Mentiras, Mentiras do Caraças, e Estatística

Dinis Pestana

Universidade de Lisboa, Departamento Estatística e Investigação Operacional
CEAUL — Centro de Estatística e Aplicações da Universidade de Lisboa[1]
CFCUL — Centro de Filosofia das Ciências da Universidade de Lisboa
Instituto de Investigação Científica Bento da Rocha Cabral

"*Dantes não havia Estatística, tinham que se contentar com mentiras*."

Stephen Leacock

"*Eh pá, tortura esses dados, que eles confessam*."

R. Coase

Pode um livro sobre Estatística tornar-se um *best-seller*? Sem dúvida, *Como Mentir com Estatística* tornou-se quase de imediato um êxito de vendas (sendo, porventura, a excepção que confirma a regra…). Muitos factores explicam esse fenómeno: um título apelativo, ilustrações humorísticas muito bem escolhidas (e estrategicamente mais abundantes nos primeiros capítulos), exemplos quase sempre referindo questões amplamente divulgadas nos meios de comunicação social, um perfume a escândalo — uma disciplina considerada científica afinal é constantemente usada para enganar o público —, pouco volume, e escrita quase terra-a-terra, de uma simplicidade e clareza notáveis. E, por outro lado, as instituições científicas

[1] A investigação do autor é financiada por fundos nacionais através da FCT — Fundação para a Ciência e a Tecnologia, projecto PEst-OE/MAT/UI0006/2011.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR Portugal

de Estatística elogiaram a obra, sendo assim respeitável e séria, mesmo quando faz sorrir (frequentemente!).

Mais de meio século volvido sobre a data de publicação, continua a ser um livro cheio de ensinamentos úteis, numa roupagem divertida. Não admira por isso que a prestigiada revista *Statistical Science* tenha dedicado um fascículo do volume de 2005 a comemorar o cinquentenário da primeira edição, com uma coleção de artigos[2] cuja leitura será menos amena do que a de *Como Mentir com Estatístic*, mas que vale a pena serem lidos, pois claramente demonstram que usar Estatística continua a ser, frequentemente, abusar da Estatística.

Uma parte substancial do livro denuncia que muita divulgação de factos e conclusões, envolvendo-se em roupagem respeitável de ciência (e em particular usando a Estatística para conferir uma aparência de rigor) usa maus dados.

O primeiro capítulo — que continua actual, e é porventura o mais útil, por ir ao cerne de muita fraude feita com uma aparente respeitabilidadeestatística — discute detalhadamente as implicações de erros de amostragem, nomeadamente os que se devem a amostras que não são representativas — comentando também que mesmo operações de amostragem planeadas com rigor facilmente se afastam do padrão de excelência que pretendem, indo as causas de tal acontecer de simples intervenção do acaso a má prática de agentes amostrais ou a tendência de os entrevistados fornecerem respostas "corretas" em vez de verdadeiras.[3]

---

[2] Darrell Huff e Cinquenta Anos de "*Como Mentir com Estatística*" (J. M. Steele).
Mentiras, Cálculos e Construções: Para Além de "*Como Mentir com Estatística*" (J. Best).
Mentir com Mapas (M. Monmonier).
Como Confundir com Estatística, ou: O Uso e Abuso de Probabilidades Concicionais (W. Krämer and G. Gigerenzer).
Como Mentir com Maus Dados (R. D. De Veaux and D. J. Hand).
Como Acusar Outro Tipo de Mentir com Estatística (C. Murray).
Efedra (S. C. Morton).
Em Busca do Laço Mágico: A Verdade sobre o Detetor de Mentiras (S. E. Fienberg and P. C. Stern).

[3] Huff usa, entre outros, um exemplo que na altura era um caso célebre, comentando a recolha de dados dos dois volumes sobre comportamento sexual masculino e feminino elaborado pela equipa liderada por Kinsey (Kinsey, A. C., Pomeroy, W. B., and Martin, C. E. (1948). *Sexual Behavior in the Human Male*, Saunders, Philadelphia; e Kinsey, A. A., Pomeroy, W. B., Martin, C. E, and Gebhard, P. (1953) *Sexual Behavior in the Human Female*, Philadelphia: Saunders.), que também, contra todas as expectativas, foi um escandaloso êxito de vendas (o recente filme sobre o Relatório Kinsey ecoa parte da polémica que esses livros geraram). Penso que é interessante comentar dois factos:

• O segundo capítulo (*Interviewing*) do volume sobre comportamento sexual masculino (que foi o primeiro publicado) discute amplamente questões de amostragem e de condução de entrevistas, e contém muitas observações judiciosas e úteis; os capítulos 3 (*Statistical Problems*) e 4 (*Validity of the Data*) são também uma interessante documental, leitura recomendável para quem queira investigar questões sociais.

