

Analizando dados

A avaliação de uma ideia à luz da [evidência](#) disponível deveria ser uma atividade simples e óbvia, não é verdade? Ou existe acordo entre os resultados e as [expetativas](#) geradas por uma ideia (ou seja os resultados suportam a ideia) ou não (ou seja, rejeitam-na). Por vezes, este processo é relativamente simples (por exemplo, perfurações de um atol coralino ou revelam uma espessa camada de coral, ou apenas uma fina camada superficial). Mas frequentemente tal não é o caso. O mundo real é intrincado e complexo e, muitas vezes, interpretar a evidência associada a uma certa ideia está longe de ser uma tarefa fácil. Para complicar ainda mais a situação, torna-se frequentemente necessário pesar múltiplas [linhas de evidência](#) de modo a determinar a validade de uma dada ideia.



Recolher dados em bruto:

```
FTCGCTATAACTGCACCATCAGAGG
ATCTAAAGTTCTCTATGGGAATACC
TAAGACAGACGATCATGATCGCATT
TGCAACTATCGCGCAACACAACACTG
ACTGACGTCATTGAGACGGCACTT
```

Analisar dados:

```
fly:  ACTTCGCTATAACTGCAC
spider: ACTTCCTATAGTTGCAA
lobster: ACTTCGCTATAACTGCAC
```

Interpretar resultados:



Consenso da comunidade:



Os [testes](#) científicos geram tipicamente aquilo que os cientistas chamam de [dados](#) em bruto — [observações](#), descrições ou medições inalteradas, tal como foram obtidas — as quais devem ser posteriormente analisadas e interpretadas. Os dados tornam-se evidência apenas depois de terem sido interpretados de modo a esclarecer o grau de rigor de uma ideia científica. Por exemplo, a investigação das relações evolucionárias entre crustáceos, insetos, diplópodes, araneídeos (aranhas) e os seus vizinhos taxonómicos, pode fornecer a sequência genética de um gene específico em cada um destes organismos. Estes são os dados em bruto mas, na realidade, o que é que eles nos dizem? As longas sequências de Ás, Tês, Gês e Cês que formam o ADN não podem, por si próprias, dizer-nos se os insetos se encontram mais próximos dos crustáceos ou dos araneídeos. Em vez disso, os dados têm que ser analisados através de cálculos estatísticos, o uso de tabelas e/ou representações gráficas. No presente caso, uma bióloga poderá iniciar a análise da informação genética alinhando sequências diferentes, chamando a atenção para semelhanças e diferenças, e efetuando cálculos com o objetivo de quantificar o grau de similitude entre as diversas sequências. Apenas então poderá interpretar os resultados e descobrir se estes dão ou não suporte à [hipótese](#) de que os insetos se encontram geneticamente mais próximos dos crustáceos do que dos araneídeos.

Além disso, o mesmo conjunto de dados pode ser interpretado de maneiras diversas. Consequentemente, um outro cientista poderia analisar os mesmos dados usando um método novo e chegar a uma conclusão diferente sobre a relação entre insetos, crustáceos e araneídeos. No fim, a comunidade científica acabará por chegar a um consenso acerca de como um conjunto de dados deve ser interpretado, mas este processo poderá levar algum tempo e normalmente envolve o uso de linhas de evidência adicionais.

CALCULANDO CONFIANÇA

Interpretar o resultado de um teste significa frequentemente que devemos lidar com incerteza e erro. "Mas espere um momento," — estará talvez pensando — "eu estava convencido que a ciência tinha por objetivo construir conhecimento e *reduzir* incerteza e erro." E tal é verdade. Contudo, quando os cientistas chegam a uma conclusão ou fazem um cálculo, frequentemente eles tentam dar uma indicação estatística da confiança que têm nos resultados. Na linguagem do dia-a-dia, incerteza e erro significam que uma resposta não é clara ou que algo foi feito de forma incorreta. Porém, quando os cientistas falam de incerteza e erro, normalmente estão a referir-se ao grau de *confiança* que eles têm num determinado valor. Por exemplo, dizer que a temperatura é de 37°C (com uma margem de incerteza de mais ou menos 0.2°C , na realidade significa que temos grande confiança de que a temperatura real se encontra entre 36.8 e 37.2°C .