

## — FUNDAMENTOS DA ESTATÍSTICA —

Prof. Raymundo Costa e Souza

### I — Universalidade do «número»

E' inegavel a importância da Estatística em qualquer setor da atividade humana que defronte problemas, para cuja solução seja indispensável um adequado tratamento numérico. Tanto no âmbito científico como no técnico, o «número» é elemento primordial que se impõe diante da necessária precisão dos resultados que se faz oportuno estabelecer e também da subsequente mensuração do grau dessa mesma precisão.

Em apoio do que afirmamos, invoquemos o pensamento de Lord Kelvin enunciado no fim do século dezenove, quando então de maneira inequívoca precisou o relevante papel que desempenham as apreciações de natureza quantitativa no desenvolvimento das ciências. Sua asserção proferida numa época de animadoras inovações no campo material e no social, transformou-se através do tempo em advertência metodológica aos investigadores em suas diferentes ciências :

«Quando podemos medir e traduzir em números aquilo de que estamos falando, já conhecemos alguma coisa a seu respeito; mas quando não podemos medi-lo, quando não podemos exprimi-lo em números, nosso conhecimento é de natureza frágil e pouco satisfatória» (1).

---

(1) — YULE, G. UDNY e KENDALL, M. G. Introdução à teoria da estatística. Rio de Janeiro, Instituto Brasileiro de Geografia e Estatística. 1948, p. 19 e 20.

Sem pretender subscrever «in totum» a idéia de Lord Kelvin nesse particular, por reconhecermos que a interpretação de muitos aspectos das ciências psicológicas, históricas e sociais, independe e não fica subordinada a conceitos quantitativos, corroboramos com sua opinião no que concerne às ciências que estudam as categorias de fenômenos cujas componentes são passíveis de expressão quantitativa, sob a forma de grandezas quer contínuas ou descontínuas.

E' fundamental o pensamento de Lord Kelvin, merecedor de paciente e minucioso exame, a fim de que possa ser assimilado o alcance transcendente do que escrevera aquêlê genial e laborioso cientista. Ele percebia que os números e os processos de mensuração vinham penetrando em todos os ramos do conhecimento; as ciências em sua quase totalidade, quer direta ou indiretamente os utilizavam, ora em fôrma primária ora modificados; não sòmente as ciências de caráter experimental mas também aquelas de características sociais recorriam a expressões quantitativas; os atributos qualitativos se tornavam mais significativos todas as vezes que apareciam associados a uma idéia numérica.

De fato, não é difícil constatar hoje como os números são úteis e sugestivos ao espírito em busca de conhecer mais e melhor acêrca das realidades que cercam o homem. Vê-se que além de formarem a base das experiências realizadas em laboratórios de física e química, o emprêgo dos números sob expressões de «grandezas» se estende às estações de experimentação biológica, aos institutos de pesquisas sociológicas e econômicas, aos círculos de estudos teóricos e aplicados, às administrações públicas e privadas, sempre criando formas inéditas de vida e harmonia, que sem êles inexistiriam, em detrimento dos pesquisadores desejosos de assentarem resultados tão positivos quanto possível.

## II — Problemas matemáticos e estatísticos

Admitida a suscetibilidade de serem os fenômenos de muitas das várias ciências, decompostos em expressões quantitativas, pretendendo a uma melhor elucidação dos seus fundamentos, estabelecemos dois princípios que orientam os pesquisadores na utilização dos conceitos numéricos em face das suas investigações.

Em razão da unidade de propósitos por parte dos pesquisadores que manipulam dados quantitativos, verifica-se pelo menos dois objetivos genéricos, independente da natureza da própria pesquisa, que precisam ser satisfeitos com certo rigor, no curso de empreendimento científico. Estes dois objetivos são em síntese :

- 1.º) — «atribuir precisão aos resultados da pesquisa»;
- 2.º) — «avaliar o grau de precisão dos resultados da pesquisa».

