COMO GANHAR UM AUTOMOVEL

Antonio Luiz Pereira - IME-USP

1 O problema original

Em seu artigo **Como Perder Amigos e Enganar Pessoas** *Eureka!* n^0 1 o Prof. Nicolau Saldanha discute quatro problemas envolvendo probabilidades, mostrando como a 'intuição' pode ser enganadora nesse assunto. A discussão apresentada é bastante interessante e esclarecedora. Entretanto, acreditamos que cabem algumas considerações adicionais no caso do problema 1. Na verdade o problema, como já observado pelo Prof. Saldanha, é antigo e provocou um grande debate inclusive na literatura especializada, onde às vezes é chamado 'O Dilema do Prisioneiro', em virtude de uma versão formulada em termos de uma decisão a ser tomada por um prisioneiro (ver por exemplo [3]). Recentemente, uma versão do mesmo problema foi apresentado pelo Prof. Augusto C. Morgado no n^o 33 da RPM.

Inicialmente vamos apresentar um enunciado ligeiramente(?) distinto do proposto pelo Prof. Saldanha (e mais próximo do que aparece em [2] e [3]), para tentar tornar mais evidente o ponto polêmico. Discutiremos depois as diferenças com o enunciado original.

O problema dos bodes (**Problema P**). *Suponha que um convidado está em um programa de televisão e deve escolher entre três portas, uma das quais esconde um automóvel e as outras duas dois bodes. O convidado escolhe uma das portas. Em seguida, o apresentador, que sabe o que as portas escondem, escolhe uma das duas restantes mostrando um bode. Ele então pergunta ao convidado: você quer trocar de porta? O problema é: é vantajoso para o convidado fazê-lo? Se o fizer, qual é sua probabilidade de ganhar o automóvel?*

Um fator complicador neste problema (e em outros do mesmo tipo), é que certas condições não estão explicitamente enunciadas, mas precisam ser utilizadas na resolução do problema. Por exemplo, *não se afirma que a probabilidade do carro estar atrás de qualquer uma das portas é a mesma.* Parece, porém, na ausência de informação contrária, bastante natural assumir esta hipótese de 'equiprobabilidade'. Nós o faremos aqui (assim como O Prof. Saldanha), exceto na seção final.

Existe ainda outro aspecto a ser examinado: o que exatamente faz o apresentador após a escolha inicial da porta pelo convidado? Para fixar idéias vamos supor que ele escolha a porta 1 (em vista da hipótese de equiprobabilidade acima isto é irrelevante). Suponha que a porta 1 esconda um bode. Nesse caso o apresentador não tem escolha; ele deve abrir, entre as duas portas restantes, aquela que contem um bode.

Agora, suponhamos que a porta 1 esconde o automóvel. O que faz o apresentador? Ele pode abrir qualquer das duas portas restantes. Como faz a *escolha* da porta a ser aberta? Queremos enfatizar que ele tem de usar algum 'critério' para escolher a porta. Estamos usando a palavra critério num sentido bem amplo; ele pode jogar um dado, consultar um oráculo, a platéia, ou mesmo usar métodos diferentes de cada vez. Vamos *assumir* que, qualquer que seja a estratégia usado pelo apresentador para escolher a porta, resulta que a probabilidade de que a porta 2 seja escolhida é p e, portanto, a probabilidade para a porta 3 é $q = 1 - p$ (estamos assumindo em particular que estas probabilidades não variam com o tempo). Aqui há um ponto importante: devemos assumir $p = q = \frac{1}{2}$, argumentando, como acima, que esta hipótese está 'implícita' no enunciado? Se esta hipótese é assumida, a probabilidade do convidado ganhar o carro, trocando de porta, é sob qualquer ponto de vista, $\frac{2}{3}$. Agora, pode ser interessante entender o que acontece *se esta hipótese não é assumida.*

Vamos agora modificar o problema original de modo a obter dois problemas distintos. Em seguida veremos como resolvê-los e discutiremos se algum deles é, em algum sentido, 'equivalente' ao problema original.

