FABIO CHALUB
Universidade Nova
de Lisboa
chalub@fct.unl.pt

5 *SIGMAS* DE **CERTEZA**

A confirmação experimental da existência do bosão de Higgs foi uma das grandes conquistas da física contemporânea. O seu anúncio, no entanto, foi cheio de cuidados: era uma partícula *compatível*, mas podia não ser a própria; é necessário ter a certeza de "5 sigmas", etc. Mas o que é uma certeza de 5 sigmas? Aliás, o que é a certeza em ciências experimentais? Considerando um exemplo simples das sondagens eleitorais, vamos discutir como se pode medir o grau de confiança de um enunciado científico, seja em ciências humanas ou mesmos nos experimentos mais precisos do mundo, como os realizados no Centro Europeu de Investigação Nuclear (o famoso CERN, na fronteira franco-suíça).

Imagine um país em que a população se divide politicamente em dois grupos exatamente do mesmo tamanho: enquanto metade da população apoia o partido PX, um grupo de precisamente o mesmo tamanho apoia o partido PY. No entanto, isto não é sabido. A fratura na sociedade só será revelada ao mundo no dia das eleições. Enquanto a votação não chega, sondagens são feitas. O método consiste em escolher um pequeno número de eleitores e a partir das suas intenções de voto tentar inferir o resultado das urnas.

Vamos supor que dez cidadão foram escolhidos de forma aleatória para participar numa sondagem: Alice, Bernardo, Cátia, David, Elisa, Filipe, Germana, Hugo e Joana. Também para simplificar nossas contas, vamos supor que ninguém evita revelar o seu voto, que sempre é num dos dois concorrentes, o Governo — PX — ou a oposição, PY.

Vamos quantificar todas as possibilidades, escrevendo em maiúsculas os eleitores de PX e em minúscula os de PY. Há apenas uma maneira de todos votarem em PX: ABCDEFGHIJ, ou seja, os dez querem reconduzir o Governo. Há uma probabilidade de $1/2$ de cada componente da amostra ser governista, pois a sociedade está dividida em duas partes iguais; logo há uma probabilidade de $\frac{1}{2^{10}} = \frac{1}{1024}$ do resultado da sondagem ser 10-0, ou seja 10 votos no Governo e 0 na oposição.

Aumentam as possibilidades quando consideramos uma amostra de 9 governistas e um oposicionista entre 10 eleitores: aBCDEFGHIJ, AbCDEFGHIJ, ABcDEFGHIJ, . . . , ABCDEFGHIj. São exatamente dez configurações possíveis, e como cada uma tem uma probabilidade de $\frac{1}{1024}$, então há uma probabilidade de $\frac{10}{1024}$ do resultado da sondagem ser $9-1$. Obter o caso geral não é difícil. O número de combinações de n eleitores de PX num conjunto de 10 eleitores é dado por

$$C_n^{10} = \frac{10!}{n!(10-n)!}\,,$$

onde $n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$ é o fatorial de n. Levando em consideração as probabilidades, temos que numa população dividida politicamente meio a meio, a probabilidade de numa sondagem com 10 eleitores, n serem governistas é dada por

$$\frac{10!}{n!(10-n)!2^{10}}\,.$$

Veja a tabela 1 para as várias probabilidades (medidas em percentagens) para sondagens com 10 e 100 eleitores. A partir deste número poderíamos calcular médias, desvios-padrão e tudo mais (o desvio padrão é o intervalo típico – que precisaremos melhor abaixo – que os experimentos se desviam do valor esperado, ou seja, da média). No entanto, como devem lembrar das primeiras linhas, nós não sabemos que a sociedade se dividiu em dois, portanto nós não temos as probabilidades acima à disposição.

Abstraindo o exemplo para o caso das ciências experimentais, temos algumas observações (que no exemplo acima é o resultado da sondagem) e queremos saber o quanto é confiável extrapolar o resultado para o conjunto completo não observado (sejam experiências que não foram feitas, mas poderiam ter sido, seja o conjunto completo da população que não foi inquirido na sondagem). Para isto, formulamos o que é conhecido como *hipótese nula* que no nosso exemplo pode ser escrita: "a sociedade divide-se igualmente, em relação à sua intenção de voto, entre os partidos PX e PY", e calculamos a probabilidade de que, na validade desta hipótese, o resultado da sondagem seja exatamente igual aquele registado ou mesmo ainda mais extremo. Suponha, por exemplo, que na sondagem com apenas 10 eleitores, encontramos uma divisão de 7-3. Então, sob a hipótese nula, a possibilidade de termos este resultado ou um ainda mais extremo (8-2, 9-1, ou 10-0) é de

$$\frac{1}{2^{10}}\left(\frac{10!}{7!3!}+\frac{10!}{8!2!}+\frac{10!}{9!1!}+\frac{10!}{10!0!}\right)=\frac{176}{1024}\approx 17.2\%.$$

