Como perder amigos e enganar pessoas

Nicolau C. Saldanha

21 de janeiro de 1998

Neste artigo apresentaremos quatro situações simples em que probabilidades enganam. Em alguns casos a probabilidade de certos eventos tem um valor diferente daquele que a maioria das pessoas parece julgar razoável, pelo menos de início; em um exemplo mostraremos como é fácil chegar a conclusões absurdas. Para que o leitor possa pensar sozinho apresentaremos primeiro quatro "enunciados", em que apresentamos cada situação e depois quatro "desenvolvimentos" onde voltamos a discutir as quatro situações na mesma ordem. Qualquer um pode usar estes exemplos para divertir-se às custas de seus amigos mas em nenhum caso o autor tem responsabilidade pela integridade física daqueles que usarem a matemática para o mal.

Enunciados

1. Em um programa de auditório, o convidado deve escolher uma dentre três portas. Atrás de uma das portas há um carro e atrás de cada uma das outras duas há um bode. O convidado ganhará como prêmio o que estiver atrás da porta; devemos supor neste problema que o convidado prefere ganhar o carro. O procedimento para escolha da porta é o seguinte: o convidado escolhe inicialmente, em caráter provisório, uma das três portas. O apresentador do programa, que sabe o que há atrás de cada porta, abre neste momento uma das outras duas portas, sempre revelando um dos dois bodes. O convidado agora tem a opção de ficar com a primeira porta que ele escolheu ou trocar pela outra porta fechada. Que estratégia deve o convidado adotar? Com uma boa estratégia, que probabilidade tem o convidado de ganhar o carro?

2. Um móvel tem três gavetas iguais. Em uma gaveta há duas bolas brancas, em outra há duas bolas pretas e na terceira há uma bola branca e outra preta. Abrimos uma gaveta ao acaso e tiramos uma bola ao acaso sem olhar a segunda bola que está na gaveta. A bola que tiramos é branca. Qual é a probabilidade de que a segunda bola que ficou sozinha na gaveta seja também branca?

3. Dois amigos querem decidir quem pagará a conta do restaurante com uma aposta. Cada um deles escolhe uma seqüência de três caras ou coroas e eles jogam uma moeda até que saia uma das duas seqüências: aquele que tiver escolhido a primeira seqüência a sair ganhou a aposta. Por exemplo, André (por ser o primeiro em ordem alfabética) é o primeiro a escolher e fica com a seqüência **ckc** (onde **c** representa cara e **k** coroa) enquanto Bernardo responde com **cck**. Eles jogam a moeda obtendo **kckkckkkkccck** e neste momento Bernardo declara-se o vencedor. Esta aposta é justa? André leva vantagem ou desvantagem por ser o primeiro a escolher? Quais são as probabilidades de vitória de cada um?

4. Aqui novamente devemos nos imaginar em um programa de auditório. Eugênio foi sorteado e tem direito a um prêmio mas ele deve escolher entre dois envelopes lacrados aparentemente iguais. O apresentador informa que cada envelope tem um cheque e que o valor de um cheque é o dobro do outro, mas não diz nada sobre o valor dos cheques nem indica qual envelope contem o cheque de maior valor. Eugênio escolhe e abre um envelope que contem um cheque de, digamos, R$ 100. Neste momento o apresentador sempre faz uma proposta ao convidado: ele pode trocar de envelope mediante uma "multa" de 5% do valor do cheque que ele tem em mãos—no caso, R$ 5. Assim, se Eugênio aceitar ele pode ganhar R$ 45 (se o cheque no segundo envelope for de R$ 50) ou R$ 195 (se o outro cheque for de R$ 200). Suponhamos que Eugênio (que fez um curso de Introdução à Probabilidade no período anterior) queira maximizar o valor esperado de seu prêmio: deve ele aceitar a troca? E se o valor do primeiro cheque tivesse sido outro, de que forma deveria isto influenciar a decisão de Eugênio? Se Eugênio trocar de envelope independentemente do valor do cheque não vale mais a pena para ele trocar de envelope antes de abrir, evitando assim a multa?

