





RUÍDO



DANIEL KAHNEMAN  
OLIVIER SIBONY  
CASS R. SUNSTEIN

# RUÍDO

*Porque tomamos más decisões  
e como podemos evitá-lo*

Tradução de  
ISABEL VERÍSSIMO





*Para Noga, Ori e Gili — DK*  
*Para Fantin e Lélia — OS*  
*Para Samantha — CRS*

# ÍNDICE

---

<i>Introdução: Dois tipos de erro</i>	11
<b>PARTE I: ENCONTRAR O RUÍDO</b>	21
1. Crime e castigo ruidoso	23
2. Um sistema ruidoso	35
3. Decisões singulares	49
<b>PARTE II: A MENTE É UM INSTRUMENTO DE MEDIÇÃO</b>	55
4. Questões de juízo	59
5. Medir o erro	73
6. A análise do ruído	89
7. Ruído ocasional	101
8. De que forma os grupos ampliam o ruído	119
<b>PARTE III: RUÍDO EM JUÍZOS PREDITIVOS</b>	135
9. Juízos e modelos	139
10. Regras silenciosas	153
11. Ignorância objectiva	169
12. O vale da normalidade	181
<b>PARTE IV: COMO ACONTECE O RUÍDO</b>	195
13. Heurísticas, enviesamentos e ruído	197
14. A operação de associação	215
15. Escalas	229
16. Padrões	245
17. As fontes de ruído	257



<b>PARTE V: MELHORAR OS JUÍZOS</b>	269
18. Melhores juízes para melhores juízos	273
19. Remoção de enviesamentos e higiene de decisão	287
20. Sequenciar informações nas ciências forenses	299
21. Selecção e agregação de prognósticos	317
22. Directivas em medicina	335
23. Definir a escala nas avaliações de desempenho	351
24. Estrutura na contratação	367
25. O protocolo de avaliações mediadoras	383
<b>PARTE VI: RUÍDO IDEAL</b>	399
26. Os custos da redução do ruído	405
27. Dignidade	417
28. Regras ou padrões?	431
<i>Análise e Conclusão: Levar o ruído a sério</i>	445
<i>Epílogo: Um mundo menos ruidoso</i>	463
<i>Apêndice A: Como realizar uma auditoria de ruído</i>	465
<i>Apêndice B: Uma lista de verificação para um observador de decisões</i>	473
<i>Apêndice C: Corrigir previsões</i>	477
<i>Agradecimentos</i>	483
<i>Notas</i>	485



## INTRODUÇÃO

# Dois tipos de erro

Imagine que quatro grupos de amigos vão a um clube de tiro. Cada grupo forma uma equipa composta por cinco pessoas. Os membros da equipa partilham uma espingarda e cada um dispara um tiro. A Figura 1 mostra os resultados.

Num mundo ideal, todos os tiros acertariam no centro do alvo.

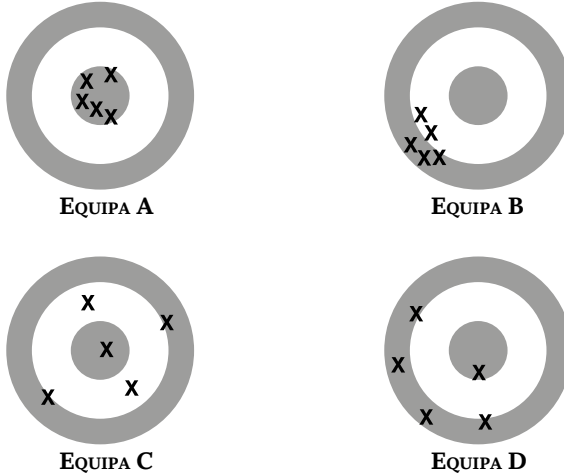


Figura 1: Quatro equipas

O resultado da Equipa A não se afasta muito do objectivo. Os tiros desta equipa estão muito agrupados à volta do centro, num padrão que é quase perfeito.

Dizemos que a Equipa B apresenta um *enviesamento*<sup>1</sup>, porque os seus tiros falham sistematicamente o alvo. Como a figura ilustra, a consistência do enviesamento apoia uma previsão. Se um dos membros da equipa disparasse mais um tiro, apostaríamos que acertaria na mesma área que os primeiros cinco. A consistência do enviesamento também convida a uma explicação causal: talvez a mira da espingardada daquela equipa estivesse desalinhada.