Portanto exite uma probabilidade de 17.2% de, numa sociedade dividida igualmente pelos dois partidos, por razões puramente aleatórias na escolha de uma amostra de eleitores, uma sondagem com 10 pessoas resultar em 7 ou mais votantes no partido do Governo. Pior ainda, há uma probabilidade de 34.4% de uma sondagem com 10 pessoas apresentar 7 apoiantes de um dos dois partidos, mesmo quando a sociedade como um todo está divida em grupos partidários de mesmo tamanho.

A conclusão é que a sondagem acima não nos permite deduzir que o PX será o vencedor, pois com alta probabilidade este resultado pode ter sido obtido apenas pelas flutuações estatísticas da escolha da amostra. Se fizermos uma sondagem com 100 pessoas e tivermos um resultado de 70-30, as conclusões são diferentes, apesar de todos os números serem uma simples multiplicação por 10 da sondagem restrita anterior.

PX	Probabilidades (10 eleitores, em %)	Probabilidades (100 eleitores, em %)
0-5%	0.098	10^{-20}
6-15%	0.98	10^{-11}
16-25%	4.4	10^{-5}
26-35%	12	0.18
36-45%	21	18
46-55%	25	68
56-65%	21	13
66-75%	12	0.090
76-85%	4.4	10^{-5}
86-95%	0.98	10^{-12}
96-100%	0.098	10^{-22}

Tabela 1: Na hipótese nula de uma sociedade dividida em grupos iguais, estas são as probabilidades, numa sondagem com 10 eleitores (centro) e com 100 eleitores (direita), de os vários resultados possíveis medidos em percentagem de apoiantes do partido PX. Note como aumenta a probabilidade de a sondagem capturar o resultado correto. (A soma de cada coluna difere de 100% devido aos arrendodamentos.)

Em particular:

$$\frac{1}{2^{100}}\sum_{70}^{100}\frac{100!}{n!(100-n)!}=$$
$$=\frac{49756171168061176633478360}{1267650600228229401496703205376}$$
$$\approx 0.003925069823\%$$

Portanto, a veracidade da hipótese nula é compatível com o resultado da sondagem ou um resultado mais extremo em menos do que 0.01% dos casos. Podemos então rejeitar a hipótese nula e supor que a sociedade não está dividida em dois, concluíndo uma provável vitória do partido PX. É importante notar que nenhum dos números discutidos acima fornece a probabilidade de vitória de PX nem a probabilidade de a hipótese nula ser verdadeira. De facto, chamamos ao complementar destes números "nível de confiança", no primeiro caso de 100% − 34.4% = 65.6% e no segundo caso de 99.99%.

Há, evidentemente, um conceito mal definido. Qual o número mágico que nos permite ter uma razoável certeza de que podemos rejeitar a hipótese nula? Não há resposta única. Nas sondagens eleitorais, assim como em grande parte do trabalho das ciências humanas e sociais, este valor é fixado em 5%. Por exemplo, se fizermos experimentos com seres humanos, ou mesmo com ratinhos, para medir se em dadas circunstâncias eles fazem *isto* ou *aquilo*, é costume considerar como hipótese nula que o seu comportamento é aleatório e ver qual a probabilidade de a hipótese nula reproduzir os dados observados ou mais extremos. Se esta for inferior a 5%, dizemos que a mesma foi rejeitada e temos uma conclusão a mostrar ao mundo. Caso contrário, nada afirmamos.

As ciências físicas costumam ter um padrão muito mais rigoroso. Aqui entram os tais 5σ (diz-se "cinco sigmas") de que tanto se falou nos experimentos do bosão de Higgs [1]. A hipótese nula é algo como "não há partícula de massa entre 125 GeV e 126 GeV no experimento" (o giga-eletrão-volt ou GeV é uma unidade de energia frequentemente usada em física de partículas como medida de massa – lembre-se da relação entre massa e energia antevista por Einstein; os valores usados decorrem de experimentos preliminares que indicaram que se o bosão de Higgs existisse, teria massa num certo intervalo e da afinação destes valores-limites pelo próprio experimento). Finalmente temos de considerar que os dados experimentais podem ser decorrentes da existência de um bosão na faixa procurada ou resultado apenas de movimentos aleatórios de outras partículas que por puro acaso geraram nos aceleradores do CERN o mesmo padrão predito pelo físico irlandês.

