



Valério Gozzano ()*

A Justeza das Medidas Estatísticas

(*) Especialista. Professor de Estatística na
Universidade de Sorocaba — UNISO.



RESUMO

A Estatística é essencial ao desenvolvimento de todas as Ciências Concretas, principalmente das Ciências de Comportamento. Nestas a aplicação do método experimental é exclusivamente estatística. Importa, pois, lembrar que fundamento probabilístico dos métodos estatísticos colimam-se em estruturas seguras e justas.

ABSTRACT

Statistics is essential to the development of all Concrete Sciences, mainly to Behavioral Sciences. Here, the application of the experimental method is exclusively Statistics. So, it is important to remember that the probability foundation of statistic methods culminate in structures that are both safe and fair.

Medições e erros em ciências constituem um capítulo vasto que envolve diretamente alguma filosofia (o que não é tão comum no campo) e gera uma concepção do assunto que sempre parece moderna, mas não é. Talvez a aceitação explícita do grau de incerteza da medição e da medida resultante não se haja tornado atávica. A asserção de absoluta certeza intelectual (filosófica, em si) pode gerar ao adolescente (cronológico ou cultural) um período mais ou menos longo de perplexidade. Abstrações de uma ação tão pragmática quanto a de uma medição direta? Isto poderia acrescentar novos problemas. Mesmo estudantes de Matemática não cogitam da natureza intelectual (abstrata) de sua disciplina. Por vezes eles se interessam pelas “ações” de uma Álgebra Linear ou de uma Geometria Euclidiana, mas têm enorme dificuldade com a natureza da Álgebra em si ou com Geometrias Não-euclidianas. A engenhosidade dos enredos práticos das matrizes e as mensurações superficiais e volumétricas deveriam assegurar-lhes que, ainda que no mesmo campo matemático, preferem o conforto do concreto ante o abstrato conflitante. Paradoxalmente, podem ter dificuldades em compreender que as relações e operações, entretanto, se definem pela funcionalidade.

Porém o estudante de Ciências Concretas, mesmo que tão “exatas” quanto a Física, tem de haver-se com outro problema, o de habituar-se com o procedimento estatístico, além do determinístico, das leis, mormente nos assuntos modernos e contemporâneos. Uma lei sempre sugere um comportamento abstrato, impecável, isento das “poluições” probabilísticas. A concepção de que a Estatística trata o incerto da maneira mais “certa” possível, raramente será espontânea.

Se a imposição das medidas estatísticas às leis concretas vem de longa data, talvez ocorra há menos tempo que permita seu ingresso no genético. Não se mede sequer o comprimento apenas uma vez, como no comércio de tecidos, mas sim vários comprimentos do mesmo evento, para obter a média, a variância, o erro padrão, um intervalo de situação, nunca absolutamente definido; todos estes compromissos requerem algo mais que a absorção intelectual imediata. Um optante por cursos de humanidades freqüentemente não terá sequer interesse em penetrar nas concepções probabilísticas das ciências. Os que se envolvem em carreiras médicas, por exemplo, ou mesmo por engenharias, podem meramente operar com o incerto de modo totalmente empírico sem considerar as implicações abstratas do probabilístico, a não ser raramente. Mas toda vez que

aqui e ali, movido por motivos ao menos práticos, se considera a Estatística, uma questão há de estar presente: a da validade de uma medida indireta. Esse problema deveria ser especialmente crucial para as Ciências do comportamento, cujos comportamentos limites (“leis”) são totalmente estatísticos. Entretanto, via de regra, quem mais receia esse método, que recorre intensamente às matemáticas, é o cientista que se envolve com a natureza humana, mesmo que um biólogo exclusivamente fisiológico. Acrescente-se a natureza peculiar do método probabilístico e completa-se a repulsa.

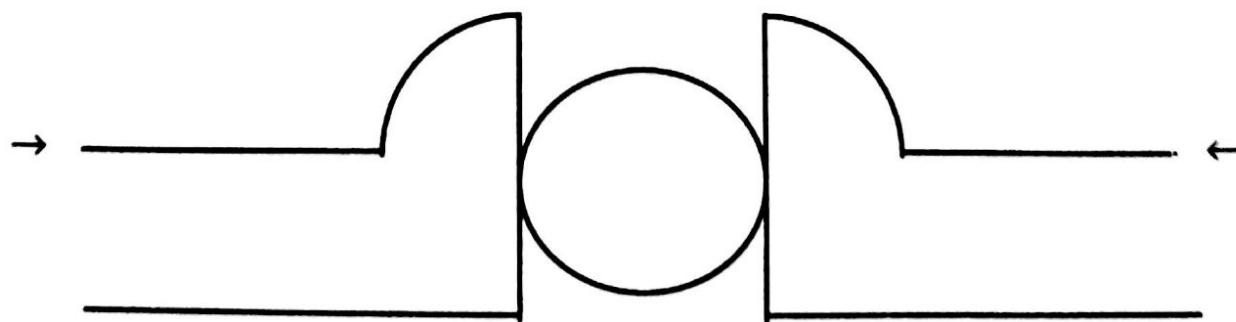
Na verdade, os métodos estatísticos freqüentemente podem ser usados sem que se analisem seus conteúdos matemáticos. Mas o mero uso do método poderá constituir-se em uma barreira irremovível. Poder-se-ia assumir que as técnicas estatísticas fossem solicitadas apenas dos que a conhecessem, mas o solicitante há de conhecer a essência quantitativa do problema e deve poder inferir ou, ao menos, compreender inferências.