Chamamos à Equipa C *ruidosa*, porque os seus tiros estão muito espalhados. Não existe um enviesamento óbvio, pois os impactos estão aproximadamente centrados no alvo. Se um dos membros da equipa disparasse mais um tiro, não conseguiríamos prever com exactidão onde acertaria. Além disso, não nos ocorre uma hipótese interessante para explicar os resultados da Equipa C. Sabemos que os seus membros são maus atiradores. Não sabemos porque são tão ruidosos.

A Equipa D apresenta enviesamento e ruído. Como a Equipa B, os seus tiros estão sistematicamente desviados do centro do alvo; como a Equipa C, os seus tiros estão muito espalhados.

Porém, este livro não é sobre tiro ao alvo. O nosso tópico é o erro humano. Enviesamento e ruído — desvio sistemático e dispersão aleatória — são diferentes componentes do erro. Os alvos ilustram a diferença.

A carreira de tiro é uma metáfora para o que pode correr mal no juízo humano, sobretudo nas diferentes decisões que as pessoas tomam em nome de organizações. Nestas situações, encontraremos os dois tipos de erro exemplificados na Figura 1. Alguns juízos apresentam enviesamento; estão sistematicamente desviados do alvo. Outros juízos são ruidosos, pois, se esperávamos que estivessem de acordo, as pessoas que os emitem afastam-se do centro do alvo.

---

<sup>1</sup> Em estatística, «enviesamento» é uma medida da assimetria da distribuição dos dados. (N. da T.)

Lamentavelmente, há muitas organizações que são afectadas por juízos enviesados e ruidosos.

A Figura 2 ilustra uma importante diferença entre enviesamento e ruído. Mostra o que veríamos na carreira de tiro se nos fosse apresentada apenas a parte de trás dos alvos de que as equipas se serviram, sem qualquer indicação do centro onde tentavam acertar.

No verso do alvo é impossível saber se é a Equipa A ou a Equipa B que está mais perto do centro do alvo. Contudo, é fácil perceber que a Equipa C e a Equipa D são ruidosas, ao contrário da Equipa A e da Equipa B. Na realidade, tanto a Figura 1 como a Figura 2 dão-nos a mesma ideia sobre dispersão. Uma característica geral do ruído é que é possível reconhecê-lo e medi-lo, apesar de nada sabermos a respeito do alvo ou do enviesamento.

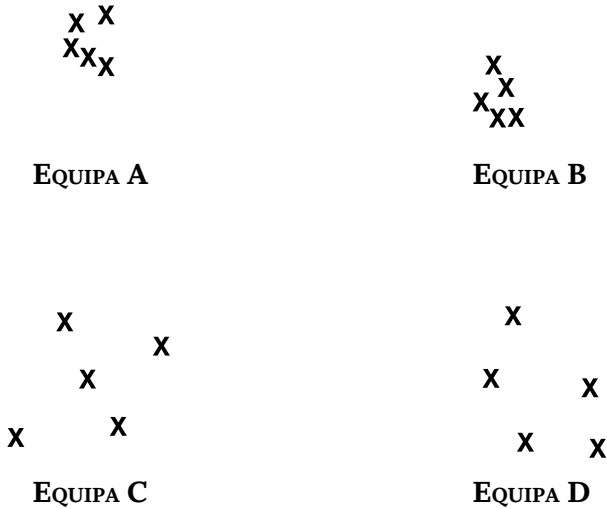


Figura 2: *Perspectiva da parte de trás do alvo*

A característica geral do ruído que acabámos de referir é essencial para os nossos propósitos neste livro, pois muitas das nossas conclusões são inferidas de juízos cuja verdadeira resposta é desconhecida ou até incognoscível. Quando médicos apresentam

diferentes diagnósticos ao mesmo doente, podemos estudar a sua discordância sem sabermos de que padece a pessoa. Quando os executivos de um estúdio cinematográfico avaliam o mercado tendo em vista o lançamento de um filme, é possível estudar a variabilidade das respostas sem fazermos ideia de quanto é que o filme rendeu ou se chegou a ser produzido. Não precisamos de saber quem tem razão para medir a variação de juízos referentes ao mesmo caso. Para medir o ruído temos apenas de olhar para a parte de trás do alvo.

Para compreender o erro no juízo é essencial entender as noções de enviesamento e ruído. Por vezes, como veremos, ruído é o problema essencial. No entanto, no debate público sobre o erro humano e em organizações no mundo inteiro o ruído raramente é reconhecido. O enviesamento é a estrela da companhia. O ruído é um figurante, que normalmente nem sequer entra em cena. O tópico do enviesamento foi discutido em milhares de artigos científicos e em dúzias de livros para o público geral, mas a questão do ruído é referida poucas vezes. Este livro é a nossa tentativa de corrigir esse desequilíbrio.