O primeiro destes objetivos, dominando a pesquisa desde o início, reveste o caráter de problema matemático, onde se aplicam relações abstratas a fatos concretos. Problema básico desse tipo é o emprêgo dos conjuntos numéricos à descrição e interpretação dos fenômenos científicos, figurando como caso mais imediato a enumeração dos «conjuntos discretos», formados por fatos ou objetos isolados, por exemplo, o número de pessoas que integram uma nação, o número de filhos pertencentes a certa família, o número de homens desempregados existentes numa cidade, em que a contagem dos fatos unitários requer o uso dos números inteiros (2). Temos assim como unidade o número «um», correspondente a um fato ou objeto destacado; dizemos, por conseguinte, que naquela coletividade havia  $X$  habitantes, ou que uma família conhecida tinha  $Y$  filhos, ou que montava a  $Z$  os desempregados da localidade à qual nos referimos; as unidades são respectivamente o indivíduo-habitante, o indivíduo-filho e o indivíduo-desempregado.

Estes «conjuntos discretos» podem receber a denominação de «variáveis descontínuas».

Entretanto, a simples enumeração dos «conjuntos discretos» não é suficiente para uma apreciação completa dos fenômenos reais, porquanto a observação científica revela a permanência de propriedades induzidas das nossas sensações, achando-se dentre elas: as propriedades de natureza física, sob modalidades de temperatura, comprimento, massa, tempo, tonalidade musical etc.; as propriedades de natureza econômica, percebidas através das formas de valor, grau de satisfação, capacidade aquisitiva etc.; e outras. Da identificação dessas propriedades naturais às sucessões numéricas, emerge o conceito de «grandeza», assumindo grande importância em relação à evolução dos métodos e processos científicos. Encontramos então na amplitude da investigação científica, grandezas físicas (temperatura, comprimento, massa etc.), grandezas econômicas (preço, salário, volume de vendas etc.), e assim por diante.

Na maioria das vezes, as grandezas são chamadas «variáveis contínuas».

---

(2) — ALLEN, R. G. D. *Análisis matemático para economistas*. Madrid, M. Aguilar, 1946, p. 10 e 11.

O segundo dos objetivos citados, aparecendo sòmente depois que o primeiro tenha sido cumprido, envolve um «problema de estatística» pròpriamente dito. E' uma nova etapa da investigação, sucedendo à determinação das variáveis contínuas e descontínuas, que fôra o objeto fundamental da fase anterior, quando procuravamos apenas atribuir precisão aos resultados da pesquisa.

Se é o desejo de consignar precisão aos resultados perquiridos, o que conduz os pesquisadores dos mais variados campos da ciência, a exprimirem «quantitativamente» os fatos concernentes aos setores que lhes interessam, são êles posteriormente estimulados a medir até que ponto êsses mesmos resultados são precisos, situando então «estatisticamente» os fatos já observados.

Realizando o seu trabalho de avaliar o grau de precisão dos resultados, o pesquisador apela inevitavelmente para a teoria da inferência estatística, a qual utiliza uma técnica especial baseada na semelhança que necessariamente existe entre os elementos que estão sendo examinados, integrantes do fenômeno, e a amplitude total de elementos da mesma espécie que constituem o fenômeno global. Ergue-se assim uma nova indução em alicerces quantitativos, estribada na: cálculo de algumas características sôbre uma parte isolada dos dados, e na generalização dos resultados dessas características ao conjunto de dados da mesma natureza. Na técnica da «inferência ou indução estatística», a parcela destacada na qual executamos nossas operações, é denominada «amostra», reservando-se o nome «universo» ao conjunto, acêrca do qual pretendemos estender os resultados.

Uma operação dêsse tipo consiste na estimativa da magnitude de uma variável, contudo que os valores estimados para o universo a partir da amostra, caiam dentro de intervalos mais ou menos definidos, variando a precisão dos resultados em função da amplitude dêsses intervalos.

E' exatamente estabelecendo os limites de tais intervalos e calculando que probabilidade têm os valores estimados de ficarem circunscritos a êsses limites, que o investigador chega a dispor de um expediente útil e precioso, destinado a mensurar o grau de precisão dos resultados quantitativos provenientes das observações científicas.

