidade, tenha probabilidade* $1/10^5$ *de ser o assassino. No entanto, há um teste de DNA; se Simpson não é o assassino, a chance do teste afirmar que os DNAs de ambos são iguais é de apenas uma em* 10 *mil; se Simpson é o assassino, o teste afirmará que os DNAs de ambos são iguais com certeza.*
O resultado acaba de voltar do laboratório: o teste afirma que os DNAs de Simpson e do assassino são iguais. Baseado apenas nesta evidência, qual a chance de Simpson ser o assassino?

3.6 Respostas dos Exercícios

Resp. 1 $\frac{1}{6}$

Resp. 2 *Árvore. Dá* $\frac{8}{68}$.

Resp. 3 *Tá errado pra caramba. Se fossem* 60% *das mulheres e* 55% *dos homens, seriam* 115% *dos brasileiros? Não se somam laranjas com bananas assim! Faríamos o problema com uma média ponderada de* 12% *e* 35%*, ponderada pela quantidade de homens e mulheres no Brasil. Se for meio a meio, então seria*

$$\frac{12+35}{2}\% = 23.5\%$$

dos brasileiros.

Resp. 4 *Ambas as probabilidades são* 50%*. São independentes, mas não são mutuamente excludentes.*

[3] http://www.ibope.com.br/opp/pesquisa/opiniaopublica/download/imprensa_torcidas_1_mencao.pdf

Resp. 5 *Agora as probabilidades são 78.4% para Kuerten vencer o jogo; 58% de acabar em dois sets. E, sabido que acabou em dois sets, Kuerten sobe para* $\frac{49}{58} = 84.483\%$ *de chance de vencer, então estes eventos não são independentes. Também não são mutuamente excludentes – Kuerten pode vencer 2 a 0.*

Resp. 6 $\Pr(A) = \frac{7}{27}$; $\Pr(B) = \frac{2}{3}$; $\Pr(A \text{ e } B) = \frac{2}{27}$; $\Pr(A|B) = \frac{2/27}{2/3} = \frac{1}{9}$. *Não são independentes, nem excludentes.*

Resp. 7 *Desenhe uma árvore, pense na probabilidade de NÃO haver par algum de aniversário repetidos, depois use a lei do complemento. A probabilidade de HAVER uma coincidência é:*

$$1 - \frac{365.364.363....(365 - n + 1)}{365^n}$$

Surpreendentemente, $n = 23$ *já dá mais de 50% de chance.*

Resp. 8 *a) Jogando tudo duma vez,* $\Pr(\textit{prêmio}) = \frac{10}{100} = 10\%$. *Jogando uma vez por semana, a chance de não ganhar nada é* $\left(\frac{99}{100}\right)^{10} = 90.4382\%$, *então a chance de ganhar alguma coisa é*

$$1 - \left(\frac{99}{100}\right)^{10} = 9.5618\% < 10\%$$

Se você só quer ganhar ALGUMA coisa, melhor jogar tudo de uma vez.
b) Use cálculo, encontre o mínimo de $f(x)$ *em* $(-1, \infty)$, *que será* $f(0) = 0$.
c) Agora as probabilidades de ganhar são

$$\begin{aligned} \textit{Jogando tudo de uma vez} &: \frac{n}{N} \\ \textit{Jogando em n sorteios} &: 1 - \left(1 - \frac{1}{N}\right)^n \end{aligned}$$

Mas, tomando $x = -\frac{1}{N}$ *no item (b), conclui-se que*

$$f\left(-\frac{1}{N}\right) = \left(1 - \frac{1}{N}\right)^n - \left(1 - \frac{n}{N}\right) \geq 0$$

isto é

$$1 - \left(1 - \frac{1}{N}\right)^n \leq \frac{n}{N}$$

Melhor jogar tudo de uma vez!