Em decisões concretas, a quantidade de ruído é muitas vezes escandalosamente elevada. Seguem-se alguns exemplos da alarmante quantidade de ruído em situações onde a precisão é importante:

- *A medicina é ruidosa.* No diagnóstico de um mesmo doente, diferentes médicos formulam diferentes juízos sobre se este sofre de cancro de pele, cancro da mama, doença cardíaca, tuberculose, pneumonia, depressão e uma série de outras doenças. O ruído é especialmente intenso na psiquiatria, uma área onde o juízo subjectivo é sem dúvida importante. Contudo, também nos deparamos com um considerável ruído em áreas onde tal não seria expectável, como na análise de exames imagiológicos.
- *As decisões sobre a custódia de menores são ruidosas.* Os assistentes sociais dos serviços de protecção de crianças e jovens

têm de avaliar se as crianças estão em risco de abuso e, em caso afirmativo, se devem ser retiradas à família. O sistema é ruidoso, dado que alguns assistentes sociais apresentam maior tendência que outros para retirar a criança à família e colocá-la em acolhimento familiar ou numa instituição. Anos mais tarde, um número maior das desafortunadas crianças que foram colocadas em famílias de acolhimento ou em instituições por estes técnicos inflexíveis têm vidas mais complicadas: taxas de delinquência mais elevadas, taxas de natalidade mais altas na adolescência e rendimentos mais baixos.

- *Os prognósticos são ruidosos.* Os analistas profissionais apresentam previsões extremamente variáveis sobre o provável volume de vendas de um novo produto, sobre o crescimento provável da taxa de desemprego, sobre a probabilidade de falência de empresas em dificuldades e sobre praticamente tudo o resto. Não só discordam uns dos outros, como também entram em contradição consigo mesmos. Por exemplo, quando foi pedido aos mesmos programadores informáticos em dois dias diferentes que calculassem o tempo necessário para completar a mesma tarefa, o número de horas apresentado diferiu, em média, 71%.
- *As decisões de concessão de asilo são ruidosas.* A entrada de um requerimento de asilo nos Estados Unidos depende de um processo semelhante a uma lotaria. Um estudo de casos atribuídos de forma aleatória a diferentes juizes concluiu que um juiz autorizou a entrada de 5% dos requerentes e outro autorizou a entrada de 88%. O título do estudo diz tudo: «A Roleta dos Refugiados.» (Veremos à frente muitos exemplos de roleta.)
- *As decisões de contratação de pessoal são ruidosas.* Os entrevistadores de candidatos a emprego fazem avaliações muito diferentes das mesmas pessoas. As avaliações de desempenho

desses funcionários também são extremamente variáveis e dependem sobretudo da pessoa que as elabora e não tanto do desempenho que está a ser analisado.

- *As decisões para o estabelecimento de uma fiança são ruidosas.* Depende em parte do juiz seleccionado se uma pessoa acusada pode pagar uma fiança e aguardar o julgamento em liberdade ou se tem de ficar em prisão preventiva. Alguns juízes são muito mais clementes que outros. Os juízes também diferem claramente na avaliação dos arguidos que apresentam maior risco de fuga ou de reincidência.
- *As ciências forenses são ruidosas.* Fomos ensinados a pensar que a identificação através de impressões digitais é infalível. Porém, ocasionalmente, os técnicos que as examinam divergem quando estão a decidir se uma impressão digital encontrada num local de crime corresponde à de um suspeito. Não só os especialistas discordam, como por vezes os mesmos especialistas tomam decisões inconsistentes ao serem confrontados com a mesma impressão digital em diferentes ocasiões. Uma variabilidade semelhante está documentada noutras áreas das ciências forenses, até mesmo na análise de ADN.
- *As decisões para a concessão de patentes são ruidosas.* Os autores de um importante estudo sobre pedidos de patentes realçam o ruído envolvido: «A aceitação ou rejeição de uma patente pelo instituto da propriedade industrial está significativamente relacionada com o examinador que é escolhido para analisar o pedido.» Esta variabilidade é sem dúvida problemática do ponto de vista da equidade.

Todas estas situações ruidosas são a ponta de um grande icebergue. Para onde quer que olhemos nos juízos humanos, é provável encontrarmos ruído. Para melhorarmos a qualidade dos nossos juízos, temos de ultrapassar o ruído e o enviesamento.