• Devido à polémica que esta investigação gerou, foi solicitado um parecer sobre a propriedade de uso de metodologias estatísticas à *American Statistical Association*, que entregou esse trabalho a três dos mais proeminentes estatísticos de sempre, Cochran, Mosteller e Tukey. O relatório que prepararam (Cochran, W. G., Mosteller, F., and Tukey, J. W. (1954). *Statistical Problems of the Kinsey Report*, American Statistical Association, Washington é actualmente difícil de obter na íntegra, sendo no entanto facilmente acessível o texto de conclusões e um dos apêndices técnicos, sobre princípios de amostragem que foram ambos publicados no *Journal of the American Statistical Association*,

Por outro lado, muitas vezes as informações destinam-se a proteger interesses de quem as dá, em vez de pretenderem agradar a quem as ouve. Há muitos anos (andava eu nos meus 18 anos, e estava longe de pensar em Estatística) li no *Diário de Notícias* a deliciosa história de uma prostituta de Viena que aceitou uma bem intencionada boleia, mas quando se dirigiam ao local onde recompensaria o motorista de tão gentil comportamento, tiveram um acidente, que levaram a senhora ao internamento hospitalar durante alguns dias. E não é que a ingrata levou o solícito condutor a tribunal, exigindo desmedida indemnização por perdas no exercício da actividade, estimada a valores de artigo de luxo! E ganhou, e ficou a rir-se, mas não muito tempo, porque as autoridades fiscais austríacas depressa lhe enviaram uma intimação para pagar impostos, com multas e coimas devidas ao não ter declarado o início de actividade e pago os correspondentes impostos, baseando-se na estimativa que ela tinha apresentado a tribunal. Esta historieta ilustra bem que numa questão aparentemente trivial o interessado muda de interesses conforme as circunstâncias, e se num casos exagera os seu proventos, noutros (e especialmente nos contactos com o fisco) há uma enorme tentação de os depreciar.

Amostragem inadequada é fértil húmus para o crescimento de ciência da treta. Se eu disser que pedi a dois colaboradores que entrevistassem 200 cidadãos interrogando-os sobre como iriam votar no referendo sobre liberalização da lei sobre interrupção voluntária da gravidez que na altura ia ocorrer, e que um deles reportou que 99.5% eram favoráveis à alteração legislativa, e o outro que apenas 11% tinham declarado ser favoráveis a essa alteração, a Estatística parece muito pouco fiável; mas se juntar que os primeiros foram entrevistados num jantar de homenagem à Drª Odete Santos (uma deputada que se envolveu fortemente na campanha a favor da liberalização da lei), e que a segunda "amostra" foi recolhida à saída da missa no Mosteiro dos Jerónimos, facilmente se percebe que não é a Estatística que é pouco fiável, quem tomou a decisão de recolher essas amostras enviesadas, em sentidos opostos, é que está a usar mal Estatística.

Há muitos anos dei-me ao trabalho de coleccionar pequenas notícias na revista Expresso, sobre "progressos " na Ciência, numa secção intitulada "Antes do Tempo", se não me falha a memória. A maior parte dessas notícias seguia o tipo de matriz

> *Uma equipa da universidade de ... pediu a 46 voluntários que usassem a mesma roupa durante 24 horas, e que ao tirá-la a colocassem num saco de plástico hermeticamente fechado. O saco era depois dado a cheirar a cada um dos outros 45 indivíduos, pedindo-se que cada um deles classificasse o cheiro de 0 (abominável) a 10 (agradável, inebriante, despertando todos os desejos, masmo os mais secretos e inconfesséveis), concluindo-se que ...*

Este tipo de estudos (ciência da treta, como muito bem são cada vez mais frequentemente qualificados) enferma de vários males, nomeadamente o

---

respectivamente nos volumes **48** (1953), p. 673-715, e **49** (1954), p. 13-35.) é uma primorosa introdução ao bom uso da Estatística na indagação científica, e consequentemente à metodologia da investigação científica, cuja leitura recomendo a todos os profissionais de áreas científicas.

facto de o uso de voluntários muito provavelmente recrutar indivíduos que têm um comportamento ou opções diversas dos da população (por exemplo, em muitas situações são os "exibicionistas" que se oferecem), e o facto de a dimensão da amostra não ter sido cientificamente determinada[4] para os fins em vista (e infelizmente nem sempre há fins em vista a orientar um planeamento experimental, há muita ciência da treta que resulta de um "andar à pesca" do que possa parecer interessante). Este tipo de coleções de dados quase nunca pode ser considerado representativo, e por isso não serve para estabelecer progressos no conhecimento. E, por outro lado, a graduação usando aquela escala ordinal, é muito subjectiva, não tendo cada número o mesmo sentido para todos os respondentes.

O uso de amostras de voluntários (amostras de conveniência — que muitas vezes são muito inconvenientes para a seriedade do estudo que se está a realizar) é comum, mas tirar dessas colecções de dados (repugna-me chamar-lhes amostras) conclusões para uma população é ilegítimo, e só por mera sorte essas conclusões são acertadas.