Nenhum proveito poderá ser obtido no exercício de uma ciência, se substituirmos as apreciações estatísticas por uma considerável nebulosa, informe, mal conhecida, ainda que bem intencionada.

No entanto, uma concepção necessária para valorizar adequadamente a Estatística há de verificar a justeza de suas medidas. Quem se disporia a desentranhar os mistérios de uma ciência arquitetada sobre medidas não-justas?

A Física contemporânea é eminentemente estatística, mas possui um monumental passado “determinístico” (entendido como produtor de leis invariáveis). Mas o recurso a medições simplíssimas mostra a qualidade essencial à justeza.

Por exemplo, consideremos uma medida de espessura, realizada por um paquímetro aferido, um leitor competente, com precisão satisfatória.



O processo consiste, exclusivamente, em aproximar as aletas ao objeto de modo correto, justo, isto é: mínimo. Semelhantemente, a luva pode medir a mão

à qual se adapte de modo mínimo; menor, não contém a mão; maior, nada mede. Associamos, pois, à justeza o caráter de mínimo. Isto é, toda medida justa procede de um critério que assegure a obtenção do valor mínimo. Esta também é a regra em Estatística, embora as medidas sejam indiretas e embora haja medidas em uso que não a respeitem. São medidas mais ou menos explicáveis: coerência com um procedimento amplo, opção pela facilidade de cálculo, irrelevância da precisão, rapidez, etc.

Freqüentemente as medidas estatísticas são indiretas, construídas a partir de medidas diretas. Assim, uma mediana, que prescinde até de um cálculo, quando determinada por contagem, não tem a menor possibilidade de ser, em si, explorada matematicamente. Os cálculos obtidos a partir dela (ou da moda), enquanto exercendo funções de média, carecem de justificativa matemática e têm alcance muito limitado. A concepção da média equívale a uma definição perfeita — a medida cujos erros por falta e por excesso são compensados. O que equivale a dizer que a soma dos desvios obtidos da média é nula.

Para $d = x - V$, sendo $d = \text{desvio}$, $x = \text{dados}$, $V = \text{valor}$

$$\sum d = \sum x - NV; \text{ se } V = M = \sum x / N \text{ se terá } \sum d = 0$$

$$\text{Ora, sendo } \sum d^2 = \sum x^2 - 2V \sum x - NV^2$$

será a primeira derivada parcial em relação a V :

$$F = -2 \sum x - 2NV$$

$$\text{e para } F = 0 \text{ será } V = \sum x / N = M$$

e a segunda derivada será $2N > 0$, o que garante o caráter de mínimo ao valor da média.

A afirmação de que a soma dos desvios quadrados obtidos da média constitui um valor mínimo é o princípio de toda a análise estatística (técnicas de mínimo quadrados, erros, chi-quadrado, análise de variância ou co-variância, etc). A garantia da justeza, cunho essencial da medida, é obtida através da garantia do valor mínimo, diretamente ou não.

As considerações de medidas não-justas em Estatística parecem-nos mais freqüentes do que, por exemplo, em Física. Isto provém da disciplina, que labora com o inseguro, o incerto, com os fenômenos probabilísticos. As leis que produz hão de considerar a aceitação e a rejeição. A consequência imediata é a dificuldade de trato com a matéria. E também, por vezes, o abuso do que haja de subjetivismo... Não é simples conceber medidas como a mediana e a moda,

embora sejam tão fáceis, talvez por serem definidas por posição ou por repetição. Mas, possivelmente, o motivo preponderante seja o fato de que as pessoas têm uma maneira de pensar marcadamente intuitiva, ao defrontarem-se com o nem tão espontâneo. A verificação, pois, de que a mediana não compensa seus erros pode ser, paradoxalmente, mais eficaz em compreender por que ela não satisfaz como a média. Se, posteriormente, se usa a mediana em lugar da média, por conveniências legítimas, torna-se claro que isso é indevido. Medianas e separatrizes são muito convenientes para escalas, nas quais nada mais pretendem que fixar posições. A escolha da média, ao contrário, possibilita qualquer exploração matemática — se pretendida o for. Assim, embora o desvio médio seja calculado através da média, a possibilidade de sua aplicação é muito restrita por sua abstração dos sinais dos desvios, justificável pelo fato estatístico, mas não pelos cânones matemáticos. Já o desvio padrão (ou a variância) resulta uma medida perfeita — também na justeza.