Este livro está dividido em seis partes. Na Parte I, exploramos a diferença entre ruído e enviesamento e mostramos que as organizações públicas e privadas podem ser ruidosas, por vezes de forma chocante. Para avaliar o problema, começamos com juízos em duas áreas. A primeira envolve sentenças do foro criminal (daí o sector público) e a segunda envolve seguros (daí o sector privado). À primeira vista, estas duas áreas não poderiam ser mais diferentes. Porém, no que diz respeito ao ruído, têm muito em comum. A apoiar este ponto de vista, introduzimos a ideia de uma auditoria do ruído, concebida para medir a quantidade de discordância entre profissionais tendo em conta os mesmos casos no seio de uma organização.

Na Parte II investigamos a natureza do juízo humano e exploramos a forma de estabelecer a margem de precisão e de erro. Os juízos são susceptíveis de enviesamento e ruído. Descrevemos uma notável equivalência nos papéis dos dois tipos de erro. Ruído ocasional é a variabilidade de juízos formulados para o mesmo caso pela mesma pessoa ou grupo em diferentes ocasiões. Uma quantidade surpreendente de ruído ocasional acontece em discussões de grupo por factores aparentemente irrelevantes, como, por exemplo, quem é o primeiro interveniente a falar.

A Parte III aprofunda um tipo de juízo que tem sido muito estudado: o juízo predictivo. Exploramos a vantagem fundamental de regras, fórmulas e algoritmos em relação aos seres humanos para fazer previsões: ao contrário da crença popular, não é tanto a perspicácia superior das regras, mas o facto de estarem isentas de ruído. Abordamos o limite fundamental da qualidade do juízo predictivo — a ignorância objectiva do futuro — e como conspira com o ruído para limitar a qualidade da previsão. Por fim, abordamos uma questão que terá quase de certeza colocado a si mesmo nesta altura: «Se o ruído é tão ubíquo, porque não reparei nele antes?»

A Parte IV centra-se na psicologia humana. Explicamos as causas principais do ruído. Entre elas contam-se as diferenças interpessoais resultantes de uma série de factores, incluindo

a personalidade e o estilo cognitivo, variações idiossincráticas na ponderação de diferentes considerações e os diferentes usos que as pessoas dão às mesmas escalas. Exploramos o que leva as pessoas a estarem alheadas do ruído e não ficarem surpreendidas com acontecimentos e juízos que não podiam ter previsto.

A Parte V explora a questão prática de como podemos melhorar os juízos e evitar o erro. (Os leitores que estão acima de tudo interessados nas aplicações práticas da redução de ruído poderão preferir saltar a discussão dos desafios da previsão e da psicologia do juízo nas Partes III e IV e passar directamente para esta parte.) Investigamos os esforços para resolver o ruído na medicina, nas empresas, no ensino, no governo e noutras áreas. Propomos diversas técnicas de redução do ruído, agrupando-as sob o rótulo de *higiene de decisão*. Apresentamos cinco estudos de caso em áreas onde existe muito ruído documentado e onde as pessoas fazem esforços sustentados para reduzi-lo, com diversos níveis de sucesso, que são elucidativos. Os estudos de caso incluem diagnósticos médicos, avaliações de desempenho, ciências forenses, decisões de contratação e prognósticos em geral pouco fiáveis. Concluimos com a apresentação de um sistema a que chamamos *protocolo de avaliações mediadoras*: uma abordagem polivalente da avaliação de opções que incorpora diversas práticas fundamentais de higiene de decisão e procura produzir juízos menos ruidosos e mais fiáveis.

Qual é o nível adequado de ruído? A Parte VI aborda esta questão. Talvez de forma contra-intuitiva, o nível adequado não é zero. Em algumas áreas, não é exequível eliminar o ruído. Noutras, é demasiado dispendioso. Noutras ainda, os esforços para reduzir o ruído comprometeriam importantes valores concorrentes. Por exemplo, as tentativas de eliminar ruído podem enfraquecer a moral e dar às pessoas a sensação de que estão a ser tratadas como peças da engrenagem de uma máquina. Quando os algoritmos fazem parte da resposta, suscitam uma série de objecções, e vamos falar sobre algumas delas aqui. No entanto, o nível actual de ruído

é inaceitável. Recomendamos às organizações privadas e públicas que realizem auditorias de ruído e que se esforcem mais, com uma seriedade sem precedentes, para reduzir o ruído. Se o fizerem, as organizações poderão diminuir a injustiça generalizada — e reduzir custos em muitas áreas.