Voltemos ao nosso exemplo, o das sondagens eleitorais. Uma boa estimativa é que as várias combinações possíveis C_k^n podem ser aproximadas pelas distribuições normais quando a população e as amostras são grandes. Esta é uma função caracterizada por dois parâmetros (a média μ e o desvio padrão σ) e dada explicitamente por

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} .$$

Esta expressão é conhecida como "gaussiana". A figura 1 mostra o gráfico desta função, e também ilustra o facto de que σ mede o grau de dispersão em torno da média μ. Na figura 2 vemos como os dados experimentais estão próximos do que seria de esperar sem um bosão com massa x, exceto quando x é próximo a 126 GeV.

A probabilidade de termos um resultado que se desvie da média de um valor igual ou superior a 5σ é o integral da gaussiana entre $-\infty$ e $\mu - 5\sigma$ somado ao integral entre $\mu + 5\sigma$ e $+\infty$. Isto dá menos de 10^{-6}. Quando um físico de partículas afirma que uma partícula existe ele quer dizer que a probabilidade de ele ter observado um determinado valor (ou ainda, um valor mais extremo) nos seus experimentos sem que a partícula existisse era inferior a um em um milhão. E é devido a este rigor na produção de enunciados que podemos acreditar mais na equipa do CERN do que nas sondagens em véspera de escrutínio eleitoral!

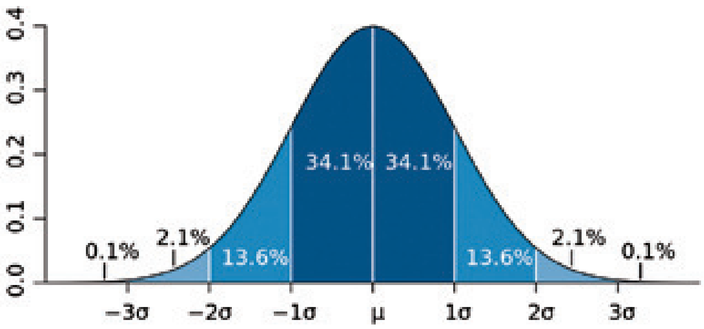

Figura 1: Distribuição normal e probabilidades. Se exigirmos uma precisão de 1σ, então o nível de confiança na rejeição da hipótese nula é de 100% − 2 × 34.1% = 31.8%; o 5σ não é mostrado no gráfico. Fonte: Wikipedia.

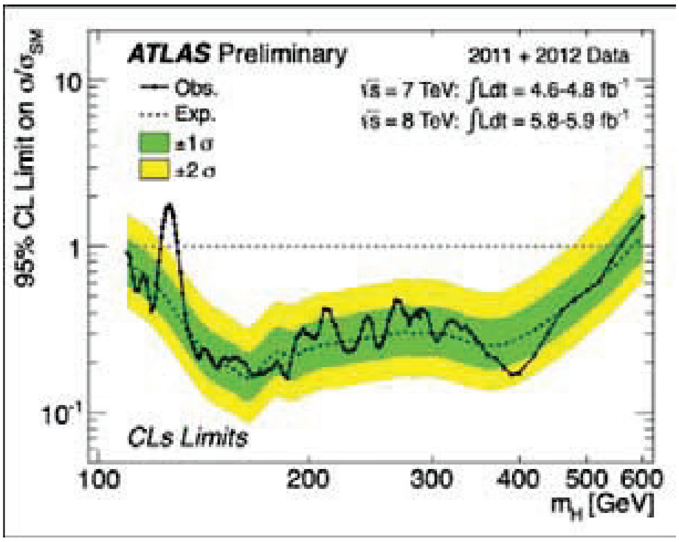

Figura 3: Resultados do experimento Atlas. Na maior parte das energias estudadas o resultado observado pode ser considerado consequência de flutuações aleatórias, exceto próximo aos 126 GeV, onde o desvio chega a 5σ tornando altamente improvável obter estes dados na ausência de um bosão com esta massa aproximada. Fonte: ATLAS Experiment ©2012 CERN

REFERÊNCIA

[1] "Últimos Resultados da Experiência ATLAS na Procura pelo Higgs". Comunicado de imprensa do experimento ATLAS no CERN, disponível em *http://www.atlas.ch/news/2012/HiggsStatementATLAS-PortuguesePortugal.pdf*.