Alguns anos atrás, um dos telejornais nacionais noticiou a introdução de uma cirurgia em Portugal destinada a aumentar, por recolocação, o "micropénis", causa de infertilidade indesejada em muitos casais. A equipa médica que tinha introduzido essa técnica cirúrgica afirmava que 18% dos portugueses (de sexo masculino, suponho, embora isso não estivesse especificado) têm micropénis — uma estimativa cuja validade parece fortemente contestável, pois não temos conhecimento de nenhum estudo sério sobre o assunto na população portuguesa, e aquela percentagem é tanto mais estranho quanto a estimativca internacionalmente aceite é 0.6%. Provavelmente este erro era muito vantajoso para justificar pedidos de subsídios, mas era tão grosseiro — percentagem 30 vezes maior do que a estimada internacionalmente! — que não duvidamos que nenhuma agência

---

[4] A flutuação amostral é um factor a ter em conta, e o capítulo 3 de Como Mentir com Estatística aborda com exemplos expressivos as conclusões disparatadas ou mesmo fraudulentas que são produzidas por amostras de dimensão insuficiente. Um dos pontos fortes da teoria da amostragem é ensinar-nos como usar uma estimativa da variabilidade do estimador que se vai usar para estimar os parâmetros que nos interessam para determinar quantas obsevações devem ser recolhidas de tal forma que, com uma probabilidade elevada pré-estabelecida, se tenha o grau de aproximação pretendido em relação ao verdadeiro valor do parâmetro. Se não se fizer o trabalho prévio de determinar que dimensão deve ter a amostra, as conclusões não serão porventura legítimas.
Mas por outro lado — e é um assunto pouco esclarecido em muitas referências — apostar em amostras de grande dimensão para ter a certeza que o número de observações necessárias foi excedido, também não é recomendável, porque se cai numa orgia de conclusões sem sentido baseadas em pequenas diferenças, que são significativas de acordo com a teoria estatística, mas de facto totalmente irrelevantes.
O célebre escritor de ficção científica Isac Asimov (que tinha um doutoramento em Bioquímica, boa formação científica, e naturalmente um apreço por Probabilidade bem patente na famosa série “Fundação”, que explora a ideia de ser inventado um ramo da Probabilidade, a que Asimov chama Psico-História, capaz de prever matematicamente evoluções sociais e até certo ponto de controlá-las), escreveu um conto irónico, que de memória penso que se chama *Election Day*, em que o super-computador Multivac precisa de entrevistar um único indivíduo, que escolheu usando um complexo cruzamento de critérios, para determinar, sem eleição, quem seria eleito presidente mundial. É ficção científica, e daquela que creio que nunca se realizará: quanto maior a variabilidade, maior tem que ser a amostra capaz de a representar. Só num fenómeno constante uma amostra de tamanho 1 serviria.

de financiamnto se compadeceu dos cerca de mais de meio milhão de diminuídos. Decerto a estimativa baseou-se numa amostra de indivíduos observados em consultas de infertilidade, e não numa amostra aleatória de portugueses adultos, o que leva a que a indução dessa amostra não representativa para a população em geral não tenha qualquer significado. E, para os leitores que acharam descabido o meu comentário acima "(de sexo masculino, suponho, embora isso não estivesse especificado)", repare que se a amostra fosse contituída por infantes poderia obter-se aquele número. É por estas e por outras que numa investigação científica séria se começa por, face ao enunciado dos propósitos, estabelecer detalhadamente um protocolo de como os dados vão ser recolhidos.

O tamanho do pénis é uma fonte abundante de estudos tolos. Uma das histórias mais saborosas sobre a questão tem que ver com um estudo de antropologia biológica que provava que os bosquímanes do sul de Angola tinham sido mais bem apetrechados pela natureza do que os bosquímanas sul africanos. Mas a diferença estimada era tão elevada que ocasionou dúvidas, e quando a chefe da equipa exibiu a anotação dos dados que tinha recolhido, um conhecedor da geografoa humana de Angola disse estufacto: "A colega recolheu uma amostra maior do que a população!" Quando se foi aclarar o assunto, percebeu-se que

• a "amostra" recolhida era de conveniência, e ainda por cima "comprada", no sentido em que tinham andado veículos do exército português equipados com autofalantes a berrar pelo deserto de Moçâmedes que os membros da população que fossem à cidade seriam recompensados dom uma galinha;

• para que a senhora não entrasse em contacto visual com uma parte tão resguardada da anatomia masculina, foi decidido que dois pobres magalas estariam de plantão com um lençol entre a observadora e o que ela ia medir, pelo que com uma mão segura essa parte da anatomia, com a outra o instrumento de medição (os erros de aproximação assim cometidos eram irrelevantes face ao erro sistemático que adiante se descobre, nesta ehistória tão escondida);

• alguns bosquímanes tiveram a excelente ideia de ir esconder a galinha, e voltar para receber segunda galinha, terceira galinha, etc. Com certeza esse exemplo foi seguido por todos os que se aperceberam que podiam assim alegremente enriquecer o seu património. Mas quando voltavam já sabiam ao que vinham, e para muitos deles estas medições ulteriores não eram em repouso, se me entendem, não era só o património galináceo que aumentava.