Desenvolvimentos

1. A resposta correta é que trocando de porta, a probabilidade de ganhar o carro é 2/3 enquanto não trocando a probabilidade é apenas 1/3. Uma forma simples de ver isto é a seguinte: trocando de porta, o convidado ganha desde que a primeira porta que ele escolher esconda um dos dois bodes, como pode-se facilmente perceber. A melhor estratégia para o convidado é portanto trocar sempre e assim sua probabilidade de ganhar fica sendo 2/3.

O erro comum aqui é achar que após a eliminação de uma porta (que foi aberta pelo apresentador, revelando um bode) há uma simetria entre as duas outras portas e a probabilidade de cada uma esconder o carro é 1/2. Não existe entretanto tal simetria pois a porta escolhida pelo convidado não poderia, pelas regras, ser tocada pelo apresentador enquanto a outra *poderia* ter sido aberta, mas não foi.

Este processo de fato era seguido em um programa nos Estados Unidos. Uma longa e áspera discussão ocorreu na imprensa quanto a qual era o valor correto da probabilidade e pessoas que deveriam ser capazes de resolver um problema trivial como este passaram pela vergonha de publicar soluções erradas. Julgamos melhor esquecer os detalhes deste episódio deprimente.

2. A resposta correta é 2/3 (e não 1/2). As seis bolas seriam de início igualmente prováveis mas sabemos que a primeira bola escolhida foi branca: assim as três bolas brancas têm igual probabilidade. Estamos interessados em saber a cor da companheira de gaveta de cada bola branca: em dois casos é branca, em um caso é preta. Assim, a probabilidade de que a segunda bola seja branca é 2/3, como já afirmamos.

Um raciocínio comum mas *errado* é dizer: as gavetas são igualmente prováveis mas obviamente não escolhemos a gaveta que contem duas bolas pretas. Portanto teríamos probabilidade 1/2 de termos escolhido a gaveta com duas bolas brancas e 1/2 de termos escolhido a gaveta com uma bola de cada cor; no primeiro caso a segunda bola é branca e no segundo caso a bola é preta. Assim, a resposta seria 1/2.

O que há de errado neste raciocínio? O erro está em dizer que as duas gavetas possíveis são igualmente prováveis. Inicialmente a probabilidade de cada gaveta é de fato a mesma (inclusive para a gaveta com duas bolas pretas) mas ao tirarmos uma bola e constatarmos que ela é branca isto deixa de ser verdade. Isto é bem óbvio para a gaveta com duas bolas pretas: passou a ser impossível termos escolhido esta gaveta. Entre as duas outras gavetas, entretanto, há uma diferença que está sendo ignorada no raciocínio do parágrafo anterior. Se pré-escolhermos a gaveta com duas bolas brancas temos certeza de passar no teste: uma bola escolhida ao acaso nesta gaveta será sempre branca. Por outro lado, se pré-escolhermos a gaveta com uma bola de cada cor ainda temos probabilidade 1/2 de sacarmos uma bola preta, o que estaria em contradição com o enunciado. Assim, a probabilidade de termos escolhido cada uma destas duas gavetas é 2/3 e 1/3, respectivamente. Podemos a partir deste ponto facilmente rededuzir a resposta correta de 2/3.

É fato empírico desencorajador que muitas pessoas teimam em dizer que a probabilidade é 1/2 mesmo após esta explicação. O seguinte exemplo serve como exercício para aqueles que entenderam a explicação e é uma espécie de redução ao absurdo do raciocínio "rival". Temos novamente três gavetas, uma com vinte bolas brancas, uma com vinte bolas pretas e a terceira com dez bolas de cada cor. Abrimos uma gaveta e, sem olhar, retiramos ao acaso dez bolas: elas são todas brancas. Qual a probabilidade de que as dez bolas restantes sejam também brancas?