Com essa ambição em mente, terminamos cada capítulo com algumas breves propostas ao estilo de citação. O leitor poderá usar essas afirmações *ipsis verbis* ou adaptá-las a quaisquer questões que sejam importantes para si, quer estejam relacionadas com saúde, segurança, educação, dinheiro, emprego, entretenimento ou outra coisa. Compreender o problema do ruído, e tentar resolvê-lo, é um trabalho em curso e um esforço colectivo. Todos temos o ensejo de contribuir para este trabalho. Escrevemos este livro com a expectativa de podermos aproveitar as oportunidades que surjam.



## PARTE I

---

# ENCONTRAR O RUÍDO

Não é aceitável que pessoas em tudo semelhantes, condenadas pelo mesmo crime, acabem por receber sentenças drasticamente diferentes — digamos, cinco anos de prisão efectiva para uma e pena suspensa para outra. Todavia, a verdade é que situações parecidas com a que usámos neste exemplo acontecem em muitos lugares. É certo que o sistema de justiça criminal também está impregnado de enviesamento, mas no capítulo 1 vamos focar-nos no ruído — e em particular no caso que um famoso juiz trouxe para a ribalta, por o considerar escandaloso, e lançando uma cruzada que de certa forma mudaria o mundo (se bem que não o suficiente). A nossa história passa-se nos Estados Unidos, mas acreditamos que histórias similares acontecem (acontecerão) em muitas outras nações. Em algumas dessas nações, é provável que o problema do ruído seja pior ainda que nos Estados Unidos. O que pretendemos ao usar o exemplo das sentenças judiciais é em parte demonstrar que o ruído pode produzir uma grande injustiça.

As sentenças do foro criminal apresentam um nível de drama particularmente elevado, mas a nossa preocupação estende-se ao sector privado, onde os riscos também podem ser significativos. Para exemplificar este ponto de vista, no capítulo 2 analisamos uma grande companhia de seguros. Nela, os técnicos de seguros têm a tarefa de estabelecer prémios para potenciais clientes e os avaliadores devem decidir o valor das reclamações. Seria previsível que estas

tarefas fossem simples e mecânicas e que diferentes profissionais chegassem sensivelmente aos mesmos valores. Para testar essa previsão, levámos a cabo uma experiência cuidadosamente estruturada — uma auditoria de ruído. Os resultados surpreenderam-nos, mas, mais importante, chocaram e consternaram os administradores da companhia de seguros. Ficámos a saber que o nível de ruído está a custar muito dinheiro à empresa. Usamos este exemplo para demonstrar que o ruído pode causar grandes perdas económicas.

Estes dois exemplos envolvem estudos com um grande número de pessoas, e essas mesmas pessoas formulam uma grande quantidade de juízos. Contudo, muitos juízos importantes são *singulares*, não são repetidos: o que fazer perante uma oportunidade de negócio aparentemente única, o lançamento ou não de um produto novo, como lidar com uma pandemia, contratar ou não uma pessoa que não corresponde ao perfil-padrão. É possível encontrar ruído em decisões únicas como estas? É tentador pensar que o ruído está ausente nestes casos. Afinal de contas, se o ruído significa uma variabilidade indesejada, como pode existir variabilidade em decisões singulares? No capítulo 3 tentamos responder a esta pergunta. O juízo formulado, mesmo numa situação aparentemente única, é apenas um numa nuvem de possibilidades. Também encontraremos muito ruído nestes casos.

O tema que emerge nestes três capítulos pode ser resumido numa frase, que será um tema central deste livro: *sempre que há um juízo, há ruído — mais do que seria de esperar*. Vamos começar por descobrir se é assim tanto.

## CAPÍTULO 1

---

# Crime e castigo ruidoso

Suponha que alguém foi condenado por um crime — roubo em lojas, posse de estupefacientes, agressão ou assalto à mão armada. Qual será a sentença?

A resposta não deveria depender do juiz a quem o caso é atribuído, das condições climatéricas ou da vitória da equipa local no jogo da véspera. Seria chocante se três arguidos em condições semelhantes, condenados pelo mesmo crime, recebessem penas radicalmente diferentes: pena suspensa para um, dois anos de prisão para outro e dez anos de prisão para um terceiro. No entanto, uma tal atrocidade pode ocorrer em muitos países — não apenas no passado distante, mas também nos nossos dias.