Não admira que um membro da audiência tenha exclamado num tom convicto "A colega manipulou os dados". E com alguma desfaçatez, alguém comentou o comentário exclamando com voz expressiva "Os dados?!"

E já que usámos este exemplo, a talhe de foice chamamos a atenção que muitas vezes factos verdadeiros são seguidos de conclusões erradas. Sabemos que a maior prosperidade e capacidade de alimentar melhor as crianças, a altura média das populações tem estado a aumentar; possivelmente foi isso que esteve na base da notícia de que os fabricantes europeu de preservativos tinham decidido fabricá-los com mais um centímetro de comprimento de que anteriormente, reconhecendo que também o comprimento do pénis está a aumentar. É uma interpretação

abusiva, pois o que aconteceu foi que os industriais se aperceberam de que há actualmente muito mais emigrantes africanos, e que devido a campanhas de prevenção de doenças sexualmente transmissíveis, muitos tinham passado a usar preservativo. Note-se, porém, que segundo alguns especialistas, a convicção de que tamanho do pénis e raça estão associados é apenas um preconceito, mas a questão é controversa.

Se Huff fosse vivo, e actualizasse o seu livro, provavelmente teria que incluir muitos novos disparates, infelizmente cada vez mais vulgarizados na prática de má Estatística, mesmo pela comunidade académica, que deveria ter mais preparação e espírito crítico. Refiro-me em particular aos inquéritos que circulam na *net*, com um apelo a que "responda e divulgue", pois é com base nas respostas que o autor vai escrever a sua tese de mestrado ou de doutoramento.

Ora a "não-resposta" é um factor que pode contribuir decisivamente para enviesar resultados. Com este tipo de inquéritos, nem a percentagem de não-resposta é conhecida, e não há qualquer possibilidade de se usar técnicas que possam mitigar esse enviesamento. Um dos mais curiosos não mandava um inquérito, mandava dois, devendo cada qual decidir se respondia ao inquérito para heterossexuais ou ao inquérito para homossexuais, garantido a quase confidencialidade, creio que só o investigador e os seus orientadores teriam conhecimento da escolha de cada respondente. Auguro muito êxito a este trabalho, que decerto vai basear-se em amostras representativas ...

Por outro lado, é de notar que quando os dados são obtidos em entrevistas, há em situações triviais uma tendência reconhecida (que Huff refere) para o respondente dar a resposta que julga que agrada ao entrevistador. Mas em situações menos triviais, em que factores como a interesse, medo, privacidade, embaraço, vergonha, ou mero desconhecimento, levam a erros, mesmo que a amostra seja enorme (ou mesmo que seja um censo da população). Por exemplo, no censo de 2001 um grupo de pressão queria que se perguntasse ao chefe de família "quantos homossexuais e lésbicas há neste agregado familiar?" Se a questão tivesse sido incluída, decerto a percentagem obtida subestimaria fortemente a percentagem real.

E se julga que inquéritos anónimos garantem a veracidade das respostas obtidas, está a avaliar mal a desconfiança de quem não se sente confortável com as questões que são feitas. E há algumas razões para isso — há alguns anos houve no EUA um escândalo que levou a que o responsável pelo organismo nacional de estatística se demitisse: descobriu-se que uma agência governamental de investigação tinha usado marcas de água no papel de respostas que pemitiam identificar quem tinha respondido o quê.

Anote-se, no entanto, que há formas sofisticadas de elaborar inquéitos de forma a levar os respondentes a confiar que não há de facto forma de saber o que responderam. Uma técnica interessante e simples é a da resposta aleatorizada: entrega-se a cada entrevistado um baralho de cartas, e pede-se que proceda a duas extrações sequencias e independentes (isto é, com reposição da primeira carta extraída no baralho), sendo a instruções do tipo:

"*Se a primeira carta que observou era preta, responda*
à pergunta A, se era vermelha responda à pergunta B.
A. Já realizou aborto em condições ilegais?

B. A segunda carta que observou era vermelha?

Sim ____

Não ____ "

Qualquer pessoa que não seja completamente obtusa percebe que nestas circunstâncias quer responda "sim" quer "não" ninguém saberá se a resposta foi à pergunta A ou à pergunta B, eliminando-se assim o temor e embaraço como causas de mentira.

Suponha-se que em 400 entrevistas feitas desta forma se obtêm 113 sins, o que nos leva a estimar (admitindo que a amostra é representativa) que a percentagem de sins na população é 113/400. Por outro lado, denotando p a percentagem desconhecida de mulheres que realizou aborto em condições ilegais, a probabilidade de resposta "Sim" é $\frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times p$. Equacionando as duas expressões, obtém-se a estimativa p=6.5%. É uma forma imaginativa mas simples de obter informação mais fidedigna sobre questões polémicas ou embaraçosas.