3. No nosso exemplo, Bernardo tinha probabilidade 2/3 de ganhar. Em geral, o segundo a jogar leva uma vantagem considerável e se escolher bem sua resposta pode garantir uma probabilidade de vitória de pelo menos 2/3, mas às vezes até 7/8, dependendo da primeira jogada. A Tabela 1 dá a probabilidade de vitória de Bernardo para cada par de jogadas (a coluna é a jogada de André e a linha a de Bernardo).

	ccc	**cck**	**ckc**	**ckk**	**kcc**	**kck**	**kkc**	**kkk**
ccc	—	$\frac{1}{2}$	$\frac{2}{5}$	$\frac{2}{5}$	$\frac{1}{8}$	$\frac{5}{12}$	$\frac{3}{10}$	$\frac{1}{2}$
cck	$\frac{1}{2}$	—	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{4}$	$\frac{5}{8}$	$\frac{1}{2}$	$\frac{7}{10}$
ckc	$\frac{3}{5}$	$\frac{1}{3}$	—	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{7}{12}$
ckk	$\frac{3}{5}$	$\frac{1}{3}$	$\frac{1}{2}$	—	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$
kcc	$\frac{7}{8}$	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	—	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{3}{5}$
kck	$\frac{7}{12}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	—	$\frac{1}{3}$	$\frac{3}{5}$
kkc	$\frac{7}{10}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{1}{4}$	$\frac{2}{3}$	$\frac{2}{3}$	—	$\frac{1}{2}$
kkk	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{5}{12}$	$\frac{1}{8}$	$\frac{2}{5}$	$\frac{2}{5}$	$\frac{1}{2}$	—

Tabela 1

Não reconstruiremos aqui toda a tabela: apresentaremos apenas como exemplo a situação descrita no enunciado. O leitor que estiver interessado em aprender mais sobre este problema pode consultar nosso *Precisa-se de alguém para ganhar muito dinheiro*, a ser publicado na *Revista do Professor de Matemática* do Chile mas já disponível na home page do autor: http://www.mat.puc-rio.br/∼nicolau/. O Diagrama 2 descreve bem a situacao. Os seis vertices indicam as seis situações possíveis durante o processo de jogar a moeda. O ponto indica que nenhum jogador tem como esperar fazer uso das jogadas já feitas, ou seja, ou nenhum lance ainda foi feito, ou foi lançado apenas um **k**, ou os dois últimos lances foram **kk**; como o jogo sempre começa nesta situação, chamaremos este vértice de *inicial*. O **c** indica que o último lance foi um **c** mas o anterior ou não existiu ou foi um **k**. Os vértices **cc** e **ck** indicam que estes foram os dois últimos lances. Finalmente, os vértices **cck** e **ckc** indicam que o jogo terminou; chamaremos estes vértices de *finais*.

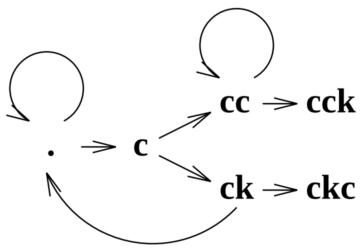

Diagrama 2

As duas setas partindo de cada vértice (exceto os finais) indicam como a situação se modifica a cada lance de moeda: elas correspondem às possibilidades de tirar **c** ou **k** em um dado momento. Queremos agora calcular a probabilidade de vitória de Bernardo dado que o jogo chegou a uma certa situação. Temos assim quatro probabilidades a serem calculadas: $p_{.}, p_{\mathbf{c}}, p_{\mathbf{cc}}$ e $p_{\mathbf{ck}}$; consideramos naturalmente $p_{\mathbf{cck}} = 1$ e $p_{\mathbf{ckc}} = 0$. Como a partir de cada vértice não final as probabilidades associadas às duas setas são iguais, temos as seguintes equações:

$$\begin{aligned}
p_{.} &= \frac{1}{2}(p_{.} + p_{\mathbf{c}})\\
p_{\mathbf{c}} &= \frac{1}{2}(p_{\mathbf{cc}} + p_{\mathbf{ck}})\\
p_{\mathbf{cc}} &= \frac{1}{2}(p_{\mathbf{cc}} + 1)\\
p_{\mathbf{ck}} &= \frac{1}{2}p_{.}.
\end{aligned}$$

Resolvendo o sistema, temos $p_{.} = 2/3$, conforme afirmamos.

