

AFFONSO CELSO PARREIRAS HORTA

Apontamentos
de
Metodologia Estatística

(SEPARATA DA "REVISTA BRASILEIRA DE ESTATÍSTICA" — ANO X — Nº 37)



RIO DE JANEIRO

SERVIÇO GRÁFICO DO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA

1949

AFFONSO CELSO PARREIRAS HORTA

Apontamentos
de
Metodologia Estatística

(SEPARATA DA "REVISTA BRASILEIRA DE ESTATÍSTICA" — ANO IX — N.º 37)



RIO DE JANEIRO
SERVIÇO GRÁFICO DO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA
1 9 4 9

APONTAMENTOS DE METODOLOGIA ESTATÍSTICA

GENERALIDADES

O papel dos números

Nada há que melhor represente o resultado da aplicação dos processos aconselhados pela ciência, que vamos estudar, do que um Anuário Estatístico. Que contém uma dessas publicações? Uma série de quadros, cujos títulos ou cabeçalhos, indicam fatos, em seguida descritos por meio de algarismos ou cifras, isto é, numéricamente. São, pois, os números a parte mais saliente da Estatística, e é justo que lhes consagremos as primeiras referências ao iniciar o estudo desta ciência.

Um ilustre professor da Universidade de Paris, FERNAND FAURE, assim caracterizou o papel dos números, nas ciências sociais: "Os matemáticos, êsses mestres dos números, costumam repetir que os números governam o mundo, "mundum regunt numeri". Isto, porém, não é exato, pois os números são criações do nosso espírito. Nada governam. Os estatísticos, menos ambiciosos, contentam-se em dizer, com GOETHE, "que os números servem para nos explicar como o mundo é governado". Esforçam-se, o mais conscienciosamente que é possível, em reuni-los e pô-los à disposição dos homens de ciência¹.

Ensinam as enciclopédias que os homens, muito antes de se aproveitarem dos dedos, para escrever, já dêles se utilizavam para contar as coisas. A superstição desde logo se apoderou dos números, e a própria filosofia grega, analisando-lhes, com a habitual sutileza, os característicos e as qualidades, deu-lhes, com PITÁGORAS e PLATÃO, um

papel demasiado importante na teoria das idéias. Só, modernamente, caiu em descrédito a opinião de que há números fatídicos.

A Estatística não foi feita para criar êsses números, e, embora lidando, sem descontinuar, com algarismos, nenhuma força ou significação cabalística lhes atribui, limitando-se a considerá-los como insubstituível instrumento de investigação, manejado pelas ciências matemáticas, e que se presta, maravilhosamente, para o estudo dos fenômenos da vida social.

SE houve, no passado brasileiro, certa preocupação pela existência de informações estatísticas — especialmente as ligadas ao estado da população — muito pouco há que assinalar no tocante à bibliografia específica, não tanto, propriamente, a que concerne à publicação de dados — de que se encarregariam as "falas", relatórios e mensagens oficiais —, mas a de metodologia. Toma, por isso, assinalada posição na história da bibliografia estatística nacional a obra de SEBASTIÃO FERREIRA SOARES, Elementos de Estatística, lançada em 1865, ou seja sete anos antes do primeiro recenseamento geral do País.

Depois do trabalho de FERREIRA SOARES, decorreriam 61 anos até o aparecimento de Apontamentos de Metodologia Estatística, de AFFONSO CELSO PARREIRAS HORTA, que se inscreveria, como a primeira, em lugar destacado em nossa bibliografia especializada. Inspirado nos ensinamentos de COLAJANNI, GABAGLIO, ARMAND JULIN, VIRGILII e BLOCK, elaborou PARREIRAS HORTA excelente manual, de que se utilizariam, com o maior proveito, quantos, no Brasil, — e notadamente na fase de recrutamento do pessoal que viria a integrar os quadros do sistema estatístico do País instituído em 1936 — necessitaram conhecer noções fundamentais de Metodologia Estatística.

Reconhecendo os relevantes serviços que os Apontamentos prestaram e podem, ainda, vir a prestar — embora sujeitos a modificações alguns de seus conceitos, — deliberou a Assembléia-Geral do Conselho Nacional de Estatística encarecer a conveniência de a REVISTA incluí-los na secção específica de um dos seus números. E é isto o que ora se faz, depois de obtida a necessária licença para a publicação.

Definição da Estatística

Feliz, portanto, sob vários pontos de vista, se nos afigura a definição da Estatística, formulada por ÉMILE LEVASSEUR, nos seguintes termos: estudo numérico dos fatos sociais.²

Para comentar essa definição, parece-nos bastante reproduzir o seguinte trecho de F. VIRGILII³:

"O homem deve ser também estudado sob outros aspectos: vive em sociedade, desenvolve a sua atividade de mil formas diversas, vale-se das forças da natureza para aumentar e distribuir de diversos modos as riquezas, realiza atos generosos e perpetra delitos, experimenta o fanatismo religioso e escraviza-se aos hábitos e às conveniências da vida prática, obedece a leis positivas e submetete-

se a organizações políticas especiais, recolhe-se a um claustro ou funda associações, atrofia os sentimentos de seu coração ou regozija-se com as alegrias do lar doméstico. Se o estudo de todos êsses múltiplos aspectos da vida social pertence a ciências diversas, a determinação quantitativa de todos êsses fatos constitui o objeto ou a matéria de uma única ciência que é, precisamente, a Estatística".

Essa determinação quantitativa dos fatos sociais deverá realizar-se segundo certas normas, cujo conjunto constitui a parte técnica ou metodológica da Estatística.

¹ Segundo JACQUES BERTILLON — *Cours Élémentaire de Statistique*, 1905, pág. 1.

² FILIPPO VIRGILII — *Manual de Estatística*, 4.ª edição, traduzida, 1908, pág. 3.

¹ FERNAND FAURE — *Éléments de Statistique*, 1906, págs. 57 e 58.

Conforme outras disciplinas sociais, a Estatística tem sido definida por meio de numerosas fórmulas, que, ao tempo da publicação do tratado de M. BLOCK⁴, já subiam a 200. BERTILLON refere-se apenas a mais de 100 definições, porém COLAJANNI⁵ eleva o total a perto de 300, fazendo, a respeito, as seguintes judiciosas reflexões:

"É por isso que, sob a definição de Estatística, existe toda uma literatura, — classificada por MOHL de extravagante e de singularidade psicológica, — já se havendo enumerado cerca de 300, na maior parte sem precisão, ou falhas, unilaterais, ou prolixas. A de WAGNER, por exemplo, é uma longa descrição. Outras, com aparência paradoxal, eram destinadas a impressionar e a serem para sempre lembradas, mesmo quando se reconhece não possuírem o menor requisito de uma boa definição. Haja vista a célebre e antiga de SCHLÖZER: "A Estatística não é senão a História em repouso; a História não é senão a Estatística em movimento."

Tão famosa quanto a de SCHLÖZER é a de NAPOLEÃO: "A Estatística é o orçamento das cousas, e sem orçamento não há salvação."

Sem nos emaranharmos no dédalo de controvérsias a que a análise aprofundada da natureza do número estatístico e da própria Estatística, como ciência e como método, arrasta a generalidade dos autores, preferimos ater-nos à opinião, segundo A. JULIN⁶ ainda hoje freqüentemente admitida, de que existe "uma ciência estatística, aplicando-se aos fenômenos sociais, e um método estatístico que tem por objeto o estudo dos fenômenos variáveis, e utilizável tanto pelas ciências naturais como pelas sociais".

O estudo numérico dos fatos sociais deu lugar à criação de um método de observação, com caracteres bem acentuados, e que, na prática, se verificou ser proveitosíssimo, para o estudo dos demais fenômenos, que interessam ao homem.

Daí, o crescente empenho com que se passou a examinar e descrever as propriedades características desse método, dedicando-lhe os autores cada vez maior atenção e espaço em seus tratados, cuja primeira parte é, exatamente, o estudo da Estatística como método de observação.

Divisão da Estatística

COLAJANNI e VIRGILII, adotando a divisão da Estatística, professada por FERRARIS, a discriminam em metodológica e expositiva, subdividindo a metodológica em teórica e técnica, e a expositiva, em descritiva e investigadora.

A Estatística metodológica teórica consiste na teoria geral da Estatística, isto é, nos princípios gerais do método, a que acima aludimos.

A Estatística metodológica técnica indica quais os processos especiais, que deve-

rão ser adotados para as pesquisas, tendo em vista a natureza peculiar do objeto a que se reporta a investigação.

A Estatística expositiva descritiva limita-se a dar notícia das manifestações quantitativas dos fenômenos observados, e a Estatística expositiva investigadora procura pôr em evidência as causas dos fenômenos, e as leis empíricas, que os regem e resultam das manifestações quantitativas.

FERRARIS observa que a Estatística metodológica técnica, juntamente com as duas subdivisões da expositiva, podem ser englobadas sob a denominação de Estatística aplicada, constituindo a Estatística teórica a primeira subdivisão da Estatística metodológica.

Ciências afins

No resumo do seu curso de Estatística, A. JULIN diz que, no conjunto das ciências, a Estatística filia-se às ciências sociais, e, pelo seu método, está ligada à Lógica.

O mesmo autor, no seu grande curso⁷ *Princípios de Estatística teórica e aplicada*, transcreve, a respeito, um significativo trecho de RÜMELIN⁸, no qual esse notável escritor determina, com muita clareza, a posição da Estatística entre as demais ciências:

"A Estatística, diz êle, é um ramo das ciências sociais, e deve ser considerado como tal, qualquer que seja a definição que se lhe dê. Todavia, esse método particular de investigações, cujo traço essencial é de contar e classificar a multidão dos casos, foi pôsto, a princípio, ao serviço da Estatística, e, em seguida, ao dos interesses do Estado; mas, em sua essência e em suas aplicações, é de alcance muito mais geral, e não está restrito a esse círculo de observações. Como todas as formas do método científico, o método estatístico também faz parte da Lógica."

Das ciências sociais, cumpre-nos destacar a História e a Economia Política, como profundamente ligadas à Estatística.