Resp. 9 *Tabela começa assim:*

	R	$\bar{R}$	
M			0.2
$\bar{M}$		0.2	
	0.7		1.0

Complete a la Sudoku:

	R	$\bar{R}$	
M	0.1	0.1	0.2
$\bar{M}$	0.6	0.2	0.8
	0.7	0.3	1.0

Como $\Pr(\textit{Ralph}) = 0.7 \neq \Pr(\textit{Ralph}|\textit{Morgado}) = \frac{0.1}{0.2} = 0.5$ *os eventos não são independentes.*

Resp. 10 *Não se somam probabilidades de eventos que não são mutuamente excludentes! A resposta correta é*

$$1-\left(\frac{5}{6}\right)^4=\frac{671}{1296}=51.77\%$$

Resp. 11 *Agora, a probabilidade é*

$$1-\left(\frac{35}{36}\right)^{24}=49.14\%$$

Resp. 12 *Seja n o número de lançamentos. A probabilidade de não obter um 6 é*

$$\left(\frac{5}{6}\right)^n$$

Queremos

$$\left(\frac{5}{6}\right)^n\leq 0.1\Rightarrow n\geq\frac{\ln 0.1}{\ln\left(\frac{5}{6}\right)}=12.629$$

Então $n=13$ lançamentos serve.

Resp. 13 $\frac{2}{3}$

Resp. 14 *Troque, pois a chance da outra porta ter o prêmio é $\frac{2}{3}$.*

Resp. 15 $\frac{2}{3}$ *de novo (é igual ao da caixa acima).*

Resp. 16 *Tabela em milhões:*

	Brancos	*Negros*	*Outros*	*Total*
Supervisão	3.4	2.15	0.95	6.5
Livre	$170-3.4=166.6$	$23.89-2.15=21.74$	$73.08-0.95=72.13$	260.47
Total	$\frac{3.4}{0.02}=170$	$\frac{2.15}{0.09}=23.89$	$\frac{0.95}{0.013}=73.08$	266.97

a) $\frac{6.5}{266.97}=\frac{1}{x}\Rightarrow x=\frac{266.97}{6.5}=41.072$
b) Dados do problema! São 9% e 2%, respectivamente.
c) $\frac{2.15}{6.5}=33.08\%$ *para negros e* $\frac{3.4}{6.5}=52.31\%$ *para brancos*
d) $\frac{21.74}{260.47}=8.35\%$ *e* $\frac{166.6}{260.47}=63.96\%$
e) Não. $\Pr(\mathit{Supervisão}|\mathit{Branco})=2\%<\frac{6.5}{266.97}=2.434\%=\Pr(\mathit{Supervisão})$

Resp. 17 *Dá 49% e uns quebrados.*

Resp. 18 *Seja $q=1-p$ a chance de um pênalti não ser gol. Então:*
a) $\Pr(U3\times 0S\mid S\ \mathit{começa})=qpqpqp=p^3q^3$ *(é o único jeito; se chegássemos ao quarto pênalti de S e S errasse então, terminaria no máximo 2×0).*
b) $\Pr(U3\times 0S\mid U\ \mathit{começa})=pqpqpq+3(qqpqpqp)=p^3q^3+3p^3q^4$ *(a segunda parcela corresponde aos casos em que U erra um dos 3 primeiros pênaltis e só confirma sua vitória no quarto pênalti batido).*
c)

$$\begin{aligned}f(p) &= \Pr(U3\times 0S)= \\ &= \Pr(S\ \mathit{começa}).\Pr(U3\times 0S\mid S\ \mathit{começa})+\Pr(U\ \mathit{começa})\Pr(U3\times 0S\mid U\ \mathit{começa}) \\ &= \frac{1}{2}\left(2p^3q^3+3p^3q^4\right)=\frac{p^3q^3(2+3q)}{2}=\frac{p^3(1-p)^3(5-3p)}{2}\end{aligned}$$

Derivando com relação a p e igualando a 0 para encontrar o máximo em $p \in [0,1]$:

$$\begin{aligned}\frac{df}{dp} &= \frac{p^2(1-p)^2}{2}\left(3(1-p)(5-3p) - 3p(5-3p) - 3p(1-p)\right) = \\ &= \frac{3}{2}p^2(1-p)^2\left(7p^2 - 14p + 5\right) = 0 \Rightarrow \\ &\Rightarrow p = 0 \text{ ou } p = 1 \text{ ou } p = \frac{7 \pm \sqrt{14}}{7} = 1 \pm \frac{\sqrt{14}}{7}\end{aligned}$$