Esta íntima conexão entre problemas matemáticos e problemas estatísticos se prende à própria essência da Estatística, que é fortemente vinculada às noções quantitativas, permitindo-nos dizer sem exagêro que a metodologia estatística é na realidade

um derivativo da aplicabilidade dos conceitos matemáticos aos fenômenos reais.

### III — Objeto da Estatística

A imponente questão de «avaliar ou mensurar o grau de precisão dos resultados da pesquisa», é problema que comporta a utilização de muitos processos e técnicas da estatística.

Antes de tudo, precisamos conceituar «estatística» a fim de nos ser mais fácil associar os seus fundamentos a alguns dos aspectos científicos que caem sob a percepção dos pesquisadores.

A Estatística consiste na «observação metódica por meio de processos e técnicas especiais, no sentido de analisar e interpretar, tão precisamente quanto possível, dados quantitativos numerosos».

Decorre deste conceito que cabe à Estatística proporcionar ao investigador um conjunto de procedimentos através dos quais seja possível examinar a estrutura de um fenômeno em forma quantitativa, após inúmeras observações acumuladas das suas modalidades. Não são, portanto, as observações unitárias, individuais, que constituem o objeto da Estatística, mas sim um numeroso grupo de observações componentes do denominado «fenômeno coletivo» ou «fenômeno em massa», diferenciados nitidamente dos meros «fenômenos singulares».

Em consequência do seu caráter sobretudo metodológico, os critérios estatísticos são aplicáveis extensivamente a todas as ciências cujos fenômenos sejam suscetíveis de apreciação quantitativa, daí o observador poder dispor de um meio valioso para aferir a precisão dos resultados numéricos.

Analisaremos dois importantes grupos de fenômenos, tomados de conscienciosa classificação do egrégio mestre da estatística na Itália, Corrado Gini, transcrita em sua obra «Curso de Estatística». A êssez dois grupos particulares adaptaremos nossa interpretação concernente ao «grau de precisão que se pode mensurar» com o concurso da estatística, toda vez que as conclusões científicas revestirem «modalidades quantitativas».

O referido tratadista estabeleceu como primeiro e segundo grupos dentro da sua classificação, segundo um critério eminentemente científico, os seguintes: (3)

---

(3) — GINI, CORRADO. *Curso de Estadística*. Madrid, Editorial Labor, S/A, 1953, p. 14 - 16.

nômeno econômico da «produção», por carecer de significação qualquer idéia dêste fator, sem uma apreciação de ordem estatística. No caso da indústria, por exemplo, não é praticável perceber-se nem analisar algumas características da produção fabril, quer em seu caráter estático ou dinâmico, senão recorrendo à aplicação da metodologia estatística.

Analisemos uma questão rudimentar de «contrôle estatístico da qualidade», processo hoje amplamente empregado pelas organizações fabris dos grandes centros industrializados, no propósito de manter os produtos manufaturados sob rigoroso padrão ou conservar a eficiência das operações em nível satisfatório. Daremos uma figuração baseada em dados hipotéticos ilustrados por Clark (5), demonstrando o cuidado que um fabricante precisa dispensar a um produto que tenha de atender a certas especificações.

O observador poderia arrolar o número de artigos produzidos no decurso da fabricação, cujas unidades reunidas ao fim das observações acumuladas, serviriam para definir a precisão que necessariamente corre em todo fenômeno de modalidades quantitativas. Ele teria dessa maneira, inventariado os itens fabricados, para uma ulterior finalidade contábil ou administrativa.

Todavia, esta operação de arrolamento não faculta de nenhum modo ao observador, avaliar o grau de precisão dos resultados numéricos referentes ao volume dos itens produzidos. Para tal, o técnico em «controle de qualidade» (quality control), usando processos estatísticos, apontará margens de erro que exprimirão a maior ou menor proporção do aproveitamento de materiais, do rendimento fabril, etc.