Quanto à História, basta-nos recordar, como o fez COLAJANNI, a célebre definição de SCHLÖZER, citada páginas acima, para assinalarmos a que ponto se pode ser levado pelas afinidades existentes entre ambas as ciências. Os próprios autores mostram-se embaraçados em bem explicar as diferenças que separam a História e a Estatística, chegando VIRGILII à curiosa afirmativa de que, a seu entender, "a diferença essencial entre a História e a Estatística consiste em que a primeira narra e a segunda descreve."

Ora, segundo os dicionários, narrar e descrever são quase sinônimos (narrar: expor minuciosamente, relatar, contar; descrever: narrar circunstancialmente, — são as definições de CÂNDIDO DE FIGUEIREDO).

COLAJANNI foi um pouco mais feliz, quando frisou as diferenças entre as duas ciências, ponderando que a Estatística "não dá a identidade das pessoas, mas apenas as encara como objetos de numeração; não reve-

⁴ MAURICE BLOCK — *Traité Théorique et Pratique de Statistique*, 1886, pág. 90.

⁵ N. COLAJANNI — *Lezioni di Statistica*, 1903, pág. 11.

⁶ ARMAND JULIN — *Principes de Statistique Théorique et Appliquée*, Tomo Primeiro, *Statistique Théorique*, 1921, pág. 37.

⁷ A. JULIN — Ob. cit., I, pág. 35.

⁸ RÜMELIN — *Problèmes d'Économie Politique et de Statistique*, trad. franç. de A. DE RIEDMATTEN, 1896.

la as peculiaridades características de cada unidade e não manifesta se são boas ou nocivas; ao passo que a História nos fornece o próprio drama. Na História, campeiam os heróis; na Estatística, as coletividades anônimas”.

Também não é fácil de dizer, de maneira clara e precisa, em que diferem a Estatística e a Economia Política. VIRGILII, ao tratar do assunto, dá-nos explicações do mesmo teor das que formula em relação à História. Os demais autores, cautelosamente, nada dizem, preferindo o silêncio a explanações redundantes e desprovidas de significação ou fundamento. Os economistas, igualmente, não se aprofundam no caso.

No grande Curso de CHARLES GIDE⁹ encontramos interessante observação a respeito:

“Notamos ainda, diz êle, duas outras ciências, ditas também ciências sociais, a História, que estuda a ordem de sucessão dos fatos, e a Estatística, que recolhe e classifica todos os fatos suscetíveis de serem expressos por números. Todavia, é preciso acentuar que essas duas ciências, ao contrário das acima mencionadas (a Moral, o Direito, a Política, a Lingüística, a Ciência das Religiões), não só não têm por objeto uma categoria especial de fatos sociais, como se aplicam a todos. Porque há uma história não só da Economia Política, como da Religião, da Moral, — e até há uma “História Natural” e uma “Paleontologia”, que é da formação das espécies vivas ou que existiram na terra. Do mesmo modo, há uma estatística judiciária, moral, política, demográfica, etc., tal como uma estatística econômica ou financeira.”

CAMILLE PERREAU¹⁰, atribuindo à Estatística o papel de principal ciência auxiliar da Economia Política, reflete que a Estatística se limita a registrar os fatos na sua sucessão e dar-lhes representação numérica ou gráfica, não se preocupando com determinar-lhes as causas, de ordinário, complexas. A Economia Política, ao contrário, deve estudar o meio e as condições em que se produzem os fenômenos, aproveitando-se, sempre que for necessário, das indicações fornecidas pela Estatística.

Servindo-se de números, para os respectivos estudos, a Estatística tem que atender, muito particularmente, às leis e aos problemas de que tratam as ciências matemáticas.

“Os grandes números, diz ARTHUR BOWLEY¹¹, não são contados, corretamente, até a unidade, porém, apenas, avaliados; e talvez possamos tomar êsse fato como uma distinção entre a Aritmética e a Estatística; pois, se a Aritmética alcança a exatidão, a Estatística tem de contentar-se com estimativas, às vezes bastante aproximadas, e em geral suficientes para os seus objetivos, porém, nunca, matematicamente exatas.”

A. JULIN dedica a êsse ponto algumas páginas do seu grande tratado, onde cita e apóia a opinião de FERRARIS de que “a Estatística reclama, de quem deseja estudá-la a fundo, certa dose de conhecimento (uma discreta cultura) das ciências matemáticas”.

Resumo histórico

Resta-nos, agora, para sair do terreno das generalidades, resumir a história da Estatística, a qual, no ponderado parecer de BENINI¹² “não deve merecer, num tratado, e muito menos num simples opúsculo, senão um lugar secundário, por isso que a história de uma ciência não é a própria ciência; pode ser (e, acrescentamos nós, sempre é) uma ordem de conhecimentos, extremamente interessantes, acêrca do progresso do pensamento humano, nesta ou naquela direção, porém, não é parte necessária dos fatos, sobre os quais o pensamento se exercitou e continua a exercitar-se”.

Apontaremos, pois, aqui, os principais lances da história da Estatística, em geral repetidos, quase textualmente, pelos escritores que se ocuparam desta ciência.

No livro clássico de GABAGLIO¹³ os leitores encontrarão o mais completo e curioso repositório de informações, sobre as vicissitudes históricas da Estatística, como ciência e como método.

Já no ano 2238, antes de Cristo, faziam-se, na China, estatísticas oficiais, conforme se depreende de uma referência de CONFUCIUS no 1.º capítulo da 2.ª parte do *Chou-King*. Nas mais remotas eras, os egípcios, assírios, persas e indianos possuíam órgãos rudimentares de coleta de dados estatísticos. Em diversos trechos, e até em livros, um dêles denominado “dos Números”, a Bíblia contém positivas provas do grande interesse em que eram tidas pelos hebreus, na Judéia, as verificações estatísticas (até mesmo recenseamentos), amiúde realizadas para fins administrativos. Na Grécia e em Roma, funcionavam registos de nascimentos e óbitos, deparando-se nos escritores de ambos êsses grandes focos de cultura e civilização, tais como PLATÃO, PLUTARCO, XENOFONTE e ARISTÓTELES, DIONÍSIO, POLÍBIO, CÍCERO, TITO LÍVIO TÁCITO, significativas e interessantíssimas alusões às práticas estatísticas, então em uso. Basta lembrar o censo e o cadastro romanos.

No seu livro sempre pitoresco, BLOCK, a esta altura da resenha da história da Estatística, “saúda” CARLOS MAGNO, por ter sido curioso de informações numéricas e ordenador de levantamentos e inventários.

Também fez jus a cumprimentos GUI-LHERME o Conquistador (1086) com seu *Liber Censualis*, ou *Doomsday Book* (cuja tradução é *Livro do dia do juízo final*), considerado por BENINI um documento estatístico-administrativo de notável importância, por ser um verdadeiro recenseamento dos vassallos e dos livres, e completo cadastro das terras, minas, etc., existentes na Inglaterra.

Voltam, agora, os eruditos a sua atenção para Veneza, com os *Atti della Republica*, as *Relazioni* dos embaixadores, e os recensea-

⁹ CHARLES GIDE — *Cours d'Économie Politique*, Tomo Primeiro, 1908, pág. 2.

¹⁰ CAMILLE PERREAU — *Cours d'Économie Politique*, Tomo Primeiro, 1914, pág. 32.

¹¹ ARTHUR L. BOWLEY — *Elements of Statistics*, 1920, pág. 3.

¹² RODOLFO BENINI — *Principii di Statistica Metodologica*, 1906, pág. 36.

¹³ ANTONIO GABAGLIO — *Teoria Generale della Statistica*, 1888.

mentos da população, também procedidos em Florença.

A organização dos Estados, na idade moderna, profundamente auxiliada pela difusão da imprensa, exigiu a realização de numerosos trabalhos de inegável cunho estatístico, e deu lugar à sua narrativa, em livros redigidos com o auxílio oficial, de que são exemplos memoráveis o de FRANCESCO SANSOVINO, *Del governo e amministrazione di diversi regni e republiche, etc.*, o de LUIGI GUICCIARDINI, *Descrizioni di tutti i Paesi Bassi, etc.*, o de GIOVANNI BOTERO, *Relazioni universali*, e o de PIERRE D'AVITY, *Les États, Empires et Principautés du Monde*.

A idéia de aproveitar, sistematicamente, essa massa considerável de informações, sobre as condições políticas e econômicas dos Estados, deu origem a duas orientações, nitidamente distintas, por parte dos sábios e dos professores das Universidades de então, formando-se duas correntes, uma na Inglaterra e outra na Alemanha, iniciadas, respectivamente, por WILLIAM PETTY e HERMAN CONRING.

Subordinada a objetivos práticos e restritos, qual, por exemplo, o de construir tábuas de mortalidade e calcular tabelas de seguros de vida, a primeira dessas escolas, ilustrada pelos trabalhos de PETTY, JOHN GRAUNT, KING e do célebre astrônomo HALLEY, tornou-se conhecida pela original denominação de Aritmética Política, empregada, pela primeira vez, por WILLIAM PETTY, no seu livro *Several essays on political arithmetic*.

Os economistas acima citados, e seguindo os mesmos processos, o holandês KERSEBOOM e o francês DÉPARCIEUX sujeitaram os números, então existentes nos registros civis de seus países, aos cálculos mais variados possíveis, a fim de extrair das relações observadas entre êles determinadas conclusões, que serviram de bases para completar lacunas, ou fazer conjecturas e estimativas concernentes a outros fenômenos, insufficientemente representados por meio de números absolutos ou relativos.

JOHN GRAUNT, tendo analisado, com toda a minúcia, os registros da cidade de Londres, foi o primeiro a divulgar a predominância dos nascimentos masculinos sobre os femininos, e a demonstrar que, do conhecimento do número de óbitos, se pode chegar, bem aproximadamente, ao número de pessoas vivas.

HALLEY criou o método de construção de uma tábua de mortalidade, deduzida, exclusivamente, da classificação dos mortos por idade.