Em todo o mundo, desde tempos imemoriais que os juízes têm usufruído de um poder discricionário para decidirem sentenças adequadas. Em certos países, os especialistas enaltecem este poder discricionário e consideraram-no justo e humano, defendendo que as sentenças do foro criminal deviam basear-se numa grande quantidade de factores que envolvem não apenas o crime, mas também o carácter e as circunstâncias do arguido. A adaptação individualizada estava na ordem do dia. Se os juízes fossem limitados por regras, os criminosos seriam tratados de uma forma desumanizada; não seriam vistos como indivíduos únicos com o direito de chamar a atenção para as especificidades da sua situação. Para muitos,

a própria ideia de devido processo legal parecia um convite a um poder discricionário judicial flexível.

Na década de 1970, este entusiasmo pelo poder discricionário judicial começou a esmorecer por um simples motivo: os surpreendentes indícios de ruído. Em 1973, um juiz conceituado, Marvin Frankel, chamou a atenção para o problema. Antes de ser juiz, Frankel foi um defensor da liberdade de expressão e um fervoroso apoiante dos direitos humanos, tendo ajudado a criar a *Lawyers Committee for Human Rights* (uma organização que é agora conhecida como *Human Rights First*).

Frankel era uma pessoa bastante empenhada. E estava chocado com o ruído no sistema de justiça criminal. Eis como descreve a sua motivação:

Se um arguido acusado de assaltar um banco federal fosse condenado, poderia receber uma pena máxima de 25 anos. Isso significaria qualquer coisa entre 0 e 25 anos. E depressa percebi que, onde essa moldura penal estava definida, o desfecho dependia menos do caso ou do arguido que do juiz, isto é, das suas opiniões, predilecções e tendências. Assim, o mesmo arguido, no mesmo caso, poderia ser condenado a penas muito diferentes dependendo do juiz que era escolhido para o julgamento.

Frankel não apresentou qualquer tipo de análise estatística para fundamentar o seu argumento. Contudo, apresentou uma série de situações muito caricatas que mostravam disparidades injustificadas no tratamento de arguidos semelhantes. Dois homens sem antecedentes criminais foram condenados por tentarem levantar cheques falsos no valor de 58,40 dólares e 35,20 dólares, respectivamente. O primeiro foi condenado a 15 *anos* de prisão e o segundo a 30 *dias*. No caso de dois crimes semelhantes de desvios de fundos,



um dos arguidos foi condenado a 117 *dias* de prisão e o outro foi condenado a 20 *anos*. Referindo inúmeros casos destes, Frankel lamentou aquilo que considerou «os amplos poderes praticamente sem controlo» de juízes federais que resultavam em «crueldades arbitrárias perpetradas todos os dias», considerando-as inaceitáveis num «regime de leis, não de homens».

Frankel apelou ao Congresso para que pusesse fim a esta «discriminação», o termo que usou para descrever aquelas crueldades arbitrárias. Referia-se sobretudo ao ruído sob a forma de inexplicáveis variações nas sentenças, mas também estava preocupado com o enviesamento sob a forma de disparidades raciais e socioeconómicas. Para combater o ruído e o enviesamento, insistiu que não deviam ser permitidas diferenças de tratamento dos arguidos em processos criminais a menos que essas diferenças fossem «justificadas por testes relevantes que pudessem ser formulados e aplicados com objectividade suficiente para garantir que os resultados serão mais que ucasses idiossincráticos de determinados funcionários, juízes ou outros». (O termo «ucasses<sup>2</sup> idiossincráticos» é um pouco esotérico; Frankel usou-o para se referir a decretos pessoais.) Muito mais que isso, Frankel apelou a uma redução do ruído através de um «perfil pormenorizado ou lista de verificação de factores que incluiriam, sempre que possível, alguma forma de classificação objectiva, numérica ou de outro tipo».

Tendo escrito isto no início da década de 1970, Frankel não foi ao ponto de defender o que chamou «substituição de pessoas por máquinas». Porém, surpreendentemente, esteve perto. Ele acreditava que «o estado de direito pede um conjunto de regras impessoais, aplicáveis a todos os níveis, obrigatórias para os juízes e para todas as outras pessoas» e defendeu explicitamente a utilização de «computadores como auxiliares de um pensamento metódico nas sentenças». Também recomendou a criação de uma comissão de sentenças.

---

<sup>2</sup> Decretos dos antigos czares. (N. da T.)

O livro de Frankel tornou-se um dos mais influentes de toda a história do direito criminal — não apenas nos Estados Unidos, mas no mundo inteiro. O seu trabalho enfermava de alguma informalidade. Foi demolidor, mas vago. Para confirmar a realidade do ruído, diversas pessoas começaram imediatamente a explorar o nível de ruído nas sentenças de processos criminais.