O erro mais natural aqui é achar que todas as seqüências são igualmente boas: isto não é verdade pois os dois últimos lances em geral serviram, sem sucesso, para tentar finalizar as seqüências e servirão agora para tentar iniciá-las. Mais surpreendente ainda é o fato que o segundo jogador *sempre* tem uma boa resposta: este jogo é um pouco como jogar par-ou-ímpar ou pedra-papel-tesoura com um dos jogadores tendo o direito de escolher sua jogada só *depois* de ver a jogada do adversário.

4. Antes de mais nada gostaríamos de lembrar que Eugênio deseja, *por hipótese*, maximizar o valor esperado do prêmio. Este critério é razoável em algumas situações e em outras não. Outro convidado poderia precisar desesperadamente de uma certa quantia, talvez R$ 100, e gostaria portanto de maximizar a probabilidade de ganhar pelo menos este valor crítico. Ainda outro convidado pode ser tão curioso que deseja saber quanto há em cada envelope mais do maximizar seu prêmio. O leitor, se fosse o convidado, talvez julgasse interessante considerar ainda outros aspectos. Podemos imaginar inúmeros critérios diferentes e em princípio cada critério gera um novo problema. Nós nos propomos aqui a estudar o problema na forma em que foi proposto e não a discutir se Eugênio, com sua opção pelo valor esperado, é um homem verdadeiramente sábio.

Neste problema, ao contrário dos outros, apresentaremos inicialmente um raciocínio falho e vamos seguí-lo até chegarmos a um absurdo deixando a análise dos erros deste raciocínio para o final. Para tornar a discussão toda mais viva acompanharemos o pensamento de Eugênio.

Ao receber a proposta de troca, Eugênio pensa: *Se ficar com este cheque meu prêmio será de R$ 100. Se trocar de cheque tenho probabilidade* $1/2$ *de ganhar R$ 45 e probabilidade* $1/2$ *de ganhar R$ 195: o valor esperado é de* $(1/2) \cdot 45 + (1/2) \cdot 195 = 120$ *reais. Como 120 é maior que 100, a troca é vantajosa.* Eugênio troca de cheque e fica felicíssimo ao ver que o outro cheque é de R$ 200: ele ganhou R$ 195!

Ao voltar para seu lugar no auditório, Eugênio continua pensando: *Na verdade vale a pena trocar qualquer que seja o valor do primeiro cheque. Se chamarmos este valor de* x *temos por um lado a opção de ficar com* x *e por outro lado a opção de arriscar, com probabilidade* $1/2$ *de ganhar* $0.45x$ *e probabilidade* $1/2$ *de ganhar* $1.95x$*. No primeiro caso o valor esperado é* x *e no segundo caso o valor esperado é* $1.2x$*. Assim, como* $x > 0$*, vale sempre a pena trocar.* Eugênio fica feliz com sua conclusão e pensa como seu curso de Probabilidade foi útil.

Mas um pouco mais tarde Eugênio começa a ter dúvidas quanto a suas conclusões: *Se vale a pena trocar de envelope sempre então não é necessário abrir o envelope e ler o valor do cheque para tomar a decisão de trocar. Neste caso eu poderia ter trocado de envelope um minuto antes e ter evitado a multa.* Eugênio fica irritado, pensando que poderia ter ganhado 5 reais a mais se apenas tivesse pensado mais rápido. Mas ele continua pensando: *Ei, espere, há algo errado! Um minuto antes os dois envelopes estavam lacrados e pareciam iguais para mim: trocar significaria apenas escolher o outro. Mas então cada vez que eu penso em um envelope tenho que trocar e nunca posso escolher nada!* Assim, ao invés de aproveitar seu prêmio, Eugênio passa a noite angustiado com seu paradoxo. Na manhã seguinte, Eugênio procura seus colegas do curso de Probabilidade com a pergunta: *o que exatamente há de errado com este raciocínio*?