Enquanto os aritméticos políticos assim realçavam, com utilíssimas aplicações práticas, o manejo dos números significativos de fatos sociais, os professores alemães persistiam em lecionar por meio de descrições e narrativas das condições políticas e econômicas dos principais Estados, então organizados.

Foi notável, a esse respeito, o longo ensinamento de HERMAN CONRING, docente da

Universidade de Helmstädt, pequeno recanto do Ducado de Brunswick.

CONRING, durante a sua vida, só publicou trabalhos sobre Direito Público, tendo, porém, deixado apontamentos suficientes para que seus discípulos, em 1730, publicassem um apanhado do curso professado em 1660, sob o título de *Exercitatio historico-politica de noticia singularis alicujus reipublicae*, e no qual, no dizer de BLOCK¹⁴, foi traçada a linha de demarcação que separa a Estatística e a Geografia, da História e da Política.

Segundo CONRING, a situação de um Estado deriva do influxo de quatro causas, qualificadas, em latim, conforme o uso da época, de "materialis", "finalis", "formalis" e "efficiens".

A causa "materialis" abrange o povo e o território; a "finalis", a tendência para melhorar as condições de vida e a possibilidade de conseguir essa melhoria; a "formalis", a forma de governo adotada; e a "efficiens", o poder público, que reside, principalmente, no chefe de Estado, e, secundariamente, nos instrumentos de execução das medidas necessárias à administração pública, por exemplo os ministros, os magistrados, etc. (englobados sob a denominação de causa "animata"), e o dinheiro (ao qual foi atribuída a designação de causa "inanimata").

O curso de CONRING repercutiu nas outras Universidades alemãs, onde SEKENDORF, STRUVIO e SCHMEITZEL também se ocuparam do assunto, mais ou menos nos moldes traçados pelo sábio de Helmstädt.

SCHMEITZEL foi o primeiro a empregar, em latim, o adjetivo "statisticum", do qual se aproveitou, anos depois, ACHENWALL, da Universidade de Göttingen, para criar o nome de "Statistik", e com êle batizar a nova ciência que, aos poucos, se vinha formando, nesses laboratórios fecundos, que eram as universidades daquele tempo. Aureolaram-no, por isso, com o título de "pai da estatística", honra que lhe não é conferida, hoje, pela unanimidade dos tratadistas, dizendo BLOCK que a primazia do emprêgo do nome cabe a SCHMEITZEL, e observando JULIN que, se ACHENWALL foi o pai do nome, não o foi da ciência.

Como quer que seja, registam os historiadores que ACHENWALL deu grande impulso ao estudo da nova ciência, cuja definição foi por êle estabelecida nos seguintes termos: "o conjunto do que é realmente notável num Estado, e dos fatores mais importantes do seu progresso e organização".

Mantendo-lhe o mesmo cunho descritivo, ACHENWALL adiantou-se a CONRING, segundo a opinião de BENINI, tornando a narrativa do que principalmente caracterizava os Estados, menos escolástica, e mais independente do Direito Público, da História, da Geografia e da Política.

SCHLÖZER, sucessor de ACHENWALL na Universidade de Göttingen, durante a segunda metade do século dezoito, deixou uma fórmula muito curiosa, na qual tentou resu-

¹⁴ M. BLOCK — Ob. cit., pág. 5.

mir o objeto da Estatística: "vires unitae agunt", querendo significar, no primeiro termo, "vires", tudo o que constitui, materialmente, o Estado, isto é, a população, o território, a agricultura, a indústria, e o comércio; no segundo, "unitae", a reunião desses elementos propulsores, sob uma forma política; e na terceira "agunt", as manifestações do governo, da administração.

Já nos referimos, linhas atrás, à célebre definição em que SCHLÖZER estabeleceu um contraste, mais original do que verídico, entre a Estatística e a História. Por aí se vê quanto deveria ter sido interessante o ensinamento do grande professor, cujos trabalhos muito concorreram para o progresso da nova ciência. Até então, era quase absoluto o predomínio das descrições sobre o ensino baseado nos números. A influência, todavia, da escola inglesa, formada pelas engenhosas concepções de PETTY, GRAUNT, etc., não tardou em atuar dentro das próprias Universidades alemãs.

Não podendo eximir-se de analisar e experimentar os inteligentes processos de investigação, postos em prática pelos aritméticos políticos, os estatísticos alemães renderam-se à evidência de que, em verdade, estava na apreciação dos fenômenos expressos em números a verdadeira razão de ser da Estatística, assim, finalmente, destacada das ciências irmãs, que são a História, a Geografia e a Política.

Muito, para isso, valeu a contribuição de JOHANN PETER SÜSSMILCH, com o seu trabalho sempre citado *Die goettliche Ordnung in den Veränderungen des menschlichen Geschlechts, etc.*, ou seja *A ordem divina nas variações do gênero humano, provada pelos nascimentos, pelos óbitos e pela reprodução dos homens*.

E, conforme expõe JULIN, SÜSSMILCH foi muito além das estatísticas, procurando demonstrar através dos quadros organizados com elementos, conscienciosamente, extraídos dos registros diocesanos da Prússia, a existência de normalidades, ou leis, queridas pela Divindade e contrariadas por obstáculos que, apesar disso, não impedem a verificação de tendências gerais.

Assinalam os historiadores que, daí por diante, houve, entre os professores das Universidades alemãs, dupla corrente, manifestando-se, uns, a favor da escola ACHENWALL, de limitar a Estatística ao estudo descritivo do Estado, e orientando, outros, em vez, os seus discípulos ou leitores, para o trabalho difícil e tão ingrato, mas às vezes extraordinariamente fecundo, do preparo de tabelas numéricas, que servissem de material para a pesquisa de explicações, mais ou menos prováveis, dos fenômenos da vida social.

Como era natural, formou-se, também, um terceiro grupo de estatísticos, aplicando, ecleticamente, ambos os processos, sem se ater, piirrnicamente, a esta ou àquela escola, mas visando, apenas, a obter os melhores resultados, com as investigações realizadas.

Observa-se, entretanto, certa predominância dos princípios postos em voga pelos adeptos de ACHENWALL quando, ao poderoso

influxo das teorias de QUETELET, passou, visivelmente, para o primeiro plano, a diretiva iniciada pelos aritméticos políticos, e profundamente aperfeiçoada pelo sábio belga, cujo livro intitulado *Lettres sur la théorie des probabilités appliquées aux sciences morales et politiques*, contém apreciações sobre o papel da Estatística, até hoje não sobrepujadas pelas de qualquer outro escritor.

Jogando com a Matemática, que para ele não tinha segredos, QUETELET firmou a Estatística em sólidas bases científicas, dando-lhe o elevado objetivo de, mediante o emprego de todos os recursos proporcionados pelo cálculo, estudar as manifestações humanas no meio social, reveladas pelos elementos numéricos, inteligentemente coligidos e inventariados, a fim de conhecer os característicos essenciais de tais manifestações, e descobrir as leis a que as mesmas estão adstritas, no seu desenrolar, à primeira vista desorientado e incoerente.

Depois de QUETELET, eminentes sábios de outros países publicaram cursos ou tratados, nos quais, com maior ou menor amplitude e clareza, são expostos os princípios que regem a Estatística, enquanto, praticamente, esta é realizada nas numerosas repartições oficiais, existentes em todos os Estados civilizados, graças ao exemplo da França, onde, desde 1796, funciona um serviço especial de Estatística; da Alemanha, possuidora de semelhante órgão desde 1805, e da Inglaterra, com um "Statistical Department" desde 1832.

Para o rápido aperfeiçoamento dessa disciplina muito concorreram os Congressos de Estatística, iniciados em 1853, em Bruxelas, graças à persistente campanha de QUETELET, e reunidos pela última vez, em 1876, em Budapest.

Fundado em 1885, o Instituto Internacional de Estatística tomou a peito a repetição de tão fecundas assembléias científicas, o que foi conseguido, com regularidade, de dois em dois anos, até 1913.

Sobrevinda, no ano seguinte, a grande conflagração dos povos, só foi possível ao benemérito Instituto recomençar as suas reuniões em 1923, na capital da Bélgica.

Relevante acontecimento histórico para a Estatística foi, sem dúvida, a criação, por parte de numerosos países, do Instituto Internacional de Agricultura, com sede permanente em Roma, e dotado de magníficas seções incumbidas da compilação, em moldes rigorosamente científicos, dos elementos fornecidos pelas estatísticas oficiais ou particulares, sendo os resultados desses trabalhos publicados, periodicamente, em substanciosos boletins e anuários.

Uma das provas mais patentes da importância e do interesse atribuídos, no século dezanove, às investigações estatísticas, reside na fundação de Sociedades de Estatística, de que são exemplos célebres a "Royal Statistical Society" de Londres, criada em 1835, a "Société de Statistique" de Paris, organizada em 1860, e as sociedades congêneres da Suíça e dos Estados Unidos.

De grande vantagem para o Brasil seria a instalação, em sua Capital, de um grêmio, cujos estatutos fossem moldados pelos das Sociedades a que acabamos de nos referir, e que constituiria mais um desses "úteis focos de discussões e estudos desinteressados", na frase de FERNAND FAURE¹⁵.

ESTATÍSTICA METODOLÓGICA, TEÓRICA E TÉCNICA

Vamos, agora, estudar as normas consagradas pela prática e analisadas e aperfeiçoadas pela teoria, para a organização de uma boa estatística.

Divisão da metodologia estatística

Em cinco partes pode ser subdividido esse estudo.

Há que considerar, primeiro, a "coleta dos fatos", fórmula adotada pelo Sr. AUGUSTO DIAS CARNEIRO para traduzir a "rilevazione", de VIRGILII e demais autores italianos, ou o "relevé", dos escritores de língua francesa.

BOSCO, em suas magistrais *Lezioni di Statistica*¹⁶, explana o conceito da "rilevazione", dizendo quase tratar-se de "rilevare", como faz o topógrafo em relação aos vários acidentes do terreno, as particularidades da vida social e dos fenômenos coletivos. Ora, em topografia, e em língua portuguesa, é empregado nesse sentido o vocábulo "levantamento", o qual, portanto, também poderá ser aplicado ao conjunto das operações que formam o "relevé" ou a "rilevazione".