Em 1974 foi realizado um dos primeiros estudos de grande envergadura deste tipo, liderado precisamente pelo juiz Frankel. Foi pedido a cinquenta juizes de diversos distritos que deliberassem as penas a atribuir a arguidos em casos hipotéticos resumidos em relatórios pré-sentença idênticos. A descoberta fundamental foi que a «ausência de consenso era a norma» e que as variações nas sentenças eram «impressionantes». Um traficante de heroína podia ser condenado a uma pena de prisão efectiva entre um e dez anos, dependendo do juiz. As penas para um assaltante de bancos variavam entre cinco e 18 anos de prisão. O estudo concluiu que, num caso de extorsão, as penas oscilaram entre uns surpreendentes 20 anos de prisão efectiva e multa de 65 mil dólares a uns meros três anos de prisão, sem multa. E o mais surpreendente de tudo foi que, em 16 de 20 casos, não houve unanimidade em relação à adequação de uma pena de prisão efectiva.

A este estudo seguiram-se outros, e em todos foram encontrados níveis igualmente chocantes de ruído. Em 1977, por exemplo, William Austin e Thomas Williams realizaram um inquérito a 47 juizes, pedindo-lhes que analisassem os mesmos cinco casos, que envolviam pequenas infracções. Todas as descrições dos casos incluíam resumos das informações usadas por juizes em condenações reais, como o texto da acusação, os depoimentos, o cadastro criminal anterior (se houvesse), o contexto social e elementos de prova relacionados com o carácter. A descoberta fundamental foi a existência de uma «considerável disparidade». Num caso de roubo, por exemplo, as sentenças recomendadas variaram entre cinco anos de prisão e apenas 30 dias (com multa de 100 dólares). Num caso

de posse de marijuana, alguns juízes recomendaram penas de prisão efectiva; outros recomendaram pena suspensa.

Um estudo muito maior, levado a cabo em 1981, envolveu 208 juízes federais, que analisaram os mesmos 16 casos hipotéticos. As suas descobertas principais foram surpreendentes:

Só houve acordo unânime para impor uma pena de prisão efectiva em três dos 16 casos. Mesmo nos casos em que a maioria dos juízes concordou que era adequada uma pena de prisão, houve uma variação substancial no tempo de detenção recomendado. Num caso de fraude em que o tempo médio de prisão foi de 8,5 anos, a pena mais longa foi de prisão perpétua. Noutro caso, o tempo médio de prisão foi de 1,1 anos, mas a pena mais longa recomendada foi de 15 anos.

Por muito reveladores que sejam, estes estudos, que envolvem experiências rigorosamente controladas, atenuam quase de certeza a magnitude do ruído no mundo real da justiça criminal. Na vida real, os juízes têm acesso a muito mais informações do que aquelas que são disponibilizadas aos participantes no estudo nos breves esboços com descrições cuidadosamente especificadas. Algumas destas informações adicionais são relevantes, é claro, mas também há muitos indícios de que informações irrelevantes, sob a forma de pequenos factores aparentemente aleatórios, podem produzir grandes diferenças nos resultados. Por exemplo, os juízes consideraram que era mais provável concederem a liberdade condicional no início de um dia de trabalho ou após uma pausa para comer do que imediatamente antes dessa pausa. Se tiverem fome, os juízes são mais duros.

Um estudo que teve como objecto milhares de decisões do tribunal de menores concluiu que, quando a equipa de futebol local perde

um jogo ao fim-de-semana, os juízes tomam decisões mais duras à segunda-feira (e, em menor escala, durante o resto da semana). Os arguidos negros sofrem as consequências desse aumento de dureza de forma desproporcionada. Um estudo diferente analisou 1,5 milhões de decisões judiciais ao longo de três décadas e também concluiu que os juízes são mais severos nos dias que se seguem a uma derrota da equipa de futebol local, e não tão severos nos dias a seguir a uma vitória.

Um estudo que contemplou seis milhões de decisões tomadas por juízes em França ao longo de 12 anos constatou que os arguidos usufruem de uma maior clemência no dia do seu aniversário. (O aniversário do arguido, isto é. Desconfiamos que os juízes também poderão ser mais clementes nos seus aniversários, mas, tanto quanto sabemos, essa hipótese nunca foi testada.) Até um pormenor tão irrelevante como as condições climatéricas pode influenciar os juízes. Analisadas 207 mil decisões judiciais tomadas ao longo de quatro anos referentes a casos de imigração, foi encontrado um efeito importante nas variações diárias de temperatura: quando está calor lá fora, é menos provável que os requerentes consigam asilo. Se lhe moverem perseguição política no país onde nasceu e quiser pedir asilo noutra país, reze para que a sua audiência decorra num dia frio.