Uma questão também muito relevante, em inquéritos, é formular as perguntas de forma clara, para evitar que o respondente as interprete erradamente, e esteja a responder a qualquer completamente diferente do que lhe foi perguntado. E mesmo que nos pareça que não há qualquer ambiguidade na forma como a questão foi colocada, há surpresas, como a daquele senhor a quem perguntaram "*Então, prefere sexo oral ou sexo anal?*" ao que, após alguma hesitação, ele respondeu "*De ano a ano é pouco, mas de hora a hora não consigo!*".

É caro e atrasa, mas um estudo piloto, em que várias destas dificuldades podem ser reconhecidas e corrigidas, compensa, e além disso pode servir para fazer uma estimativa preliminar da variabilidade, que serve para calcular que tamanho deve a amostra ter.

No que se refere a amostragem, há um interessante ponto a sublinhar no que respeita os comentários de Huff: quando analisa os inquéritos sobre comportamento sexual feitos por Kinsey, tem a argúcia de explicitar que nessa investigação a amostragem tem três níveis: na seleção da amostra de inquiridos, na seleção das perguntas a fazer, e na seleção que cada inquirido faz do seu rol de comportamentos, para responder a cada uma das questões. Há uns anos, perguntei por brincadeira na primeira aula de um curso de amostragem se os alunos aceitavem ser classificados por amostragem. A resposta foi unanimemente "não". Depois de folhear as respostas, anunciei esse resultado, e perguntei-lhes se alguma vez tinham tomado consciência de que as questões do exame focavam uma parte (não aleatoriamente escolhida) da matéria do curso, e que dada a limitação de tempo (e recursos, para aqueles que têm tanta confiança no próximo como neles próprios, e nem sempre está próximo quem convem) levava a que cada resposta também fosse em geral uma amostra não aleatória do que sobre cada assunto poderiam responder. Foi uma pergunta retórica, não fiquei à espera de resposta — o que não quer dizer que não tenha as minhas convicções, e como dizia o outro "*Tenho as minhas convicções, não venha perturbar-me com dados*".

Huff naturalmente exemplifica largamente com questões que têm que ver com inquéritos de opinião, porque é um campo em que abundam exemplos ridículos e maliciosos. Mas muitos dos problemas, particulamente os de qualidade, representatividade, tamanho da amostra, colocam-se quer

para estudos de opinião quer para qualquer outro tipo de investigação, e para dados quer qualitativos quer quantitativos. E o mesmo se pode dizer da qualidade dos dados. O capítulo 7 de *Como Mentir com Estatística* começa com uma citação:

> "*Quando for um pouco mais velho" – disse um dia um juiz na Índia a um jovem funcionário britânico - "não citará as estatísticas indianas com um tão grande grau de confiança. O governo está muito interessado em amontoar estatísticas – ele recolhe-as, acrescenta-as, aumenta-as à enésima potência, tira-lhe a raiz cúbica e prepara diagramas maravilhosos. Mas o que nunca deve esquecer é que cada um desses números começou por ser escrito pelo vigia da aldeia (o chowty dar), que regista o que lhe dá na gana."*

Por vezes isso toca as raias do absurdo, o que é fácil de detectar. Há anos um amigo meu estava interessado no estudo de riscos em operações por laparascopia ao útero. As paredes do útero começam por ser afastadas enchendo-o de uma solução salina, e por isso dá-se osmose de iões Na^+, que migram para o sistema circulatório e podem causar picos de hipertensão (devo estar a dizer uma carrada de asneiras, mas isso não estraga a história). Por isso ele usou como variáveis resposta a natrémia — concentração de iões na+ no sangue — no fim da operação, uma hora depois, duas horas depois, três horas depois, 8 horas depois, 24 horas depois e 48 horas depois, para ver se o tempo que a operação tinha durado (e consequentemente as biomembranas uterinas tinham estado sujeitas ao contacto com a solução salina) influenciavam as variáveis resposta. Para ter dados suficientes, encarregou-se de toda a parte trabalhosa (análises do sangue), mas recorreu aos colegas da maternidade pedindo-lhe que fornecessem o tempo de duração da operação. Alguns apontaram minutos e segundos, outros minutos apenas, e um apontou constantemenete 15 minutos ou 20 minutos. Este padrão salienta-se imediatamente, e percebe-se logo que as medições foram imprecisas. Mas se alguém falsear todos os dados com cuidado, é difícil perceber (apesar de haver uma "lei de Benford" que estabelece que em muitos fenómenos quantitativos reais os números contêm mais algarismo 1 do que 2, e mais 2 que 3, etc., numa escala aproximadamente logarítmica, enquanto os aldrabões para fingir que os dados são reais têm tendência a dar a todos os algarismos a mesma frequência).