Como $f(0) = f(1) = 0$ e devemos ter $p \in [0,1]$, temos o máximo local em

$$p^* = 1 - \frac{\sqrt{14}}{7} \approx 0.4655$$

Substituindo de volta em f, temos uma chance máxima de

$$f(p^*) = \frac{\left(1 - \frac{\sqrt{14}}{7}\right)^3 \left(\frac{\sqrt{14}}{7}\right)^3 \left(2 + 3\frac{\sqrt{14}}{7}\right)}{2} = \frac{2}{7^5}\left(7 - \sqrt{14}\right)^3\left(3 + \sqrt{14}\right) \approx 2.775\%$$

Resp. 19 *Sim, sim, sim e não.*

Resp. 20

$$\begin{aligned}\Pr(A|B) &= \Pr(A) \Rightarrow \Pr(B|A) = \Pr(B) \Rightarrow 1 - \Pr\left(\bar{B}|A\right) = 1 - \Pr\left(\bar{B}\right) \Rightarrow \\ &\Rightarrow \Pr\left(\bar{B}|A\right) = \Pr\left(\bar{B}\right) \Rightarrow \Pr\left(A|\bar{B}\right) = \Pr(A)\end{aligned}$$

Resp. 21 *Difícil este, leia com MUITO cuidado: eu jogo $n+1$ moedas, você joga n. Note que eu tenho mais coroas* ***ou*** *mais caras do que você (pois eu tenho mais moedas), mas não ambos (porque eu só tenho UMA moeda a mais, não dá para eu ganhar em caras e em coroas também). Por simetria, a chance de eu ter mais caras é igual à chance de eu ter mais coroas. Assim, a probabilidade é $\frac{1}{2}$.*

Resp. 22 *a) Sim*

$$A \downarrow B \Rightarrow \Pr(B|A) < \Pr(B) \Rightarrow \Pr(A \text{ e } B) < \Pr(A)\Pr(B) \Rightarrow \Pr(A|B) < \Pr(A) \Rightarrow B \downarrow A$$

b) Falso. Por exemplo, se $A = C$, está obviamente errado.
c) Falso. Por exemplo, se $\Pr(ABC) = 10\%$ e $\Pr\left(A\bar{B}\bar{C}\right) = \Pr\left(\bar{A}B\bar{C}\right) = \Pr\left(\bar{A}\bar{B}C\right) = 30\%$, note que A repele B e que C repele B, mas A e C juntos ATRAEM B.

Resp. 23 $S = \{VVV, VVA, VAV, AVV, AAV, AVA, VAA, AAA\}$ *com probabilidades respectivamente de $\frac{1}{27}$, $\frac{2}{27}$, $\frac{2}{27}$, $\frac{2}{27}$, $\frac{4}{27}$, $\frac{4}{27}$, $\frac{4}{27}$ e $\frac{8}{27}$.*
a) $\Pr(X?X) = \frac{1+2+4+8}{27} = \frac{15}{27} = \frac{5}{9}$
b) $\Pr(XXX|XX?) = \frac{1+8}{1+2+4+8} = \frac{9}{15} = \frac{3}{5}$

Resp. 24 *a) Mais provável de ter 0 filhos, com 29% de chance.*
b) Mais provável é 1 filho, com $\frac{0.22}{1-0.29} = \frac{22}{71} = 30.99\%$ de chance.
c) Supondo uma amostra fictícia de 100 mulheres, são

$$(29)(0) + (16)(1) + (22)(2) + (15)(3) + (8)(4) + (4)(5) + (3)(6) + 7 + 8 + 9 = 199 \text{ filhos}$$

