por **Fabio Chalub**
[Universidade Nova de Lisboa]

Novas Informações da Teoria da Informação

A imagem abaixo pode dar-nos a impressão de estarmos a entrar na praia do Meco. No entanto o significado desejado é exactamente o oposto: num escritório para tirar cartas de condução, próximo de uma das praias mais famosas do mundo, a Direcção-Geral de Viação local tenta impor aos utentes um código de vestimenta que dê ao lugar a respeitabilidade que considera devida. Que tem a matemática com isto?

O que fazer?

Tudo o que se refere a comunicação é antigo como o ser humano. Talvez seja a habilidade de trocar informações a característica que mais nos diferencia do resto do reino animal. No entanto, os estudos matemáticos sobre a comunicação só começaram a ser sistematizados há exactos 60 anos, quando Claude Shannon e Waren Weaver publicaram o livro *"A Teoria Matemática da Comunicação"* [2]. Na verdade, o livro é uma reedição alargada de um artigo do ano anterior cujo título era parecido: *"Uma Teoria Matemática da Comunicação"*. A mudança da primeira palavra e a rápida transformação de um artigo científico em livro, ainda hoje reeditado, mostra a importância deste trabalho.

Shanon e Weaver definem três "problemas da comunicação": i) *o problema técnico*: quão exactamente estão a ser transmitidos nos símbolos? ii) *o problema semântico*: com que precisão os símbolos transmitem a mensagem desejada? iii) *o problema da efectividade*: será que o sentido transmitido fará o recipiente da mensagem comportar-se da forma adequada? Na figura 1 os problemas ii) e iii) estão claramente postos e o leitor pode divertir-se a pensar nos significados alternativos e nas suas consequências.

Os símbolos transmitidos podem ser um conjunto linear discreto, como as letras que compõem esta página, ou até mesmo uma função de três variáveis, sendo uma o tempo e duas que definem o plano da televisão. Vamos pensar no primeiro exemplo, por ser ele mais simples. Suponha que a transmissão dos caracteres é feita através de um canal que troca todas as letras A por uma outra qualquer. Sendo esta a letra mais comum em português, é natural que o efeito para a transmissão da informação seja maior do que se o mesmo acontecesse no X. Desta forma, as letras têm um conteúdo informativo distinto.

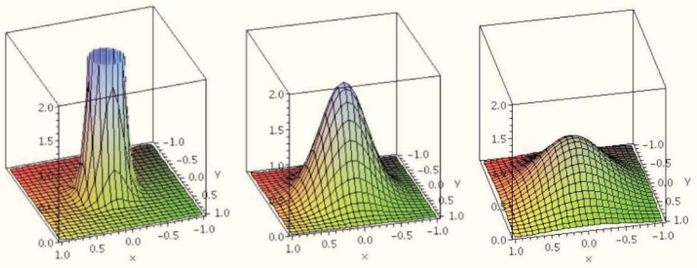
Como medir este conteúdo informativo? Shannon recorreu aos trabalhos de Ludwig Boltzman em física estatística no século 19, para definir a *entropia da informação*, ou, como a conhecemos actualmente, a *entropia de Shannon*. Para Boltzman, a entropia mede a desordem de um sistema; para Shannon, o que está a ser medido pelo inverso da entropia é a informação. Uma sucessão de caracteres alfabéticos completamente desordenada é aquela da qual não podemos obter nenhum significado. Para transmitir alguma informação é necessário colocar as letras certas nos lugares certos.

Na verdade, não exactamente, pois muitas vezes há redundância no texto (só em algumas linguagens artificiais não existe redundância). Desta forma é possível comprimir informação sem a perder. Torna-se mais económico guardá-la no disco rígido do computador ou enviar um ficheiro por *e-mail*. É o que fazem os programas de compressão de dados, como o *zip*. O limite teórico de compressão sem perda é dado pelo *teorema da codificação de fonte*, do próprio Shanon.

A expressão para a entropia é obtida da seguinte forma: considere n possíveis símbolos, ou expressões, ou seja o que for que tenha conteúdo informativo, cada um destes com probabilidade $p(1)$, $p(2)$, até $p(n)$ de ser o escolhido. Como estamos a falar de probabilidade, evidentemente temos de ter $p(1)+p(2)+...+p(n)=1$. A entropia é dada por

$H=-p(1)\log p(1)-p(2)\log p(2)-...-p(n)\log p(n)$.