Mas mesmo sem aldrabice, há enganos de registo e de transcrição, que podem ser muito difíceis de detectar. Quando o Reitor da UL recebeu um ofício do Presidente do Conselho Científico da FCUL em que se dizia "*Junto tenho a honra de enviar cópia de um peido do Professor Doutor N..., que mereceu aprovação unânime da comissão coordenadora do Conselho Científico*", não ficou decerto a pensar se o que merecia essa aprovação era o original ou a cópia, nem ficou deslumbrado com o avanço tecnológico que tinha permitido copiar algo tão etéreo. Provavelmente percebeu logo que era um "pedido" que tinha perdido o "d". E não é preciso explicar a ninguém a história do prior da ordem que, quando um noviço lhe disse que estava a copiar a regra da ordem de uma cópia, e que não copiando do original qualquer erro se ia perpetuar; percebendo que o comentário tinha razão de ser foi verificar a cópia do noviço pelo original, e não havia meio de voltar

dos subterrâneos. Quando os outros monges começaram a ficar aflitos e foram procurá.lo, estava com a nuca toda ensanguentada por ter arrancado punhados de cabelo, e ainda estava a dar bofetadas a si mesmo, e a gemer "*Era regra de caridade, não era regra de castidade! era regra de caridade, não era regra de castidade!"* Omissão de letras, duplicação de letras, troca de letras — todas essas gralhas são em geral fáceis de corrigir, porque há um contexto que permite em geral uma descodificação rápida.

Se tiver um registo do tipo 37, 65, 49, 53, 71, 42, 466, 39, 45,…, "sei" que o valor 466 está errado, e até suponho que o erro foi uma duplicação do algarismo 6; com a série 795, 810, 730, 735, 830, 756, 910,…, parece descortinar-se um padrão que leva a pensar que 756 deve ser corrigido para 765; com 475, 424, 437, 409, 68,…, eu corrigiria 68 para 468 — mas sempre com o cuidado de ir verificar se há registos em apontamentos que apoiem essas coreções. Mas ler números (e a Estatística é, também, a ciência que ensina a ler números) é uma ciência e arte mais complexa e ambígua, em que nem sempre á fácil descortinar padrões. A operação inicial de "limpeza" dos dados é demorada e cara, De Veaux and Hand[5] afirmam que pode se pode estimar que os recursos gastos num projecto com a limpeza dos dados são 60% a 95% do total! E citam trabalhos de outros autores, indicando que mesmo com todos os cuidados 1% a 10% dos dados em bases de dados de grandes organizações estão errados.

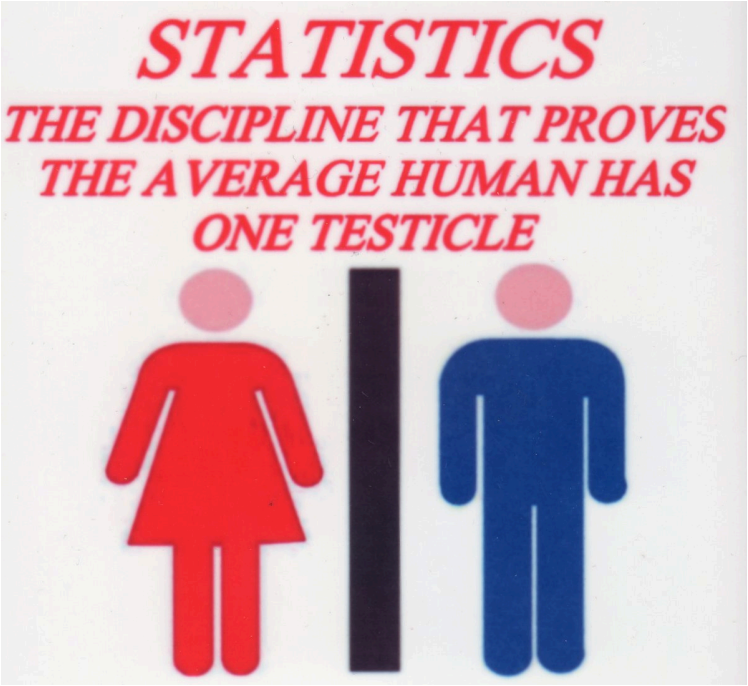
Creio que já estamos todos cansados de chover no molhado dos erros de amostragem, um tema que é abordado de diversos ângulos em diferentes capítulos deste livro. Passando a outro tema, convém anotar que no capítulo 2 Huff aborda a questão de como resumir os dados com um único número. Em inglês é fácil transmitir a ideia, porque o termo "*average*" tem um sentido vago, e tanto pode ser média como mediana, ou mesmo moda. Em português, o termo média é quase exclusivamente interpretado como média aritmética. O leitor deve abordar esse capítulo tomando um conceito abrangente do que é média (por exemplo, o de que média é um valor intermédio da amostra, entre mínimo e máximo, podendo mesmo coincidir com qualquer deles[6]). Se o leitor aceitar que *média=average* é um *valor intermédio*, o texto torna-se mais claro.

Qualquer valor intermédio é uma indicação da localização dos dados, mas é evidente que a indicação de um valor intermédio é uma informação muito escassa; Huff discute muito bem como um indicador da escala é essencial para perceber corretamente a informação que qualquer média transmite. De qualquer forma, a estafada ideia de que a média pode ser completamente inadequada está bem expressa no esmalte de uma caixa que um amigo me ofereceu quando me aposentei, e que reproduzo para os olhos do leitor.