O erro de Eugênio está logo no início, quando aceita, sem aliás sequer questionar, que a probabilidade do segundo cheque ser maior é $1/2$. O leitor deve estar muito surpreso: é quase como se de repente disséssemos que cara e coroa têm probabilidades diferentes. Por isso daremos uma explicação relativamente longa para tentar convecer.

Começaremos fazendo algumas digressões considerando o que um outro convidado, o João, que nunca estudou probabilidade mas que tem bom senso, faria em algumas situações extremas. João não acompanha todos os sorteios mas mesmo assim ele certamente tem *alguma* noção, por vaga que seja, de qual a faixa dos prêmios. Assim, se o valor do primeiro cheque fosse muito baixo, João certamente pensaria: *Não é possível, ou pelo menos não é provável, que o segundo cheque seja ainda menor. Assim, quase certamente eu peguei o envelope com o cheque de menor valor (além de ter tido o azar de vir em um dia em que os prêmios foram baixos) e aposto que o outro cheque é maior: vou trocar.* Por outro lado, se o valor do primeiro cheque fosse muito alto, seu pensamente seria: *Que sorte, hoje os prêmios estão ótimos! E é muito improvável que o segundo cheque seja ainda maior! Vou ficar com este cheque mesmo!* Assim, João *não* atribui probabilidades iguais às duas possibilidades (o segundo cheque ser maior ou menor), e as probabilidades que ele atribui (inconscientemente) a estes dois eventos dependem do valor do primeiro cheque.

Bem, este era o João e não o Eugênio: ao considerá-lo nos desviamos temporariamente do problema original e do contexto que nos impusemos no primeiro parágrafo deste desenvolvimento pois João nem sabe o que é o valor esperado e seus critérios não são os de Eugênio. João atribuiu subjetivamente probabilidades diferentes aos dois eventos; Eugênio (que aliás não se defrontou com situações extremas) atribuiu probabilidades iguais. Será que em algum sentido é errado atribuir sempre probabilidades iguais?

Sim, atribuir probabilidades sempre iguais é não apenas errado mas contraditório com a Teoria da Probabilidade que Eugênio tenta usar. Para entender isto, vamos representar cada configuração inicial de envelopes por um par ordenado (x_1, x_2) de números reais positivos: x_1 é o valor do cheque no primeiro envelope escolhido pelo convidado e x_2 é o valor do segundo cheque. Assim, o espaço amostral $\Omega \subseteq \mathbb{R}^2$ é a união de duas semi-retas abertas partindo da origem, como mostrado na Figura 3. A história que contamos envolvendo Eugênio corresponde ao ponto $(100, 200)$, também indicado. Ao abrir o primeiro envelope definimos o valor de x_1 e ficamos restritos à interseção de Ω com uma reta vertical, ou seja, aos dois pontos $(x_1, x_2 = 2x_1)$ e $(x_1, x_2 = x_1/2)$.

Eugênio implicitamente aceita que a probabilidade condicional a um valor qualquer fixo para x_1 destes dois pontos é $1/2$. Assim, ele deve aceitar que

$$P(\{(t, 2t); t \in [T, 2T)\}) = P(\{(t, t/2); t \in [T, 2T)\})$$

para qualquer número positivo T, onde $P(C)$, $C \subseteq \Omega$, denota a probabilidade de que (x_1, x_2) esteja em C. Por outro lado, a simetria inicial entre os envelopes diz que

$$P(\{(t, 2t); t \in [T, 2T)\}) = P(\{(2t, t); t \in [T, 2T)\}).$$

Sejam

$$A_n = \{(t, 2t); t \in [2^n, 2^{n+1})\},$$
$$B_n = \{(2t, t); t \in [2^n, 2^{n+1})\},$$

onde n é um inteiro qualquer; as identidades acima nos dão $P(A_n) = P(B_{n-1})$ e $P(A_n) = P(B_n)$, respectivamente. Por indução, $P(A_n) = P(B_n) = P(A_0)$ para todo n. Observemos