A chance de ser filho único é $\frac{16}{199} = 8.04\%$

Resp. 25 *a)* $\left(\frac{5}{6}\right)^5 = 40.19\%$
b) $\frac{5}{6}\frac{4}{6}\frac{3}{6}\frac{2}{6} = 9.26\%$
c) 5 *provas no mesmo dia com chance* $\left(\frac{1}{6}\right)^4 = \frac{1}{1296}$
4 *provas num dia e uma no outro: há* 5 *opções para a "outra prova". As outras* 4 *caem juntas com* $\left(\frac{1}{6}\right)^3$ *de chance, e sobram* $\frac{5}{6}$ *para a prova singular. Total:*

$$5\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)^3 = \frac{25}{1296}$$

Total: $\frac{26}{1296} = 2.01\%$

Resp. 26 *Gemigemi é Kuerten e Xarapova é Ralph do problema acima. Então:*
a) $\Pr(\textit{Berrando vencer}) = 78.4\%$
b) $\Pr(\textit{2 sets}) = 58\%$
c) $\Pr(\textit{Berrando 2 a 0}) = 49\%$
d) $\Pr(\textit{2 a 0} \mid \textit{Berrando venceu}) = \frac{49}{78.4} = 62.5\%$

Resp. 27 *Montamos uma tabela novamente; para facilitar, supomos um total fictício de* 2000 *brasileiros na amostra entrevistada. Com os dados do problema, já temos:*

	Bota	*Fla*	*Flu*	*Vasco*	*Outros*	*Total*
Com Grau Superior	$4\%(160) = 6.4$	$10\%(160) = 16$	$2\%(160) = 3.2$	$8\%(160) = 12.8$		160
Sem Grau Superior						
	40	300	40	100		2000

Agora é só completar a tabela usando somas e subtrações:

	Bota	*Fla*	*Flu*	*Vasco*	*Outros*	*Total*
Com Grau Superior	6.4	16.0	3.2	12.8		160
Sem Grau Superior	33.6	284.0	36.8	87.2		1840
	40.0	300.0	40.0	100.0	1520.0	2000

a) A chance de ele não torcer pelos grandes cariocas é de $\frac{1520}{2000} = 76\%$.
b) A chance de ele ser um flamenguista com grau superior é de $\frac{16}{2000} = 0.8\%$.
c) $\frac{3.2}{40} = 8\%$ *dos tricolores têm grau superior;* $\frac{6.4}{40} = 16\%$ *dos botafoguenses têm grau superior.*
d) $\frac{284}{1840} = 15.43\%$ *dos brasileiros sem grau superior torcem para o flamengo.*
e) De acordo com estes dados, "ser tricolor" e "ter grau superior" são surpreendentemente independentes. Direto do enunciado, $\Pr(Flu) = 2\%$ *e* $\Pr(Flu|Superior) = 2\%$ *também! Mutuamente excludentes eles não são – há tricolores com grau superior.*
Comentário: de fato, a pesquisa do IBOPE entrevistou 2000 brasileiros e as porcentagens apresentadas estão levemente arredondadas. As porcentagens gerais tem um "erro de 2.2 pontos percentuais" – razoável para determinar a maior torcida do Brasil, mas não razoável para determinar se Botafogo ou Fluminense têm mais torcedores. Dentre os de grau superior, eram apenas 160 entrevistados, e os erros possíveis ao tentar usar esta amostra como uma representação do Brasil aumentam muito (ali o erro é de 15 pontos percentuais!).

Resp. 28 *Monte uma árvore ou uma tabela. Seja* $A =$ *"Simpson é o assassino" e* $+ =$ *"Teste afirma que Simpson e assassino tem o mesmo DNA". Os dados do problema são*

$$\Pr(A) = \frac{1}{10^5}$$
$$\Pr(+ \mid A) = 1$$
$$\Pr\left(- \mid \bar{A}\right) = \frac{1}{10^4}$$

Assim, a tabela de probabilidades para o Sr. Simpson é (usando que $\Pr(+ \text{ e } A) = \Pr(A)\Pr(+|A)$, *por exemplo):*

	A	$\bar{A}$	*Total*
+	$\frac{1}{100\,000}$	$\frac{99\,999}{100\,000}\frac{1}{10\,000}$	$\frac{1}{100000}\left(1+\frac{99999}{10000}\right)$
-	0	$\frac{99\,999}{100\,000}\frac{9999}{10\,000}$	*Blah*
Total	$\frac{1}{100\,000}$	$\frac{99\,999}{100\,000}$	1

Enfim, o que se deseja calcular é

$$\Pr(A \mid +) = \frac{\frac{1}{100\,000}}{\frac{1}{100000}\left(1+\frac{99999}{10000}\right)} = \frac{10\,000}{109\,999} \approx 9.091\%$$