Diremos, pois, indiferentemente, levantamento, ou coleta dos fatos.

A êste propósito, é ainda de BOSCO a advertência sutil de que o levantamento é de "fatos" e não de "dados", visto coletar-se a "unidade" do fato, e não o "dado", que só aparece depois, com o agrupamento das unidades, na fase ulterior da organização da estatística, denominada elaboração.

A segunda parte do Estudo da Estatística metodológica é, mais geralmente, atribuída ao exame das questões atinentes à crítica, a que deve ser submetida qualquer operação estatística.

Pondera F. VIRGILII que, a rigor, a exposição dos preceitos da "crítica" poderia ser feita após o estudo das normas requeridas pela "elaboração", não obstante também necessitar de conhecer as cautelas aconselhadas pela "crítica" quem houver de fazer passar pelo período da "elaboração" uma determinada estatística. Em todo o caso, não parece que haja motivo preponderante para justificar uma modificação na ordem, como dissemos, geralmente aceita pelos tratadistas.

Ocupará, pois, o terceiro lugar, no estudo do processo estatístico, o exame do que diz

respeito à elaboração dos dados, isto é, à respectiva "apuração", por processos manuais ou mecânicos, obedecendo a plano preestabelecido e realizando-se, praticamente, pela transformação, em cartolinas, mapas e quadros primários, dos números adrede coletados.

A quarta parte abrangerá a exposição dos fatos, através dos dados, convenientemente resumidos em quadros, contendo números absolutos e relativos, e mediante gráficos que assinalem, bem ao vivo, os resultados da operação estatística.

Finalmente, a quinta parte englobará as observações que forem cabíveis acerca da interpretação, a ser dada aos resultados do trabalho, cientificamente planejado, levantado, examinado, apurado em seus pormenores, e exposto em síntese, e que ponha em destaque os frutos de tão complexas e delicadas operações.

LEVANTAMENTO OU COLETA DOS FATOS

A própria natureza dos fatos sociais, que se queiram pesquisar pelo método estatístico, indicará ao observador qual o processo de levantamento, ou de coleta, mais conveniente, ou mesmo, único, a ser escolhido, ou aplicado, para aquêle fim.

Levantamento direto

Se as manifestações dos fenômenos individuais, a serem analisados por meio da contagem estatística, permitem o respectivo conhecimento, um por um, com igual certeza, quanto aos seus característicos essenciais, ou, pelo menos, principais, é evidente que o levantamento quantitativo e qualitativo desses fatos será realizado diretamente, pelo observador, que dêles tomará a devida nota em instrumento de coleta, previamente preparado, e obedecendo a tôdas as prescrições assentadas pelo autor do plano a ser executado.

O levantamento direto é, pois, o processo naturalmente empregado, sempre que se queira e se possa realizar a investigação estatística de fatos sociais, cujo estudo interesse à sociedade, coletando-os diretamente, nas mesmas condições de tempo e de espaço, preestabelecidas.

Dir-se-á que estamos definindo o levantamento direto, com a simples repetição desse adjetivo. Aceitando a censura, excusamos-nos de rebatê-la, pois o mesmo fizeram, até agora, todos os tratadistas, impossibilitados como, decerto, se viram de explicar, por outra forma, a significação por si só bastante clara do termo "direto".

Alguns exemplos, todavia, dissiparão quaisquer dúvidas a respeito. Começemos pela estatística mais importante, o recenseamento da população. Inicia-se ela pelo levantamento direto dos indivíduos que habitam um território, alcançando-os um por um, por meio de boletins individuais ou de listas de família, distribuídos de maneira a não ficar excluída pessoa alguma existente no país, e na data prefixada.

¹⁵ FERNAND FAURE — Ob. cit., pág. 20.

¹⁶ AUGUSTO BOSCO — *Lezioni di Statistica*, recolhidas e compiladas por P. MENZARINI e A. TAMBURINI, Primeira Parte, "Metodologia Estatística", 1909, pág. 151.

Também diretas são as coletas dos nascimentos e óbitos, fatos êsses obrigatoriamente registados em obediência ao Código Civil, e, assim, passíveis de serem mencionados, com os mesmos pormenores, nos instrumentos de coleta, fornecidos aos cartórios. Direto é ainda, o levantamento dos eleitores, devidamente alistados, das hipotecas inscritas em registos de imóveis, e dos institutos de ensino, sujeitos à fiscalização oficial.

A rigor, diz BENINI¹⁷, só se compreende o levantamento como direto, e consistindo numa enumeração de casos. Pouco importa, afirma êsse grande estatístico, que os casos sejam poucos ou muitos, tomados a êsmo ou escolhidos tendo em vista os característicos quase-típicos, coletados com precisão ou mediante estimativas e medidas aproximadas.

O que os outros denominam levantamento indireto ou conjectural é, pròpriamente, "indução" ou "dedução", e deve ficar na base da Estatística, considerada, não como forma de observação, mas, como forma de raciocínio. Se eu calculo, pondera ainda o escritor, a riqueza privada de um país, levando, apenas, em conta, os dados relativos à sucessão hereditária; se, pelo número conhecido das casas, deduzo o dos habitantes, etc., — eu raciocino, não observo, não coeto fatos.

Em que pese à extraordinária autoridade de quem expendeu os conceitos acima transcritos, somos inclinados a considerar, com a generalidade dos estatísticos, como fundada e merecedora de ser conservada a divisão dos levantamentos em diretos e indiretos ou conjecturais, visto exprimir, perfeitamente, o contraste entre uma e outra coleta de fatos, a distinção decorrente dos dois adjetivos, e de serem, ainda, extremamente úteis, e, às vêzes, indispensáveis, os levantamentos conjecturais, sobretudo nos países de vastos territórios ou grandes aglomerações humanas.

Levantamento indireto ou conjectural

Procede-se a um levantamento conjectural, quando se calculam os números representativos de determinados fenômenos sociais em proporção a outros fatos coletados diretamente; ou quando se avaliam aquêles números pelas informações prestadas por pessoas entendidas e fidedignas; ou, também, tendo em vista a verificação quantitativa de outros fenômenos que com êles se relacionem, por coexistência ou sucessão.

No primeiro caso, diz-se que o levantamento é por proporcionalidade, e, nos demais, por estimativa.

A proporção poderá ser da parte ao todo, como se faz com a estatística postal, cujo desenvolvimento anual é baseado nas contagens, geralmente realizadas durante uma semana; ou de um fenômeno a outro, desde que exista uma relação estreita entre ambos, se possível, de causa a efeito, como se dá com o cálculo da população pelo número dos nascimentos e óbitos, e o das condições morais do país pelo número de crimes, de filhos ilegítimos, etc.

As estimativas ainda são mais conjecturais, por isso que partem de apreciações, mais ou menos aproximadas, feitas até "a olho", no dizer de COLAJANNI¹⁸, ou por analogia, como sejam, na primeira hipótese, a maior parte das estatísticas de produção agrícola, e, na segunda, o cálculo da população pelo número de casas ou pelas quantidades de gêneros importados e consumidos.

Esta hipótese muito se assemelha à conjectura por proporcionalidade, embora os autores só a mencionem como exemplo de estimativas por semelhança ou analogia.

Cumpra, todavia, acentuar que todos os princípios que presidem à correta execução dos levantamentos diretos devem ser respeitados ou servir de orientação no tocante a quaisquer levantamentos indiretos ou conjecturais, os quais, por muito tempo, hão de ser preferidos ou impostos pelas circunstâncias.

Os levantamentos conjecturais têm sido alvo de irônicas ou mesmo acerbas críticas por parte dos professores de Estatística, que trazem sempre à baila as curiosas estimativas de VILLANI, VAUBAN, YOUNG, NECKER, LAVOISIER, e outros, para assinalar os erros que ocasionam tais levantamentos.

VILLANI calculou a população de Florença pelo consumo do pão.

VAUBAN fêz determinar a extensão média ocupada, em cada légua quadrada de certos trechos do território, pelas terras de cultura, vinhas, pastagens e matas e quais as quantidades produzidas naqueles pontos. Generalizando os dados obtidos, aplicou-se a tôda a superfície da França, avaliando, assim, a população agrícola provável.

ARTHUR YOUNG, agrônomo inglês, anotou, numa carta geográfica da França, tudo quanto observara em suas minuciosas excursões. Cortou em pedacinhos êsse mapa, pesou os retalhos, e da proporção entre os pesos deduziu a das superfícies cultivadas e respectivas produções, nas diferentes províncias em que se subdividia a França.

NECKER presumiu a população por meio de coeficientes de natalidade, e LAVOISIER, em 1790, avaliou a produção agrícola tomando por termo de comparação o número de arados então existentes.

É evidente que o caminho seguido por VAUBAN foi mais seguro, embora ainda inaceitável pela amplitude que lhe deu, do que os trilhados por YOUNG e LAVOISIER.

Aliás, VAUBAN, além de grande administrador, foi um "maravilhoso estatístico", na opinião de LÉON SAY¹⁹, e, no pensar do mesmo economista, "muitos dados estatísticos coligidos por VAUBAN resistem à crítica com que se queira analisá-los em nossos tempos".

Assim, esquecem-se êsses censores, do grande serviço prestado pelos citados estadistas ou sábios, proporcionando valiosíssimos elementos de apreciação numa época em que não era admissível qualquer outro processo estatístico, e dando ensejo a que, pela verificação dos erros, que homens de tanto enge-

¹⁷ R. BENINI — Ob. cit., pág. 52.

¹⁸ N. COLAJANNI — Ob. cit., pág. 47.

¹⁹ Segundo F. FAURE — Ob. cit., pág. 14.

nho não puderam evitar, os seus sucessores, no governo ou na ciência, introduzissem os necessários aperfeiçoamentos na teoria e prática da estatística.