## **Reduzir o ruído nas sentenças**

Na década de 1970, os argumentos de Frankel, e as descobertas empíricas que os sustentavam, chamaram a atenção de Edward M. Kennedy, irmão do presidente John F. Kennedy, assassinado em 1963, e um dos membros mais influentes do Senado dos Estados Unidos. Kennedy ficou chocado e horrorizado. Em 1975, introduziu legislação para a melhoria de sentenças, mas não deu em nada. No entanto, o senador foi inflexível. Realçando os elementos

de prova, continuou a pressionar no sentido da promulgação dessa legislação, ano após ano. E, em 1984, finalmente teve sucesso. Reagindo às provas de variabilidade injustificada, o Congresso promulgou a Sentencing Reform Act de 1984.

A nova lei destinava-se a reduzir o ruído no sistema, diminuindo «o poder discricionário sem entraves que a lei confere aos juízes e coordenadores de liberdade condicional responsáveis pela imposição e implementação das sentenças». Em particular, os membros do Congresso referiram a disparidade «injustificadamente grande» nas sentenças, citando especificamente conclusões de que na área de Nova Iorque as penas em casos idênticos podiam variar entre três anos e 20 anos de prisão. Como o juiz Frankel tinha recomendado, a lei criou a Sentencing Commission, cuja principal missão era clara: criar directivas para a aplicação das sentenças, que seriam obrigatórias e estabeleceriam uma moldura penal rigorosa para as penas do foro criminal.

No ano seguinte, essa comissão estabeleceu as tais directivas, que se baseavam genericamente na média de penas para crimes semelhantes de acordo com uma análise de dez mil casos concretos. Stephen Breyer, juiz do Supremo Tribunal, então profundamente envolvido no processo, defendeu a utilização da prática adoptada anteriormente, referindo a inultrapassável discordância no seio da comissão: «Porque não se reuniu a Comissão e se esforçou para racionalizar esta coisa e não apenas analisar a história? A resposta concisa é: não conseguimos. Não conseguimos porque há uma imensidão de argumentos muito bons a apontar em direcções opostas. (...) Tentem classificar todos os crimes que existem por ordem de mérito passível de punição. (...) Depois, observem os resultados dos vossos colegas e vejam se estão de acordo. Digo-vos desde já que não vai acontecer.»

Segundo as directivas, os juízes têm de considerar dois factores para proferir as sentenças: o crime e os antecedentes criminais do arguido. Os crimes são classificados de acordo com 43 «níveis

de infracção», dependendo da sua gravidade. Os antecedentes criminais do arguido englobam sobretudo o número e a gravidade das condenações anteriores. Reunidos o crime e os antecedentes criminais, as directivas oferecem um intervalo relativamente curto de sentenças, com o limite máximo do intervalo autorizado a exceder o limite mínimo até seis meses ou 25%. Os juízes podem afastar-se do intervalo por referência ao que consideram circunstâncias agravantes ou atenuantes, mas os desvios têm de ser justificados num tribunal de apelação.

Muito embora as directivas sejam obrigatórias, não são inteiramente rígidas. Não vão tão longe como o juiz Frankel queria. Conferem bastante espaço de manobra aos juízes. Todavia, diversos estudos que usam uma série de métodos e se concentram em diferentes períodos históricos chegam à mesma conclusão: as directivas reduzem o ruído. Em termos técnicos, «reduzem a variação líquida de penas atribuível à circunstância fortuita da identidade do juiz que profere a sentença».

O estudo mais aprofundado foi realizado pela própria comissão. Sentenças aplicadas a assaltos a bancos, tráfico de cocaína, tráfico de heroína e desfalques bancários em 1985 (antes de as directivas entrarem em vigor) foram comparadas com as sentenças transitadas em julgado entre 19 de Janeiro de 1989 e 30 de Setembro de 1990. Os arguidos foram associados de acordo com os factores considerados relevantes para as sentenças ao abrigo das directivas. Em cada sentença, as variações entre juízes foram muito mais reduzidas no período mais recente, após a implementação da Sentencing Reform Act.

De acordo com outro estudo, em 1986 e 1987 a diferença esperada na duração das penas entre juízes era de 17%, ou 4,9 meses. Esse número caiu para 11%, ou 3,9 meses, entre 1988 e 1993. Um estudo independente que abrangia diferentes períodos encontrou um sucesso semelhante na redução das disparidades entre juízes, que foram definidas como as diferenças nas sentenças médias entre juízes com um número semelhante de casos.