Problema **P1**

O mesmo enunciado do problema **P** *exceto que o convidado deve decidir se vai trocar de porta,* **antes** *que o apresentador abra, de fato, uma das portas restantes, mostrando um bode.*

Problema **P2**

O mesmo enunciado do problema **P** *mas agora assumindo que o convidado conhece exatamente a estratégia do apresentador (conhece p).*

Cremos que seja razoável optar por resolver o problema **P1** para responder à pergunta original; afinal *se o convidado não conhece a estratégia do apresentador*, nada vai ganhar sabendo que, digamos, a porta 3 (e não a 2) foi aberta e pode muito bem definir uma estratégia "a priori", talvez antes mesmo do jogo começar.

Por outro lado, há fortes argumentos indicando que o problema a ser resolvido é **P2**. Afinal é *um dado* do problema que uma porta foi aberta e então esta informação pode ser usada, por exemplo por um observador externo que conheça a estratégia do apresentador. As probabilidades calculadas com esta informação não serão acessíveis ao convidado, se ele não conhecer a estratégia do apresentador *mas traduzirão dados da situação real*; afinal esta estratégia é um elemento do problema, ainda que desconhecido do convidado. Além disso,

o convidado pode *mesmo desconhecendo a estratégia do apresentador* computar as probabilidades como *função* de p. Ainda que este cálculo não lhe forneça um número, pode indicar (e de fato indicará como veremos!) superioridade de uma estratégia sobre outra. Finalmente, observemos que
a) a pergunta é: 'qual é a probabilidade de ganho?' e não 'qual é a probabilidade calculável pelo convidado?'.
b) o convidado poderia, em princípio, conhecer algo a respeito da estratégia do apresentador, por exemplo, por ter assistido várias vezes o programa de televisão!

Numa linguagem um pouco mais técnica, resolver o problema **P1** significa calcular a probabilidade de ganho do carro, dado que *alguma das portas restantes* vai ser (ou foi) aberta pelo apresentador e resolver o problema **P2** signifa calcular a mesma probabilidade, dado que *uma porta específica (*2 *ou* 3*)* foi de fato aberta.

Para tornar a exposição mais completa e (esperamos) mais clara, vamos primeiro apresentar a solução do problema **P1**, de forma ligeiramente distinta da apresentada no artigo do Prof. Saldanha.
Antes da abertura da porta pelo apresentador, os resultados possíveis na situação descrita no problema (ou seja o espaço amostral) são:

$$\{ABB2, ABB3, BAB3, BBA2\}$$

onde A indica automóvel, B indica bode, a posição das letras indica o número da porta e o número no final indica a porta que foi aberta pelo apresentador, depois que o convidado indicou a porta 1.
A probabilidade de cada um dos pontos do espaço é, (com a hipṕotese de equiprobabilidade acima) na ordem: $\frac{1}{3}p$, $\frac{1}{3}q$, $\frac{1}{3}$, $\frac{1}{3}$.

Agora, se o convidado trocar de porta, ele perderá nos casos 1 e 2 e ganhará nos casos 3 e 4. Segue que a probabilidade de ganho usando a estratégia de trocar a porta é: $\frac{1}{3} + \frac{1}{3} = \frac{2}{3}$. (e a de perder é, claramente, $\frac{1}{3}p + \frac{1}{3}q = \frac{1}{3}$).

Agora, no caso do problema **P2**, o que queremos calcular é a probabilidade do convidado ganhar o automóvel trocando de porta, *dado* que uma porta foi efetivamente aberta. Vamos supor, que a porta 3 foi aberta (o cálculo é análogo se supusermos que a porta 2 foi aberta). Queremos, portanto encontrar a probabilidade (condicional) do evento G_t = "o convidado ganha o automóvel trocando de porta", dado o evento $E3$ = "a porta 3 foi aberta pelo apresentador".

Temos

$$\begin{aligned} P(G_t|E3) &= \frac{P(G_t \cap E3)}{P(E3)} \\ &= \frac{P(BAB3)}{P(\{BAB3, ABB3\})} \\ &= \frac{1/3}{1/3 + q/3} \end{aligned}$$

$$= \frac{1}{1+q}$$

Concluímos, portanto, que a probabilidade de ganho do automóvel, trocando de portas, *depende da estratégia usada pelo apresentador*, um resultado considerado surpreendente pela maioria das pessoas com as quais discutimos este problema.

Algumas observações parecem ser pertinentes neste ponto.

1. Se $q = 1/2$, então a resposta é a mesma para os problemas $P1$ e $P2$. Neste caso (e só nele), o conhecimento de qual porta foi efetivamente aberta é irrelevante sob qualquer ponto de vista.