Levantamentos contínuos, periódicos e ocasionais

Os levantamentos podem ser contínuos, intermitentes ou periódicos, e ocasionais, conforme se efetuem sem descontinuar, ou somente em certas épocas, ou uma determinada vez.

É claro que será contínua a coleta do movimento da população (nascimentos, casamentos e óbitos), do comércio exterior (importação e exportação) e do movimento migratório (imigração e emigração).

Será intermitente ou periódica a grande coleta exigida pelos recenseamentos, e serão ocasionais todos os levantamentos que o exame de determinados assuntos, interessando à coletividade, torne convenientes para a devida orientação, sobretudo, dos governantes e legisladores.

Em geral, mesmo existindo um registo ininterrupto de fatos sociais, permitindo coletas, para fins estatísticos, estas só se fazem, periódica ou ocasionalmente, conforme os programas de trabalho dos observadores, particulares ou oficiais, e os recursos disponíveis para esse fim.

O comum, por esse motivo, é o levantamento periódico, mensal ou anual, mediante o qual se coligem os elementos destinados à organização de numerosas estatísticas, abordando os principais aspectos da vida social.

Levantamentos automáticos e reflexos

Costuma-se, ainda, encarar os levantamentos sob o ponto de vista do automatismo da sua execução, dizendo-se que será automática a coleta contínua, baseada em registos sempre anotados, à medida que se dão os fatos, e reflexa, querida, ou proposital, a coleta periódica, cujo funcionamento estará, sempre, dependendo de iniciativas dos observadores, mediante inquéritos estatísticos especiais.

Essas distinções carecem de importância e dispensam maiores comentários.

Plano do levantamento

Este deve reportar-se, agora, aos requisitos de um bom levantamento estatístico, cujo ponto de partida reside num plano ou programa de ação, traçado por quem conheça, razoavelmente²⁰, o assunto sobre que versará a coleta, dentro dos limites de tempo, de espaço, de especialização e de precisão, que um exame minucioso indique deverem ser observados durante a operação estatística.

É verdade que, a esse respeito, ainda não existe sistematização completa dos prin-

cípios a serem acatados pelo lançador de um inquérito.

Na capacidade, revelada em tentâmens anteriores; na prática, adquirida em trabalhos realizados sob sua responsabilidade, após haver labutado sob direção superior; na assimilação dos processos usados em casos semelhantes, e, sem dúvida, no dom, que demonstre possuir, de efetuar com êxito esses difíceis empreendimentos, — encontrará o planejador de uma investigação estatística luzes suficientes para organizar, em todos os seus pormenores e visando ao máximo resultado com o mínimo custo, o plano da pesquisa geral ou limitada, centralizada ou descentralizada, dar-lhe o impulso inicial, e prosseguir, em seguida, na direção dos trabalhos até satisfatória conclusão.

"Um plano ou programa de ação, diz-nos quem os tem, brilhantemente, idealizado²¹ e tornado realidade, abrangerá a determinação precisa do objeto do levantamento, o estudo do custo da operação, um ensaio de apresentação dos dados, e, nos casos complicados e novos, disposições para proceder-se a uma experiência de recenseamento. Além disso, fixar-se-á a época do levantamento, os prazos previstos para o seu termo, a unidade geográfica em que se passam os fatos, as vias e meios de execução, as causas de erros e os meios de fazê-los desaparecer."

Unidades a coletar

E o mesmo autor, desenvolvendo, magistralmente, esse ponto de metodologia estatística, acentua a necessidade de ser bem definida, na acepção vulgar, ou jurídica, ou econômica, a "unidade" que se pretende coletar, de modo a não lhe faltar qualquer das seguintes qualidades essenciais:

- 1.^a) pertencer à ordem dos fenômenos coletivos;
- 2.^a) poder ser reduzida a uma expressão numérica;
- 3.^a) ser diferenciada das outras cousas.

Tais são as condições indispensáveis para o êxito da investigação, devendo, mesmo, dar-se à "unidade", a ser coletada, uma definição estatística especial, se, ao exame atento do observador, não parecer suficiente a definição que lhe fôr dada, habitualmente.

Só raras vêzes se poderão coletar tôdas as unidades estatísticas, havendo casos, todavia, em que o levantamento perderá tôda a utilidade, se não alcançar tôdas as manifestações em aprêço. A estatística dos salários será possível, mesmo que se não colemem todos os salários pagos no País, não sendo, entretanto, aceitável uma estatística financeira das Unidades de uma Federação, com falta de um ou mais Estados.

É bem compreensível a distinção entre os fenômenos estatísticos e os dinâmicos, ou de movimento. Conforme se possam conhecer esses fenômenos, com maior ou menor facilidade e profundidade, se decidirá do intervalo entre as investigações sucessivas do mesmo grupo estudado.

Fazendo essa reflexão, BENINI²² analisa, em seguida, as limitações decorrentes das

²⁰ FERNAND FAURE — (Ob. cit., pág. 68), condena, como exagerada, a opinião de LIESSE, de ser necessário ao estatístico um profundo conhecimento da ciência ou da arte a que se relacionam os números que vai interpretar.

²¹ A. JULIN — Ob. cit., I, pág. 109.

²² R. BENINI — Ob. cit., págs. 42 e seguintes.

quantidades a serem coletadas, em função do tempo disponível e dos espaços considerados, bem como do conhecimento preciso dos fatos, terminando com a forçada conclusão de "caber ao homem prático encontrar, mesmo por tentativas, o sistema de limites mais conveniente, tendo em vista os interesses científicos e administrativos empenhados na pesquisa, as despesas, os obstáculos prováveis e outras circunstâncias aindas".

Para coletar as unidades estatísticas, que, no dizer de MESSE DAGLIA, repetido por VIRGLII²⁵, devem ser importantes, relativamente ao gênero de observação e ao fim a que devem servir; verídicas, isto é, provirem de fontes dignas de fé; precisas, isto é, colhidas na sua mais exata expressão; e uniformes, isto é, não contendo nem mais nem menos do que aquilo que se pediu, torna-se mister empregar determinados instrumentos ou fórmulas, constituídos por perguntas, cuja resposta constituirá a completa caracterização numérica dos fatos a investigar.

Fórmulas de coleta

Questionário é o nome mais intuitivo e genérico dessas fórmulas de coleta.

Para JULIN²⁴, o termo "boletim" é mais amplo do que o "questionário", devendo este ser reservado para os inquéritos de pequena envergadura, e aplicar-se aquêle aos formulários das grandes estatísticas, como, por exemplo, os recenseamentos.

Acrescenta, porém, o mesmo autor, que essa distinção nada tem de absoluto, o que nos permite dar preferência ao termo "questionário", cuja raiz é muito mais rica de significação, e exprime, incontestavelmente, com outra clareza, o conteúdo desses instrumentos de coleta.

Apresenta-se também com o aspecto de "quadro" a fórmula estatística traçada em linhas verticais e horizontais, constituindo colunas, com inscrições laterais e superiores, denominadas estas de "cabeças", e aquelas sem nome especial, desde que não sejam em forma de "quesitos" ou perguntas.

Nesses formulários deverá concretizar-se todo o acurado estudo preliminar da estatística a ser levantada.

Neles se materializarão as conclusões das pacientes análises e sínteses do campo em que se vai desdobrar a operação estatística, do programa a ser cumprido, e dos resultados, conscientemente esperados com a realização do inquérito.

Nos inquéritos diretos, nem sempre são necessários questionários, boletins, ou quadros, para a coleta estatística.

documentos em que se registam os fatos, procedendo-se, sobre eles, às ulteriores operações de apuração dos elementos escolhidos

Bastará, em muitos casos, o fornecimento, ao observador, de cópias autênticas dos para a estatística.

As cópias dos conhecimentos das estradas de ferro são bons instrumentos para uma estatística dos transportes.

Do mesmo modo, são úteis para a estatística migratória as listas dos passageiros das diferentes classes, organizadas, em várias vias, pelas companhias de navegação.

A exportação de mercadorias para o exterior, ou de cabotagem, também é levantada sobre os manifestos das embarcações.

É verdade que, para manter uniforme e sem lacunas essa coleta, foi tornado obrigatório o uso de um só modelo para esses formulários, nada impedindo, todavia, que outros inquéritos se realizem, ao menos nas primeiras vezes, com o emprêgo dos originais ou de cópias dos próprios documentos em que se descrevam, numéricamente, os fatos a estudar.

Quesitos e suas condições

Os quesitos, ou as indicações do que cumprirá informar, devem ser sóbrios, ou sucintos, simples e claros, específicos, isto é, com uma só interpretação, coordenados, categóricos, preenchíveis por números absolutos, e bem dispostos tipograficamente.

Não custa exigir essas qualidades para o quesito; trabalhosa, entretanto, e inçada de ingentes dificuldades é a tarefa de redigir um questionário com todos esses predicados, visto a preocupação de sobriedade e de clareza, ou, mesmo, de boa impressão tipográfica, ir, não raro, de encontro, ao não menor anseio de coordenação e de especificidade.

A sobriedade deve entender-se quer em relação ao número de quesitos, quer relativamente a cada quesito de per si, insistindo os autores na recomendação de organizar-se o questionário tão somente com os quesitos indispensáveis à coleta das informações essenciais, porquanto, no dizer de GABAGLIO²⁶, e também de COLAJANNI, mesmo interrogando pouco, grande será a fadiga para obter as respectivas respostas.

O segredo de um questionário está, segundo BLOCK²⁷, na conciliação de duas cousas quase contraditórias: pedir realmente muito, dando a impressão de pedir pouco.

Cada quesito deverá ser formado de uma frase sucinta, constituída dos termos mais inteligíveis e adaptáveis ao assunto, sem a possibilidade de parecer uma indagação tendenciosa, para fins que ponham de prevenção o informante. Daí, a necessidade de serem os quesitos, igualmente, simples e claros, em linguagem acessível até aos menos instruídos, e suscetíveis de resposta breve, embora completa.