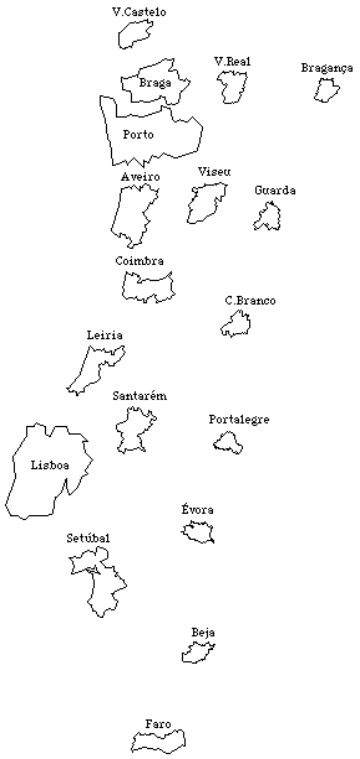
---

[5] Em "How to Lie with Bad Data", *Statistical Science* **20** (2005), p. 231-238.

[6] Espero que esta ideia, que creio que se pode reportar ao grande matemático francês Cauchy, não choque o leitor. De facto, em estatística usamos médias aritméticas, médias geométricas, médias harmónicas — que qualquer delas são casos especiais 1, 0 e -1 do que se chama média de ordem p — médias ponderadas, médias aparadas a αx100%, de que um caso especial é a mediana, etc.



Huff discute também muito bem o mau uso de gráficos, sejam eles pictogramas ou mapas. Os gráficos são muito sugestivos, e por isso podem ser facilmente usados para sugerir trapalhices. Veja-se por exemplo o mapa "explodido" de Portugal, em que cada distrito foi reescalado para indicar o número de doentes mentais desse distrito. Beja é um distrito tão pequeno, que quase apetece dizer: "se quer evitar ter uma doença mental, vá viver para Beja". Claro que o que está em causa é o distrito de Beja ter uma densidade populacional inferior à de outros distritos, se se usasse como escala o número de doentes mentais por cada 1000 habitantes provavelmente o mapa seria muito mais parecido com aquilo a que estamos habituados.



Se o leitor estiver interessado em informação mais detalhada sobre bons e maus gráficos, recomendo a leitura do interessante livro de Tufte[7].

Outra questão que Huff explora magnificamente é o das correlações espúrias, a mania de relacionar tudo com tudo. Apesar de apreciar devidamente a proposta de Karl Pearson, no seu notável livro *The Grammar of Science*, de 1896, de que que o paradigma de Ciência devia mudar, completando a busca de causalidade com a deteção de associações estatísticas, isso levado ao exagero levou um grupo de brincalhões de Lisboa a há algumas décadas fingir que Portugal estava a ser visitado por um génio da Ciência, e numa conferência de imprensa foi muito gabada a sua teoria que conseguia exactamente correlacionar qualquer variável com qualquer outra, apenas com uma excepção: aproximar o baixo contínuo do Peixinho (o compositor Peixinho, obviamente, pois baixo contínuo é uma expressão musical). Ah, mas durante esta visita essa aproximação tinha sido feita — e então exibiam um aquário com um peixinho, e chamavam um senhor baixinho, vestido de contínuo, e incitavam-no: "*Vá, aproxime-se do peixinho*".

A correlação mais disparatada foi-me inspirada por um conto de Graham Greene, no delicioso livro Empresta-nos o seu Marido, e Outras Comédias da Vida Sexual. Nesse conto, escrito na primeira pessoa, o narrador confessa que nunca conheceu os prazeres do amor, porque nunca tinha conseguido vencer o medo que o médico do colégio em que estudara lhe transmitira, com a teoria de que as relações sexuais causam cancro, pois 100% das pessoas que morrem de cancro ou particaram relações sexuais ou são filhas de pessoas que praticaram relações sexuais. Graham Greene é aliás um Mestre, e faz em *Viagens com a Minha Tia* uma descrição que se pode aplicar à Estatística que ecoa as palavras de Laplace, de que a Probabilidade era o bom senso transcrito em linguagem matemática (uma personagem, que escreve sucessões de números do tipo

28 de Julho

07:15	17 s	
10:45	37 s	
12:30	50 s	
13:15	32 s	
13:40	50 s	
14:05	20 s	
15:45	37 s	
18:40	28 s	
10:30	?	(esqueci-me de cronometrar)
Total	4m 31 s	

acaba por confidenciar que aponta as horas a que mija, e a duração, e depois junta:

> "*É só multiplicar por sete. Dá meia hora por semana. Vinte e seis horas por ano. Claro que a*

[7] Tufte, E. R. (1983). *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, Conn.