2. Como $\frac{1}{2} \leq \frac{1}{1+q} \leq 1$ então a troca de portas é vantajosa, a menos que $q = 1$, caso no qual é indiferente a troca ou não de porta. Segue que o convidado pode concluir que deve trocar de porta *mesmo desconhecendo a estratégia do apresentador.*

3. Se $q = 1$ então a porta 3 é sempre a escolhida pelo entrevistador, caso possível. Sua abertura, portanto, não aumenta as chances do convidado exceto pelo fato de estar eliminando automaticamente uma possibilidade. Esta é uma razão intuitiva para que $P(G_t|E3)$ seja $1/2$ neste caso, uma resposta que aparece frequentemente nas tentativas de solução do problema.

   Por outro lado, se $q = 0$ a abertura da porta 3 imediatamente implica que estamos no caso $BAB3$ (pois senão o entrevistador escolheria a porta 2). Isto quer dizer que a troca de portas implica ganho certo neste caso.

4. A probabilidade (não condicionada) calculada no problema **P1** pode ser obtida como uma 'média ponderada' das probabilidades condicionadas aos eventos E2= "a porta 2 foi aberta pelo apresentador" e E3. De fato, como E2 e E3 são eventos disjuntos e complementares em nosso espaço amostral, temos:

$$\begin{aligned} P(G_t) &= P(G_t|E2 \cup E3) \\ &= P(G_t|E2)P(E2) + P(G_t|E3)P(E3) \\ &= (\frac{1}{3} + \frac{1}{3}q)(\frac{1}{1+q}) + (\frac{1}{3} + \frac{1}{3}p)(\frac{1}{1+q}) \\ &= \frac{2}{3} \end{aligned}$$

5. Como já observamos, se a porta indicada pelo convidado é outra que não a 1, o mesmo raciocínio se aplica. Temos apenas que lembrar que a estratégia do apresentador poderia mudar, ou seja, poderia *depender* da escolha inicial, e seria determinada por probabilidades p', q' não necessariamente iguais a p e q.

Para terminar esta seção, seria talvez adequado discutir qual das duas soluções acima é a mais adequada à situação apresentada. Cremos que, na falta de uma explicitação completa de todas as hipóteses necessárias, o modêlo probabilístico a ser adotado não fica completamente determinado e, portanto, o problema fica sujeito a mais de uma interpretação. Na nossa opinião *com o enunciado que propusemos acima* a solução proposta para o problema **P2** é a mais adequada. Por outro lado, na formulação proposta pelo Prof. Saldanha, como se pergunta qual é a probabilidade da 'melhor estratégia' e não qual é a probabilidade em um caso específico, nos inclinamos (como o fez o Prof. Saldanha) a considerar o problema **P1** (isto é a probabilidade de uma estratégia 'a priori') como o mais adequado.

2 Uma pequena generalização

Em toda a discussão acima, estivemos assumindo que a probabilidade de que o bode se encontre em qualquer das portas é a mesma. É interessante considerar o caso em que isto não ocorre porque ele pode levar a uma situação na qual os problemas **P1** e **P2** conduzem a *diferentes estratégias para o convidado.* Em outras palavras, pode ser que a probabilidade de ganho do carro, trocando de porta, seja maior do que $\frac{1}{2}$ antes da abertura de fato da porta pelo entrevistador e menor do que $\frac{1}{2}$ depois disso.

Vamos então denotar por:

- q_i a probabilidade de que o automóvel esteja na i-ésima porta.
- $p^j_{i,k}$ a probabilidade de que o apresentador abra a k-ésima porta, dado que o convidado indicou inicialmente a j-ésima porta e o automóvel se encontra na i-ésima porta.

onde $1 \leq i, j, k \leq 3$.

As hipóteses do problema implicam em certas condições para os $p^j_{i,k}$. Por exemplo, como estamos assumindo que o apresentador não abre a porta que contém o carro nem a porta já aberta pelo convidado, devemos ter $p^j_{i,k} = 0$, se $k = i$ ou $k = j$. Além disso, como alguma das portas restantes após a escolha do convidado, deve ser escolhida pelo entrevistador, devemos ter $\sum_{k \notin \{i,j\}} p^j_{i,k} = 1$.