Após esta definição o trabalho de Shannon passa a ter cariz essencialmente matemático. Antes de formular a sua conjectura recentemente demonstrada e já generalizada, vamos analisar um exemplo simples, conhecido como o *modelo dos passos de bêbado*. Depois de alguns copos a mais, uma pessoa põe-se a andar, mas dá passos totalmente ao acaso, sem uma direcção privilegiada. Em média, não sai do lugar. Esta é a maneira informal de dizer que a distribuição de probabilidade de sua posição no espaço tem média zero. No entanto, se considerarmos a média do quadrado da distância, esta tem de ser positiva. Mais exactamente, a raiz quadrada da média do quadrado da distância (média determinada a partir de um grande número de bêbados a andarem ao acaso) é proporcional à raiz quadrada do número de passos dados por cada pessoa (supomos todos os passos do mesmo tamanho). Como o número de passos é directamente proporcional ao tempo, concluímos que a dispersão do *movimento Browniano* é proporcional à raiz quadrada do tempo. Este nome foi dado originalmente ao movimento aleatório de partículas de pólen dispersas na água, descrito pelo botânico britânico Robert Brown, explicado por Albert Einstein e que inspirou o matemático Mandelbrot a criar os fractais. Antes mesmo de Einstein, Luis Bachelier usou ideias parecidas no estudo da bolsa de valores.



Distribuição de probabilidade para o movimento Browniano em função do tempo: t=0.01 (esquerda), t=0.05 (centro), t=0.1 (direita). À medida que o tempo passa a dispersão aumenta, mas o ponto médio continua o mesmo: a origem. O eixo vertical é o mesmo nas três figuras (o que omite a parte mais alta do gráfico na figura da esquerda).

Finalmente chegamos à conjectura de Shannon: considere N eventos ("os bêbados"), cada um descrito por uma função X_1, X_2, ..., X_N, e considere a soma destas N variáveis e divida pela raiz quadrada de N. Chamamos ao resultado Z_N. O teorema central do limite garante que Z_N se aproxima de uma distribuição gaussiana quando N cresce. Além disto, e esta é a conjectura de Shannon, a entropia da variável aleatória Z_N é uma função não crescente de N.

Se $p(x)$ descreve a distribuição de probabilidade de um certo evento (no caso acima, o evento descrito pela função Z_N), então a sua entropia é dada pelo integral de $-p(x)\log p(x)$ em todos os valores possíveis de x, generalizando a fórmula descrita anteriormente.

No final da década de 50 foi finalmente demonstrado que a conjectura de Shannon era verdadeira quando comparávamos a entropia de 2 eventos com a de 1 evento. Foram necessários mais 45 anos para que finalmente, em 2004, 4 cientistas em 4 países diferentes a conseguissem demonstrar em toda a generalidade [1]. Mais 3 anos e foi finalmente possível caracterizar os casos em que a entropia é uma função decrescente (e não apenas não crescente) [3].

A demonstração em [1] usa um pouco de tudo: cálculo variacional, cálculo vectorial, análise funcional, semi-grupos, probabilidades e muitas, mas muitas, desigualdades entre funções. No entanto, para quem esperava a necessidade de desenvolver técnicas novas e profundas na solução de um problema com mais de 50 anos, a demonstração pode ser decepcionante. A verdadeira arte foi trabalhar conceitos já consolidados há décadas de forma criativa.

A extensão feita em [3] já é mais sofisticada. Foi necessário usar uma versão não-comutativa da teoria das probabilidades, desenvolvida por Voiculescu no início da década de 90. Foi a assim chamada teoria de probabilidades livres que permitiu não apenas a caracterização acima descrita, mas também a extensão do resultado anteriormente demonstrado para o caso de várias variáveis aleatórias em paralelo (como se muitos bêbados partissem de muitos pontos da cidade em simultâneo). M

Referências

[1] **Arstein, S., Ball, K., Barthe, F., Naor, A.** (2004). "Solution of Shannon's problem on the monotonicity of the entropy" *J. Am. Math. Soc* 17:975-982.

[2] **Shannon, C. E., e Weaver, W.**(1949). *The Mathematical Theory of Communication*. University of Illinois Press

[3] **Shlyakhtenko, D.** (com um apêndice por **H. Schultz**) "Shannon's monotonicity problem for free and classical entropy" (2007). *Proc. Nat. Acad. Sci.* 104(39) 15254-15258.