vida a bordo não serve de bitola. Bebe-se muito entre as refeições. E está-se sempre a beber cerveja ... Veja este tempo aqui: 1 minuto e 55 segundos. É mais do que a média, mas eu ao lado tinha anotado dois gins. Há também uma data de variações com que eu não tinha contado e, daqui em diante, vou passar a tomar também nota da temperatura. Veja aqui, 25 de Julho: 6 minutos e 9 segundos inc., quer dizer, incompleto... Fui jantar fora em Buenos Aires e deixei o livro de apontamentos em casa. E aqui, no dia 27 de Julho: só 3 m e 12 s ao todo, mas se se lembrar soprava um vento forte e frio no dia 25 e eu saí para jantar sem sobretudo…

Os abusos da Estatística, quer os denunciados por Huff quer os gozados pelo grupo de foliões de Harvard que atribui anualmente prémios IgNobel (e se o leitor se quer divertir procure a informação que disponibilizam na internet) acabam por dar mau nome à Estatística, e a profissão de estatístico é muitas vezes atacada. Por exemplo (não sei quem é o autor, mas a frase é interessante)

Os estatísticos são os únicos profissionais que exigem que lhes seja reconhecido o direito de mentir em 5% dos casos.

Mas possivelmente Stephen Fry tem toda a razão quando faz o elogio da mentira

"Padres e psicoterapeutas podem julgar que a verdade é revelada nos confessionários ou nas sessões de psicanálise, mas você sabe, e eu sei, e todos os homens sabem que mentimos constantemenet a toda a gente. Mentir é tão costumeiramente humano quanto usar roupas. […] *Mentir é parte essencial da condição humana. Eliminar a mentira não faria de nós mais humanos; pelo contrário, amputaria parte da nossa huma-nidade."*

Stephen Fry, *The Liar*, 1991, Mandarin, p. 306.

E denuncia que afinal os humanos estão sempre a mentir, e não apenas em 5% dos casos. Por outro lado, é amplamente reconhecido que as decisões quer institucionais quer pessoais são feitas sob incerteza, como judiciosamente se comenta num editorial da prestigiada revista *Science*:

É embaraçoso constatar que uma elevada proporção das decisões fulcrais do Governo, de negociação de tratados à gestão dos recursos, são tomadas com base em informação insuficiente e hipóteses não comprovadas. E isto não é assim tão diferente da forma como conduzimos os nossos assuntos privados.

(Editorial de *Science*)

Ora são Probabilidade e Estatística as ciências que nos ensinam a disciplinar a incerteza. E fazemos isso com maus dados, há uma área recente da Estatística denominada *Analysis of Messy Data*, que se preocupa especialmente com os dados que não se conformam com o padrão de ouro

clássico, de que a amostra aleatória de que a amostra observada é uma mera observação é um vector de margens independentes e identicamente distribuídas, e adoptando um modelo simples e matematicamente tratável, pois mesmo quando adoptamos esse modelo (sabendo que os modelos são para ser úteis, não para acreditarmos neles) sabemos que os dados são mal comportados, como alguém — não sei quem, e lamento — descreveu jocosamente:

> *O Teorema ds Dados Estuporados diz que na vida real os dados tendem a provir de distribuições bizarras e mal especificadas de variáveis aleatórias fortemente correlacionadas, obtendo-se em geral amostras de dimensões diversas, com dados em falta, observações dependentes, e um número indeterminado de valores mal registados.*

O livro de Huff é perfeito? Não, a meu ver tem muita infromação sobre amostragem, não abordando da mesma forma a outra disciplina de aquisição de dados, o planeamento experimental. E actualmente seria bom que houvesse se multiplicassem as denúncias da muita ciência da treta que se produz, também por culpa dos académicos e do falso ideal "*publish or perish*", que tem posto em causa a integridade da produção científica[8], e mesmo a forma desprestigiante como a ciência se arrisca a ser vista pelo público[9]. O já referido grupo de humoristas de Harvard edita também um jornal que é um mimo, em que aparecem caricaturas que infelizmente são um espelho de alguma "ciência" que se faz abusando da Estatística. Por exemplo, o que surgiu primeiro, o ovo ou a galinha?

Este mistério foi finalmente resolvido: escolheram-se 100 aviários ao acaso, e dois dos experimentadores fizeram telefonemas simultâneos, encomendando por via postal um deles uma galinha, devidamente assassinada,

---

8 Recomendo a leitura de "Integrity Under Attack:The State of Scholarly Publishing" (D. N. Arnold) http://www.ima.umn.edu/ arnold//siam-columns/ integrity-under-attack.pdf

9 Não é decerto por acaso que é muito citado o artigo "Why Most Published Research Findings Are False" (John P. A. Ioannidis) que pode consultar-se em http://www.plosmedicine.org/article/info:doi/10.1371/ journal.pmed.0020124

depenada, esventrada, e outro encomendando uma dúzia de ovos, a serem enviados por correio para uma dada morada. E a Estatística não deixa qualquer dúvida: em 97% dos casos a dúzia de ovos surgiu primeiro!