Suponhamos que terminada a fabricação de certo artigo, alguns itens tenham de ser excluídos por não corresponderem aos padrões exigidos para o produto, e que no curso de 30 dias consecutivos, foram registrados os números de itens eliminados, como segue: 10, 13, 12, 8, 11, 11, 13, 9, 10, 10, 12, 12, 15, 12, 6, 10, 13, 11, 11, 12, 14, 10, 7, 11, 15, 13, 12, 9, 11 e 12.

Calculando a média e o desvio-padrão dêstes números, acharemos :

$$\begin{array}{ll} m = 11,2 & 2,1 \\ m = 11,2 & d. \text{ padrão} = 2,1 \end{array}$$

Como a série acima forma uma distribuição de frequência aproximadamente normal, e sabendo-se que sob esta condição,

(5) — CLARK, CHARLES E. *An introduction to statistics*. New York. John Wiley & Sons, Inc., 1953. p. 90 e 91.

quase todos os números da série diferem da média em três vezes o desvio-padrão, concluiremos que os trinta números em sua maioria, estão situados entre:

$$\begin{array}{ccc} 11,2 - 3(2,1) & e & 11,2 + 3(2,1) \text{ ou} \\ 4,9 & e & 17,5. \end{array}$$

Vejamos agora a significação e utilidade destes limites. E' que se considerarmos o número de artigos que serão eliminados nos dias futuros, embora desconhecidos, supõe-se que enquanto continuarmos a verificação da quantidade dos itens rejeitados dia a dia, os valores da média e do desvio-padrão não variarão acen-tuadamente. Assim, os números de itens eliminados em cada dia no futuro, estarão dentro do intervalo 4,9 e 17,5.

Podemos então estabelecer o critério de que em dia algum o número de itens rejeitados deve exceder a 17; registrando-se os excessos além dêste número, é bem provável que o processo de fabricação esteja deficiente, valendo a pena procurar-se corrigir as falhas; por outro lado, se durante algum dia o número de itens rejeitados não atingir a 5, poder-mos localizar uma circunstância ocasional favorável, a qual seria conveniente perpetuar em prol do aperfeiçoamento do processo fabril.

## VI — Estatística e Economia

Prosseguindo a nossa explanação no que tange às modalidades numéricas dos fenômenos, tratemos do segundo grupo da classificação de Corrado Gini, que integra fatos suscetíveis de tratamento estatístico, quando se estuda o seu aspecto quantitativo ao lado da sua feição exclusivamente qualitativa.

E' sabido que na investigação econômica, inúmeras vezes é insuficiente a mera observação das modalidades qualitativas do fenômeno, recorrendo porisso o economista a um procedimento fundado em conceitos estatísticos, de modo a perceber o desenvolvimento do fenômeno em perspectiva ampla.

Ilustremos a asserção com a pesquisa realizada conjuntamente por Victor S. von Szeliski e L. J. Paradiso sôbre as condições de rentabilidade da indústria de sapatos nos Estados Unidos da America do Norte, tomando dados referentes aos anos de 1919 a 1921-1933. (6). A análise levada a efeito com a finali-

---

(6) — SZELISKI, VICTOR S. VON e PARADISO, L. J. Demand for boots and shoes as affected by price levels and national income. *Econometrica*. Vol. 4, n.º 4, Outubro 1936. p. 338 e 339.

dade de estabelecer as «curvas de procura» descritivas da tendência aproximada da evolução dos fatores que geravam a renda dessa atividade fabril, assentou nos princípios e técnicas da correlação estatística aplicáveis ao campo econômico; fôra orientada no sentido de atribuir precisão às conclusões da pesquisa e igualmente para avaliar o grau dessa precisão. Ora, a principal conclusão dos investigadores acima citados, foi que a renda nacional tem influência extraordinária na formação da renda auferida pela indústria de sapatos, enquanto são menos intensos os efeitos produzidos pela variação dos preços. Eles averiguaram que se a renda nacional crescesse em 10%, com os outros fatores inalterados, a rentabilidade da indústria pesquisada aumentaria em 7,4%; porém se os preços subissem em 10%, permanecendo constantes os demais fatores, a renda da mesma indústria cresceria em 4,8% apenas.