Ao contrário do problema original, agora a escolha inicial do convidado interfere nas suas chances de ganho, mesmo antes da abertura da porta pelo entrevistador e devemos considerá-la em nossos cálculos desde o início. Usaremos então uma notação diferente da anterior. Para cada $1 \leq j \leq 3$, correspondente à escolha da j-ésima porta pelo convidado, consideremos o espaço amostral:

$$\Omega_j = \{(j; i, k) 1 \leq i, k \leq 3\}$$

onde i e k têm o significado indicado acima.

Cada ponto $(j;i,k)$ em Ω_j ocorre se o automóvel está na porta i (com probabilidade q_i) e o apresentador escolhe a porta k (com probabilidade) $p^j_{i,k}$. Sua probabilidade , portanto é $P((j;i,k)) = q_i p^j_{i,k}$.

Como no problema original, o convidado ganha o automóvel, trocando de porta, se e somente se sua escolha inicial corresponde a uma porta que esconde um bode, ou seja, se $i \neq j$. Assumindo que o convidado escolheu inicialmente a porta j temos:

$$\begin{aligned}
P(G_t) &= \sum_k \sum_{i\neq j} q_i p^j_{i,k} \\
&= \sum_{i\neq j} q_i \sum_k p^j_{i,k} \\
&= \sum_{i\neq j} q_i
\end{aligned}$$

Podemos ver então que a melhor estratégia para o convidado é escolher inicialmente a porta que tem a menor probabilidade de esconder o carro pois, dessa forma, sua chance de ganhar depois de trocar de porta é máxima.

Suponhamos agora que o convidado já escolheu a j-ésima porta (não necessariamente a “melhor”) e o entrevistador já abriu uma das outras, mostrando um bode. Qual é agora, a probabilidade (condicional) de ganho do automóvel, trocando de porta? (ou seja qual é a solução do problema **P2**?)

Seja E_k o evento: “a porta k foi aberta pelo entrevistador”(assumimos aqui $k \neq j$ pois $P(E_k) = 0$ caso contrário). Lembrando que $p^j_{i,k} = 0$ se $k = i$ ou $k = j$ temos:

$$\begin{aligned}
P(G_t|E_k) &= \frac{\sum_{i\neq j} q_i p^j_{i,k}}{\sum_i q_i p^j_{i,k}} \\
&= \frac{\sum_{i\notin\{j,k\}} q_i p^j_{i,k}}{\sum_{i\neq k} q_i p^j_{i,k}} \\
&= \frac{\sum_{i\notin\{j,k\}} q_i \sum_k p^j_{i,k}}{\sum_{i\notin\{j,k\}} q_i \sum_k p^j_{i,k} + q_j p^j_{j,k}} \\
&= \frac{\sum_{i\notin j,k} q_i}{\sum_{i\notin j,k} q_i + q_j p^j_{j,k}}
\end{aligned}$$

Indicando por i^* o único inteiro entre 1 e 3 diferente de j e k, temos finalmente

$$P(G_t|E_k) = \frac{q_{i^*}}{q_{i^*} + q_j p^j_{j,k}}$$

É fácil encontrar valores de q_i e $p^j_{i,k}$ para os quais se tenha $P(G_t|E_k) < \frac{1}{2} < P(G_t)$. Por exemplo se $j = 1$, $k = 3$, $q_1 = \frac{1}{3}$, $q_2 = \frac{1}{8}$ e $p^1_{1,3} = \frac{3}{4}$ então

$$P(G_t|E_3) = \frac{q_2}{q_2 + q_1 p^1_{1,3}} = \frac{\frac{1}{8}}{\frac{1}{8} + \frac{1}{3}\frac{3}{4}} = \frac{1}{3}$$

Por outro lado:

$$P(G_t) = q_2 + q_3 = 1 - q_1 = \frac{2}{3}$$

Portanto, neste exemplo, é melhor, como estratégia geral 'a priori', trocar de porta mas, se o apresentador escolheu a porta 3, é melhor não fazê-lo!

Referências

[1] Saldanha, N.C., *Como perder amigos e enganar pessoas*, EUREKA! n^o 1, 1998.

[2] Morgado, C.M., *Os dois bodes* , RPM 33, 26-29.

[3] Morgan, J.P.V., Chaganty N.R., Dahiya R.C e Doviak M.J. , *Let's Make a Deal: The Player's Dilemma*, The American Statistician, November 1991, vol 45, n. 4, 284 - 287.