O fundamento experimental das Ciências Concretas centra-se nas técnicas estatísticas, assim como suas posteriores inferências. As ciências da natureza, mesmo estabelecendo leis “exatas”, não podem abstrair-se da insegurança das medições. Pode armar-se dos melhores instrumentos, os mais aptos observadores, as mais refinadas técnicas, que o problema permanece: não há medidas exatas, as há mais ou menos precisas. É necessário, pois, proceder a uma estimativa do erro, e a teoria dos erros insere-se na Estatística. As Ciências do Comportamento, porém, só podem determinar um comportamento limite, como se procurassem uma lei “inexata”. Apenas algum comportamento diretamente vinculado ao homem, como animal, poderá ser pensado em termos determinísticos: causas múltiplas e vinculadas só podem produzir conseqüências inseguras. Diríamos que a tendência de comportamento é intrinsecamente estatística, probabilística. Será impossível menosprezar no comportamento humano o exercício do arbítrio (mais ou menos) livre, da vontade humana. Se é possível obter, com alguma precisão, a medida de uma grandeza física com apenas uma medição ou poucas, isto não é concebível nas observações do Homem. Não somente um grande número de observações é requerido, como as medidas podem assumir valores diferentes em locais ou épocas diversos.

Considere-se uma medida de correlação em testes psicológicos. O valor da correlação só adquire um grau satisfatório de confiança, se o número de aplicações de teste é vasto, vastíssimo. A precisão aqui tem um sentido um tanto diverso daquela obtida na medição de uma grandeza física, que poderá ser obtida com aproximação inconcebível ao leigo. Os intervalos de confiabilidade dos dois exemplos são extremamente diversos, quantitativamente. O método

estatístico é o mesmo, mas considerado de maneira diferente: diríamos, pelo óbvio, que os seres animados não têm a perenidade física dos inanimados, estes praticamente imutáveis. As naturezas das ciências fundamentadas nos métodos estatísticos são, por vezes, tão diversas, que tendem a colocar os menos capacitados em matemáticas em uma atitude defensiva perigosa, que consiste em prosseguir considerando sua particular área de interesse como isenta de necessidades estatísticas, ou determináveis por inferências estatísticas dispensáveis, se se limita o interesse pelo aspecto teórico.

Um exemplo muito célebre lembra-nos, entretanto, como uma ciência nem tão “exata”, como a Biologia, pode contribuir ao “exato” dos processos estatísticos: as leis de Mendel. Até o presente identifica-se a “regressão” das equações, que só seria apropriada para fenômenos hereditários.

Como assinala a Dr.^a Eva Nick, da Universidade Federal do Rio de Janeiro, já não é possível prescindir da aprendizagem de técnicas estatísticas na formação dos profissionais das ciências do comportamento.

De todas as técnicas, dentre as mais propícias, estão os testes conhecidos como não paramétricos, que prescindem de suposições sobre a distribuição populacional. O desenvolvimento das técnicas é, em geral, tão simples, que pode ser aplicado a dados apenas ordenados, sem que haja exatidão de seus valores numéricos. Além disso, os testes não -paramétricos são aplicáveis a amostras pequenas, caso freqüente em pesquisas piloto. Aliás, a própria natureza da pesquisa pode referir-se a dados de difícil obtenção, o que reduz o tamanho da amostra.

Provavelmente, todas as leis físicas macroscópicas são estatísticas. Mesmo as leis microscópicas são apenas quantitativamente aproximadas, já que envolvem grandezas que conhecemos dentro dos limites dos erros. Como já Newton afirmava, as leis da natureza são representações objetivas das relações que existem na natureza, mas são limitadas por nosso conhecimento imperfeito dos eventos dos quais participam. E das leis comportamentais que se diria? A obtenção de uma tendência limite de comportamento com uma permanência tal, que permita considerá-la “regra”, é pouco menos que o almejado.

A Física dos comportamentos moleculares não é menos estatística. Ao pesquisador iniciante, o comportamento dos componentes atômicos pode sugerir uma vontade própria, como se eles fossem animados por um princípio desconhecido. Como abstrair-se, pois, das mesmas preocupações em Psicologia, por exemplo? Mesmo o reconhecimento estatisticamente objetivo de que a tendên-

cia comportamental pesquisada tem tal grau de incerteza, tal imprecisão, que talvez não exista, também é ciência.

REFERÊNCIAS BIBLIOGRÁFICAS

1. LEVIN, Jack. **Estatística aplicada a ciências humanas**. São Paulo : HARBRA, 1978.
2. SIEGEL, Sidney. **Estatística não-paramétrica para ciencias do comportamento**. São Paulo : McGraw-Hill, 1981.
3. WALPOLE, Ronald E., MYERS, Raymnonnd H. **Probability and statistics for engineers and scientists**. New York : McMillan Publishing, 1972.