Von Szeliski e Paradiso determinaram também as tendências comparativas da produção per capita, valor da produção per capita, valor por par, preços e renda nacional. Verificaram ao termo da pesquisa, que o consumo per capita evoluiu de certo modo estável no curso do período, variando os extremos dentro de uma amplitude cêrca de 13% acima e abaixo da média equivalente a 2,48 pares per capita ao ano.

Êstes foram alguns dos resultados reportados por aquêles econometristas, situando o aspecto quantitativo do fato, através de dados numéricos, cuja precisão procuraram estabelecer ao máximo possível. Essa apreciação original de caráter quantitativo, é mais sugestiva e útil do que as observações de fundo tão só qualitativo; conquanto as últimas pudessem deixar entrever as prováveis relações entre a renda nacional e a rentabilidade da indústria de sapatos, nunca chegaria, todavia, a estabelecer resultados claros, definidos e precisos.

Morris A. Copeland (7) põe em destaque a importância da estatística aplicada à economia política, ao asseverar que as mensurações quando comparadas, constituem uma técnica essencial à elucidação dos fenômenos econômicos, salientando as comparações no tempo, ha muito reconhecidas e estudadas sob a forma de séries cronológicas ou temporais, tais como as séries concernentes às safras agrícolas de cada ano, ou as subclasses dessas séries anuais, como as contruidas em base trimestral ou mais frequentemente, que são ainda mais úteis do que as primeiras. A

---

(7) — COPELAND, MORRIS A. *Statistics and objective economics*. Journal of the American Statistical Association. Vol. 50, n.º 271, Setembro 1955 p. 641.

grande variedade de séries estudadas pela ciência econômica, de importação e exportação, taxas de câmbio, preços de mercadorias, cotações de títulos, produção de ferro e aço, emprêgo e desemprego, estoques, construções, renda nacional etc., dão uma idéia da inofismável imprescindibilidade da aplicação dos métodos estatísticas à investigação dos fatos econômicos.

## VII — Conclusão

Alcançada esta fase da nossa exposição, tornamo-nos receptivos quanto às vantagens que advirão aos pesquisadores todas as vezes que utilizarem a metodologia estatística; a maior clareza proveniente do uso de processos numéricos para o fim de, (a) atribuir precisão aos resultados da pesquisa, e (b) medir o grau dessa precisão, conduz inequivocamente a uma descrição mais extensiva e a uma interpretação mais significativa da realidade fenomênica. Os princípios da matemática pura transferidos ao domínio da estatística, integram os processos e técnicas desse método, cuja eficiência tem sido comprovada nas pesquisas que envolvem entes quantitativos em massa, tomados de um fenômeno submetido a observação sistemática e científica.

## — BIBLIOGRAFIA —

YULE, G. UDNY e KENDALL, M. G. *Introdução à teoria da estatística*. Tradução da 13a. edição inglesa por Evandro de Oliveira Silva. Rio de Janeiro, Instituto Brasileiro de Geografia e Estatística, 1948.

ALLEN, R. G. D. *Análisis matemático para economistas*. Traduzido do inglês por Emilio de Figueiroa. Madrid, M. Aguilar, 1946.

GINI, CORRADO. *Curso de Estadística*. Tradução da edição italiana de 1946-47 por Jorge Stecher Navarra. Madrid, Editorial Labor S/A, 1953.

DEMING, WILLIAM EDWARDS. *Some theory of sampling*. New York, John Wiley & Sons, Inc., 1950.

CLARK, CHARLES E. *Some introduction to statistics*. New York, John Wiley & Sons, Inc., 1953.

SZELISKI, VICTOR S. VON e PARADISO, L. J. *Demand for boots and shoes as affected by price levels and national income*. *Econometrica*, Vol. 4, n.º 4, Outubro 1936.

COPELAND, MORRIS A. *Statistics and objective economics*. *Journal of the American Statistical Association*, Vol. 50, n.º 271, Setembro 1955.