Para evitar que esta seja ambígua, ou alheia ao objetivo do quesito, é que se exige, do mesmo, predicado da especificidade que se poderá, ainda, acentuar, acrescentando uma explicação oportuna, uma nota ou uma palavra, que desfaça a menor dúvida e obrigue o informante à verdadeira interpretação do sentido do quesito.

²⁵ F. VIRGLII — Ob. cit., pág. 28.

²⁴ A. JULIN — Ob. cit., I, pág. 131.

²⁶ A. GABAGLIO — Ob. cit., II, pág. 68.

²⁷ M. BLOCK — Ob. cit., pág. 314.

A coordenação das perguntas impõe-se como meio de fiscalizar as informações umas pelas outras, cumprindo concatená-las de forma que seja forçada a coerência recíproca, o que facilitará o exame da veracidade, ou do grau de aproximação desta.

Colimando provocar declarações indiscutíveis, tanto quanto possível, formuladas em algarismos, e impedir que se informe, vagamente, anotando-se nos questionários as expressões "pouco", "muito", "às vezes", "habitualmente", "grande", "pequeno", etc., os quesitos deverão ser categóricos, isto é, exigindo respostas positivas, numéricas, ou, quando muito, "sim" ou "não".

O esforço ou a habilidade de calcular médias ou percentagens deverá, em regra, ser poupado ao informante, cujo trabalho se resumirá a fornecer números absolutos, significando, concretamente, o caso tratado no quesito.

A boa disposição tipográfica, relativa à escolha dos tipos dos dizeres, à distância e ao arranjo destes, das linhas e das colunas, bem como ao tamanho dos espaços reservados para as respostas, será, também, alvo de cuidadosa atenção por parte do redator do questionário, cujo correto preenchimento muito lucrará com uma caprichosa e atraente impressão dada à fórmula de coleta.

Exemplos de fórmulas de coleta

Modêlo frisante de um bom instrumento de coleta é o que inserimos, em anexo, e foi empregado, com todo o êxito, pela Diretoria-Geral de Estatística, para o recenseamento da população em 1.º de setembro de 1920*.

Também correspondem às normas já enunciadas os questionários e boletins de que se está utilizando a mesma Diretoria-Geral de Estatística, para os seus trabalhos regulamentares, e dentre os quais escolhemos, a êsmo, os que juntamos ao do recenseamento.

Valiosos, ainda, pelo critério estatístico com que foram elaborados, são os boletins adotados pelo Serviço de Estatística Geral do Estado de Minas Gerais, conforme, facilmente, se depreende do exame, mesmo superficial, dos que se acham em apenso*.

Classificações

Para facilitar a resposta a um quesito, deverá o estatístico recorrer a uma seriação das hipóteses ou casos que, nas suas numerosas variedades, se enquadram no referido quesito.

As idades, as profissões, os locais, as causas, e tantos outros aspectos dos fenômenos a coletar, não poderão ser enumerados, na sua totalidade, no formulário. É mister resumir-los em classes gerais, agrupando-os conforme a respectiva analogia, e colocar, assim, à vista do informante, uma nomenclatura, tão exata e mesmo científica quanto

possível, mas sempre claramente inteligível, mediante a qual êle se sinta habilitado a prestar, com firmeza, o esclarecimento solicitado.

BERTILLON, que discorre, com tóda a proficiência, a êsse respeito²⁷, lembra o quesito das profissões, cujo total se eleva a mais de 3 000, e, por êsse motivo, foram por êle classificadas, sistematicamente, numa nomenclatura até hoje adotada, após haver obtido o assentimento do Instituto Internacional de Estatística.

Combate o mesmo autor a ordem alfabética, a que muitos estatísticos se apegam, ante as dificuldades que sempre apresenta a organização de uma boa nomenclatura.

Embora pareça um tanto exagerada a frase de BERTILLON, de que a ordem alfabética não é mais do que "uma forma da desordem", concordarão todos com êle em ser a ordem analógica mais racional e admissível, sobretudo quando combinada à disposição alfabética aplicada nas classes da nomenclatura escolhida.

Caso interessante, a êsse respeito, é a relação dos Estados do Brasil, a qual, segundo norma tradicional, respeitada pela Diretoria-Geral de Estatística, figura nas respectivas tabelas em ordem alfabética, em contrário à mais usada por outros serviços públicos e nas obras de Geografia e História, onde se encontra, geralmente, na ordem geográfica.

Se, no que concerne aos Estados, parece mais útil essa última ordem, não se dirá que seja mais aceitável relativamente aos Municípios, cujo número impossibilita qualquer ordem que não seja a alfabética.

Os questionários são enviados com circulares ou instruções, que devem possuir qualidades semelhantes às reclamadas para a redação dos documentos a que se reportam. Também servirão de preciosos ensinamentos as instruções que nortearam e acompanharam os questionários do recenseamento brasileiro de 1920.

Informantes

Numa síntese feliz das condições que devem apresentar as pessoas a quem se pede ou se exige o preenchimento de um questionário, MESSEDAGLIA ou GABAGLIO²⁸ reduzem-na a três: "sapere", "potere", "volere".

As pessoas informantes, de caráter particular ou investidas de funções públicas, humildes ou graduadas, devem compreender, em todos os seus aspectos, o assunto de que tratam os formulários e as instruções complementares, isto é, "sapere".

Além de serem capazes de assimilar o conteúdo dos quesitos e de estarem aptos para dar-lhes a resposta apropriada, os informantes devem ter a necessária liberdade, material e espiritual, de fornecer os dados re-

²⁷ J. BERTILLON — Ob. cit., págs. 57 e seguintes.

²⁸ COLAJANNI — Ob. cit., pág. 42, atribui-a a MESSEDAGLIA; GABAGLIO cita-a na pág. 69 do vol. II do seu tratado, e JULIN, Ob. cit., 1.º vol., pág. 145, considera-a como fórmula de GABAGLIO.

* Eliminados na presente reedição.

queridos, "potere"; e, por último, devem sentir-se impelidos a prestar os esclarecimentos, que lhes competirem, e a fazê-lo no devido tempo, com todo o escrúpulo e veracidade, o que se resume no verbo "volere".

Basta enunciar êsses requisitos para se concluir que, na consecução dêsse "desideratum", reside, sem dúvida, um dos encargos mais árduos de quem levanta uma estatística, sendo evidente, ainda, a probabilidade de mais rápido e melhor êxito para as coletas efetuadas por meio de agentes oficiais, da mesma esfera administrativa (por exemplo, inquérito federal a cargo de funcionários federais), em vez de particulares, em geral receiosos de facilitar a imposição de maiores tributos, ou impatrioticamente avessos, também por preguiça, a colaborar nesses trabalhos com a sua informação, por simples que seja.

O experimentado JULIN assim expõe, a êsse respeito, o fruto da sua prática profissional²⁰:

"Nos múltiplos casos em que tivemos de redigir instruções para os agentes de um inquérito, procedemos da seguinte maneira: a) de que natureza são as dificuldades do fato, contra as quais terá de lutar o recenseador? b) essas dificuldades estão expressamente previstas? c) tendo em vista tudo quanto está escrito nas instruções, como é que o agente poderá compreendê-las diversamente? Porque, acrescenta o autor, há pessoas dotadas de verdadeiro gênio de incompreensão, ou de deturpação no sentido natural das palavras. Urge, pois, se examinem, antecipadamente, as interpretações que se possam dar aos tópicos dos questionários e das circulares, somente conservando os quesitos ou instruções que resistirem a essa prova negativa."

Feita uma coleta dos fatos, será, não raro, mister proceder, logo após, a outras complementares, com o emprêgo de questionários simplificados, ou reduzidos a pontos que tenham sido omitidos nas respostas fornecidas. Novas circulares e instruções se tornarão precisas, e sua redação deverá obedecer às mesmas prescrições acima discriminadas.

Um dos maiores esforços exigidos para o andamento proveitoso de uma estatística, que abrange informantes localizados em pontos afastados do observador, está no arrolamento dêsses informantes, cuja relação, para ser completa, requer pertinácia e fertilidade de processos de investigação só avaliados, devidamente, por quem haja empreendido qualquer pesquisa naquelas condições.

Constituem êsses arrolamentos outras tantas estatísticas, a serem efetuadas com o máximo cuidado e rapidez, visto serem os alicerces dos resultados, que deverão aparecer à vista do público, sem que a maior parte dêsse faça idéia precisa do incrível trabalho despendido com a organização dos referidos arrolamentos.

Órgãos de coleta

O observador estatístico poderá ser individual ou coletivo, em caráter particular ou oficial, e, se auxiliado por observadores locais, será diligente, por estar no centro, ao

passo que os auxiliares, que se acham na periferia, se denominarão órgãos coletores.

É evidente que os particulares só poderão fazer levantamentos estatísticos em limites muito restritos, podendo êstes alargar-se quando os particulares não atuem isoladamente, mas sim associados, e em nome da agremiação a que pertençam.

Hodiernamente, as grandes empresas comerciais ou industriais dispõem de facilidades que permitem a realização de vastos e complexos inquéritos.

O estudo estatístico dos fatos que se passam sob sua direção já constituirá magnífico subsídio para êsse ramo das investigações econômicas.

Nos Estados Unidos, há sociedades que se ocupam, exclusivamente, de coligir e divulgar dados estatísticos, sendo, entretanto, controvertida a opinião dos que os consideram preferíveis aos dados coligidos oficialmente.

Inquéritos e monografias

Fazem, os autores, referências, mais ou menos minuciosas, à distinção entre as operações essencialmente estatísticas, como sejam os recenseamentos, os inquéritos, em geral, auxiliados pela aplicação do método estatístico, e as monografias.

O critério mais seguro para saber se se trata de uma dessas três variedades consiste na maior ou menor utilização de elementos numéricos, sendo a monografia o estudo de fenômenos sociais coletivos que mais emprega a descrição literária de todos os pormenores do assunto analisado, conferindo aos números a importância que não podem deixar de ter, mas não se limitando a simples manipulação de cifras ou algarismos, antes discorrendo, com tôdas as minúcias, acerca dos múltiplos aspectos dos fatos a investigar.