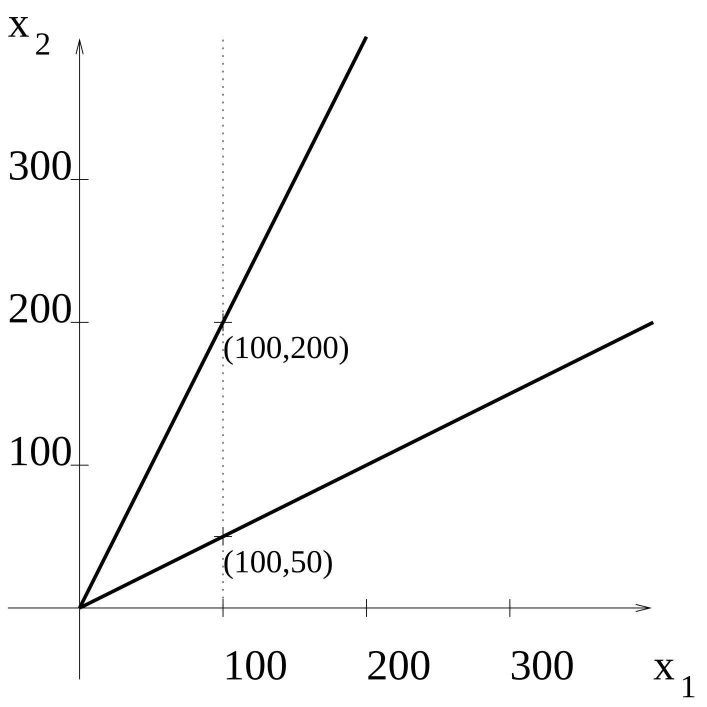

Figura 3

desde já que esta conclusão é no mínimo estranha: ela diz que a probabilidade do valor de menor cheque estar entre 64 e 128 é igual à probabilidade do menor cheque estar entre 2^{64} e 2^{65}, ou entre $2^{4199021}$ e $2^{4199022}$; no próximo parágrafo veremos que esta conclusão é não apenas estranha mas realmente absurda, mesmo ignorando o fato que um prêmio de R\$ $2^{4199021}$ é uma impossibilidade prática.

Observemos que os conjuntos A_n e B_n são dois a dois disjuntos e sua união é Ω. Se $P(A_0) > 0$ podemos tomar $N \in \mathbb{N}$ tal que $NP(A_0) > 1$ e temos

$$P\left(\bigcup_{0 \le n < N} A_n\right) > 1,$$

o que é absurdo. Por outro lado, se $P(A_0) = 0$ temos

$$P\left(\bigcup_{-N < n < N} (A_n \cup B_n)\right) = 0$$

para todo N, o que também é um absurdo pois quando N cresce este conjunto também cresce tendendo no limite para Ω donde teríamos $P(\Omega) = 0$, contradizendo $P(\Omega) = 1$. Assim, em qualquer caso temos um absurdo.

Esta explicação é um pouco técnica mas coincide perfeitamente com o "bom senso" de João: não podemos ignorar o primeiro cheque. Se seu valor for muito baixo, a probabilidade de que o segundo cheque seja maior deve em geral ser muito maior do que $1/2$ pois $P(A_n)$ deve tender a zero quando n tende a $-\infty$. Por outro lado se o seu valor for muito alto, a probabilidade de que o segundo cheque seja ainda maior deve ser muito menor do que $1/2$ pois $P(A_n)$ também deve tender a zero quando n tende a $+\infty$. E Eugênio afinal de contas precisa fazer uma avaliação sutil, dependendo de que valores são plausíveis como prêmio: até um certo valor limite vale a pena trocar, acima deste valor não.

Nicolau C. Saldanha
Departamento de Matemática, PUC-Rio
Gávea, Rio de Janeiro, RJ 22453-900, BRASIL

nicolau@mat.puc-rio.br, http://www.mat.puc-rio.br/∼nicolau/