Deram merecido renome a essa modalidade de observação social os célebres trabalhos de LE PLAY e seus discípulos e continuadores, sobre a vida de famílias de trabalhadores da mesma classe, o funcionamento de fábricas e tudo quanto dizia respeito a pequenas comunas.

O inquérito, em geral, tem um pouco de monografia e o mais possível de estatística. A escolha do processo não está subordinada a regras preestabelecidas. Depende do fim a que se destinam os resultados a serem obtidos, dos recursos disponíveis e das possibilidades de pessoal, tempo e espaço, a determinação, pelo observador, do caminho a trilhar no lançamento da pesquisa, que deva ou pretenda fazer.

Uma vez fixada a maior ou menor amplitude a dar-se ao estudo estatístico, deverá ser traçado o programa a seguir, e, na conformidade dêste, o formulário rigorosamente estatístico (questionário, boletim ou mapa), ou uma lista de pontos a serem abordados pelos investigadores (inquérito geral), ou um arrolamento de tôdas as divisões e subdivisões em que se poderá discriminar o fato social em questão (monografia).

²⁰ A. JULIN — Ob. cit., I, pág. 147.

E, de acôrdo com o formulário ou procedimento preferido, far-se-ão as circulares e instruções correlatas.

CRÍTICA DOS DADOS

Nem sempre é fácil criticar

É sedição a frase de que a crítica é fácil. Entretanto, sob certos pontos de vista, não parece essa metade do famoso verso de BOILEAU tão verdadeira quanto a segunda, sendo tão árdua uma boa crítica como qualquer das artes em que se aplique a atividade humana.

A crítica é fácil, quando feita pelo simples impulso de criticar, sem levar em conta os méritos do que se analisa, nem os esforços empregados para chegar-se a resultado razoável.

Criticar, todavia, com isenção de ânimo, e tendo em mira, apenas, o aperfeiçoamento dos trabalhos examinados, requer reais qualidades de cultura e de prática, equivalentes ou superiores às dos que realizaram os empreendimentos criticados.

É o próprio observador deve ter o espírito crítico, para acompanhar tôdas as fases da coleta ou levantamento, bem como as ulteriores, da elaboração e exposição, com a indispensável desconfiança de estar incidindo, bem como seus colaboradores, em possíveis erros ou falhas, suscetíveis de serem, a tempo, corrigidos ou sanadas. É todo um código de ensinamentos o profundo lema atribuído ao Marechal FLORIANO: "confiar, desconfiando". Mas essa desconfiança não deve ser vaga, sem método, e despropositada.

Em Estatística, há preceitos a respeitar para que sigam seu curso normal as múltiplas operações técnicas, e por eles se deverão guiar todos quantos tomam parte nesses trabalhos.

Este é um dos pontos em que a Estatística e a História quase se confundem, aplicando-se aos documentos de ambas as ciências os princípios de investigação crítica estatuídos pelas respectivas metodologias.

No justo dizer de BOSCO³⁰, muitos documentos históricos não são mais do que dados estatísticos, por serem, uns e outros, igualmente, atestados de fatos acontecidos. Os históricos, todavia, são mais variados, e os estatísticos, de natureza recente, o que importa em ser mais necessária a crítica para os primeiros, sem embargo de nada impedir que se submetam os segundos ao rigoroso exame por que devem passar os históricos.

Causas que poderão viciar os dados

Encarando a crítica no sentido mais amplo, chega-se à conclusão de que um dado estatístico poderá estar viciado por uma causa accidental ou permanente.

Neste último caso, o erro se repetirá na totalidade ou maioria dos dados, por isso que deriva de uma causa constante, regular,

ou sistemática, que será encontrada por meio de uma crítica de sinceridade.

Na hipótese de ter havido deturpação da verdade, acidentalmente, é que para isso atuou uma causa fortuita, irregular ou variável, a qual será descoberta por uma crítica de exatidão.

Esta pode recorrer a pesquisas atinentes ao próprio levantamento, às diferentes operações em que o mesmo se desdobra, e, principalmente, ao conteúdo do questionário, boletim, ou mapa, fornecido pelo informante.

Este aspecto da crítica de exatidão costuma-se denominar de verificação interna, dando-se o nome de verificação externa à apreciação das circunstâncias materiais em que se realizou a operação estatística, dos cálculos então efetuados e da forma por que foi feita a impressão de tudo quanto diz respeito ao inquérito.

Crítica de sinceridade

Vejamos em que se deve basear a crítica de sinceridade, para escoimar de erros o trabalho estatístico.

O objetivo dessa crítica deve ser tirar qualquer dúvida acêrca da boa disposição do informante em querer prestar, corretamente, o esclarecimento pedido.

O receio de concorrer para o estabelecimento de medidas fiscais, ou de auxiliar novas regulamentações do trabalho; a relutância em pronunciar-se sobre questões religiosas ou de revelar taras físicas ou morais; a falta de estímulo para se dar ao trabalho de reunir todos os elementos de uma informação coerente e certa; ou, mesmo, o defeito, inerente a certas pessoas, de serem negligentes em seus atos, ou costumarem praticar com má vontade os que muito diretamente não lhes interessam; estes mesmos inconvenientes por parte dos agentes recenseadores, que também poderão ser pouco zelosos, indolentes, e incapazes de proceder cabalmente a uma determinada tarefa do inquérito; e, por último, as próprias tendências do observador, que persista em diretivas já condenadas por outras estatísticas; eis múltiplas causas consoantes de erros, a serem, atentamente, investigadas durante o prosseguimento da coleta e nos demais trâmites da estatística.

Os informantes devem estar em condições psicológicas que os induzam a ser fiéis para com o investigador e a repelir qualquer sugestão que, insensivelmente, os arraste a viciar o respectivo questionário, ante a perspectiva, que julguem provável, de serem, afinal, prejudicados nos seus interesses econômicos e morais.

Crítica de exatidão

Lembradas, assim as causas que podiam influir no espírito dos informantes ou na solicitude e aptidão profissional dos agentes estatísticos, ocasionando defeitos que devem ser evitados para não ser empanado o brilho de custosos e demorados inquéritos, vamos apontar os fatores que, eventualmente, pode-

³⁰ A. Bosco — Ob. cit., pág. 304.

rão intervir, em detrimento da certeza dos dados coligidos e a serem elaborados para a síntese final.

O êrro accidental de que se ressinta um questionário, num dos seus quesitos, deverá ser perscrutado com o paciente exame do ponto duvidoso, sujeitando-se essa parte do formulário, bem como tudo quanto com ela se relacione, a minuciosa análise lógica, na qual o verificador deverá pôr em jôgo tôda a sua sagacidade e ciência, a fim de apurar a verdade, e, conseqüentemente, introduzir as correções que se impuserem.

Explana JULIN²¹ que a pesquisa dêesses erros fortuitos não se pode efetuar segundo regras gerais, semelhantes às que a metodologia conseguiu traçar para o estudo dos erros constantes.

Mesmo assim, o engenho de grandes estatísticos já conseguiu formular alguns conselhos, de evidente utilidade, no sentido de não ficar tão arbitrária essa revisão das informações prestadas.

E GABAGLIO, na sua admirável *Teoria Generale della Statistica*²², considera como meios mais eficientes de se controlar uma informação estatística, os competentes confrontos:

a) com os dados da mesma espécie, provenientes de levantamentos anteriores e devidamente verificados;

b) com os dados da mesma espécie, oriundos de fontes diversas, mas autorizadas;

c) com o dado da mesma espécie, resultante de cuidadosa operação de cálculo, firmada em elementos de tôda a segurança;

d) com dados contemporâneos de outras espécies, porém, apresentando relações de afinidade.

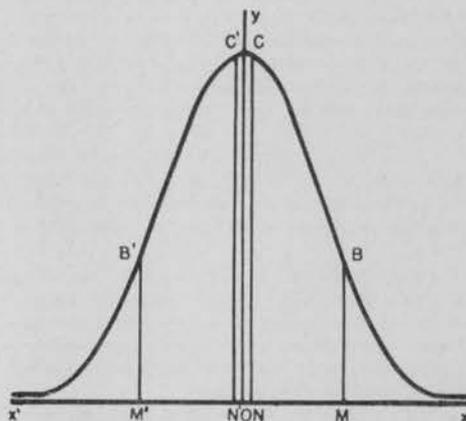
Se de tais cotejos ressaltarem pequenas diferenças, haverá motivos para atribuir credibilidade aos novos dados.

Grandes desproporções, entretanto, obrigarão a indagar se tais discordâncias provêm de êrro de observação, lacunas, duplicatas, má compreensão dos quesitos, os quais, depurados de quaisquer vícios grosseiros, deverão sofrer, ainda, uma análise por assim dizer externa, que incidirá sôbre as somas parciais e totais dos números e a boa impressão tipográfica dos mesmos, de modo a serem suprimidos os enganos, às vêzes, muito prejudiciais, de algarismos trocados nas parcelas, o que provoca contradições entre os dados individuais e a respectiva síntese final.

Para reagir contra o desânimo, que se possa apoderar de quem demasiado se impressione com tantas hipóteses ou ocasiões, involuntárias, de erros, ensina-nos o estatístico italiano MARIO MARSILI LIBELLI²³, que, sendo sempre possível, e relativamente fácil, cometer pequenos erros, será bastante difícil, embora ainda possível, perpetrar erros grosseiros. Em tórno do valor real, os erros se distribuirão indiferentemente, num ou nou-

tro sentido, para mais ou para menos. Essas deduções, a que se pode chegar com rigoroso raciocínio matemático, comprobatório da mais simples e atenta observação, constituem a lei dos erros accidentais, ou lei de GAUSS, ou curva de QUETELET.

Eis a representação gráfica dessa curva, conforme o mesmo autor:



Vê-se, aí, perfeitamente, que o êrro OM não muito grande, tem a pequena probabilidade MB, e que essa probabilidade é igual à que, em sentido negativo, M'B', decorre do êrro, em sentido contrário, OM'.

A figura torna, ainda, bastante claro, que ambas essas probabilidades são bem menores do que as indicadas em NC e N'C', correspondentes aos erros extremamente pequenos ON e ON', e que, para o êrro enorme OX, a probabilidade é mínima, embora sem chegar a zero, visto a curva jamais tocar o eixo X'X.

ELABORAÇÃO DOS DADOS

Meios de apuração

Reunidas as informações de que trata o formulário de determinada estatística, e investigada a respectiva veracidade mediante as revisões aconselhadas pelas críticas de sinceridade e de exatidão, urge apurá-las o mais rápido possível, a fim de serem divulgadas a tempo de prestar todos os serviços que justifiquem os esforços empregados e as despesas feitas com a sua realização.

A prática dessas apurações simplificou, extraordinariamente, os processos necessários para levar a térmo essa fase dos inquéritos.

As cartolinas, os mapas, e as máquinas de calcular, são os elementos principais de uma apuração efetuada com certo rigor técnico.

Cartolinas e máquinas de calcular

Refere JULIN²⁴ que, por ocasião do recenseamento italiano de 31 de dezembro de 1871, foi experimentado, pela primeira vez,

²¹ A. JULIN — Ob. cit., I, pág. 191.

²² A. GABAGLIO — Ob. cit., II, págs. 107 e seguintes.

²³ MARIO MARSILI LIBELLI — *Metodologia Statistica*, 1924, págs. 146 e seguintes.

²⁴ A. JULIN — Ob. cit., I, pág. 256.

o processo de transcrever numa ficha individual os dados constantes do boletim coletivo, daí se originando a cartolina, em seguida adotada pelos demais países. Esses retângulos de papel cartonado, impressos com traços e dizeres de maneira a receberem notações que permitam a contagem de cada elemento a ser apurado, foram, até certo tempo, preenchidos e elaborados manualmente, até que a fabricação de máquinas perfuradoras e somadoras, movidas a eletricidade, e de que são modelos universalmente conhecidos as máquinas "Hollerith", proporcionou extraordinária economia de tempo, de despesas e de cálculos, de par com imensas vantagens de simplicidade e de correção nos resultados parciais e gerais.

Outros aparelhos mecânicos, antes dos Hollerith, e procedentes da máquina inventada por BLAISE PASCAL, em 1642, haviam já facilitado, enormemente, a tarefa dos apuradores, graças à peculiaridade, que apresentavam, de realizar, automaticamente, não só as quatro operações fundamentais como todas as que decorrem das respectivas combinações.

As máquinas "Burroughs" são indispensáveis para qualquer apuração estatística de certo vulto. Somando inúmeras parcelas, concomitantemente com a sua impressão em fitas que podem ser conferidas com os números originais, tais maquinismos poupam a energia, o tempo, e a paciência, de quem os maneja, enquanto que os menos usados, porém, igualmente úteis, "Brunswick", "Triumphator", "Madas", "Millionaire", etc., facultam as multiplicações, divisões, e cálculos derivados, com os quais se conseguem todos os números absolutos e proporcionais.

Mapas de apuração

Mas, todos esses recursos mecânicos seriam malbaratados, se, para o seu inteligente aproveitamento, não pudessem contar os agentes apuradores com mapas preparados pelos lançadores das estatísticas de modo a poderem ser discriminados, coordenadamente, conforme os diferentes quesitos dos questionários, todos os dados coligidos nesses documentos.

A cartolina, também organizada para cada caso, é o instrumento mais apropriado para a separação e a contagem das unidades, efetuando-se estas por meio das máquinas acima referidas, ou se não fôr preciso, segundo o processo manual, considerado por JULIN preferível ao sugerido por BERTILON³⁵.

Este recomenda o apanhado das unidades, ou por meio de traços, a pequena distância uns dos outros (||||), ou por meio de números, em progressão aritmética (1 2 3 4 5 6 7, etc.), ao passo que JULIN, dizendo, com razão, ser incômodo esse último processo a partir da dezena, recomenda o sistema de realizar as contagens em folhas à parte

e sumárias, transportando-se, de quando em vez, para os mapas de apuração, os totais obtidos.

É evidente que as fitas das máquinas de somar substituem, admiravelmente, essas folhas avulsas.

Os mapas de apuração devem obedecer a critério uniforme, de maneira a conter as mesmas classificações, sob o ponto de vista da situação geográfica, do tempo (por dia, meses, anos, quinquênios, etc.), da idade, se se tratar de seres humanos ou irracionais, e, em geral, adotando-se nomenclaturas usuais ou preferidas pelos competentes no assunto sobre que versa o inquérito.

Reportamo-nos, aqui, às considerações expendidas no capítulo sobre a coleta dos fatos, relativamente às classificações necessárias para o melhor desdobramento dos aspectos dos fenômenos sociais suscetíveis de apreensão numérica.

Nem muito sucintas, nem muito longas, sempre que possível segundo a ordem analógica, sem repulsa absoluta da ordem alfabética, às vezes, insubstituível, tão científicas quanto fôr lícito ao saber a imparcialidade do observador, essas nomenclaturas deverão ser resolvidas após o exame aprofundado de todos os elementos abrangidos pelo inquérito, aliando-se, no seu preparo, o conhecimento de trabalhos semelhantes, contemporâneos ou anteriores, ao talento criador ou renovador do estatístico encarregado de sua orientação.

Raramente são produções de um só pensador, colaborando na respectiva feitura, em regra, teóricos e práticos, individualmente ou em caráter coletivo, as associações técnicas de estatística e as sociedades representativas das classes que se dedicam, por profissão, ao memento dos fatos a serem analisados estatisticamente.

Centralização das apurações

Efetuando-se as apurações na forma acima descrita, é claro que só um órgão central será capaz de desobrigar-se, satisfatoriamente, desses metuculosos trabalhos. Mesmo num país pequeno, de regime político unitário, sem unidades autônomas, quase independentes, a experiência tem comprovado que essas operações dão muito melhor resultado quando desempenhadas por aparelho central. Evidente, portanto, a conclusão de que nos demais países ainda mais acertada será a apuração centralizada.

Uma distinção, entretanto, convém fazer a esse respeito. Nos grandes Estados federativos, no Brasil, por exemplo, o conceito da centralização, em matéria estatística, deverá ser interpretado sem o absurdo de querer considerar, sempre, como órgão central, o pertencente à União. Verdadeira essa aceção no que diz respeito aos inquéritos mais gerais, isto é, aos censamentos, tornar-se-ia, manifestamente, contraproducente e disparatada, se relativa às estatísticas de muito menor importância. Deve-se, nesse caso, atribuir o caráter de central ao órgão da unidade federativa.

³⁵ A. JULIN — Ob. cit., I, pág. 255.

Só assim haverá possibilidade de se coligirem, em todo o País, as variadas informações de cunho estatístico, que permitam um panorama do conjunto dos fenômenos sociais acontecidos em vastos territórios, de população pouco densa e hábitos peculiares a cada região.

Ao ideal inatingível de completa centralização das apurações, cumprirá substituir o objetivo, de possível alcance, de tôdas as Unidades Federadas convergirem os seus esforços e recursos disponíveis para a consecução, em moldes bem simples e uniformes, do maior número de estatísticas demográficas, administrativas, econômicas e morais, cuja compilação, pelo órgão central, redunde numa síntese semelhante à que se acha contida nas publicações sobejamente conhecidas sob o nome de *Anuários Estatísticos*.

Números absolutos e relativos

Estudando as qualidades que devem apresentar os quesitos de um formulário, apontamos ser uma delas o respectivo preenchimento por números absolutos.

Só, excepcionalmente, portanto, se deverão apurar, na fase da elaboração, números que não sejam absolutos, isto é, relativos.

Costuma-se, por isso, denominar os números provenientes dos questionários, de primitivos, reservando-se a qualificação de derivados para os números resultantes dos primeiros, sendo que, se forem deduzidos por meio de aproximações ou comparações, serão chamados de relativos.

Os quadros de apuração, em regra, contêm, apenas, números absolutos, decorrentes das contagens dos números primitivos fornecidos pelos questionários ou boletins.

Estatísticas há, todavia, cujo interesse e utilidade ficariam em extremo diminuídos se a respectiva apuração fôsse feita apenas em números absolutos, motivando esse fato a conveniência da inclusão, nas tabelas de apuração, de colunas a serem preenchidas com números relativos.

Proporções por cento ou por mil, e coeficientes resultantes do confronto entre os dados obtidos e os representativos da população e da superfície territorial, eis números relativos cuja presença, nas aludidas tabelas, seria quase impossível de afastar-se, sob pena de ficarem inaproveitados valiosos esclarecimentos, cuja apuração conviria realizar-se desde logo. E o que dizer das médias, cuja inscrição, em seguida às linhas dos totais, se impõe, não raro, irresistivelmente, para imediata apreciação do movimento dos fenômenos estudados?

Médias

Examinemos, pois, em que consistem, e a forma por que se organizam esses números relativos, cuja importância cresce de vulto ao empreender-se a exposição e a interpretação dos dados devidamente coligidos e apurados.

F. VIRGILII³⁶ assinala ser-nos tão familiar o uso das médias, que as empregamos quase a êsmo, e, lembrando o seu uso pelos antigos, atribui a êstes a distinção das médias em aritméticas, geométricas e harmônicas.

FAURE³⁷ e COLAJANNI³⁸ observam haver quase tantas definições das médias, quantas da Estatística. O segundo apresenta umas cinco, ao passo que FAURE prefere defini-las pelo conceito mais cabível, conforme êle próprio reconhece, à média aritmética: "A média é a soma de muitas quantidades da mesma natureza, dividida pelo número das mesmas quantidades."

Segundo BLOCK³⁹, a média é a síntese de um grande número de fatos, — ou de casos individuais, unidades, — cujas diferenças ou afastamentos ficam, assim, compensados; é o ponto relativamente fixo, o centro em torno do qual oscilam tais diferenças. É um número abstrato, sem ser fictício, visto representar cousas reais, embora omitindo as qualidades acidentais e só refletindo as essenciais ou permanentes.

Com o progresso da Matemática, foram estudados vários teoremas relativos às diversas propriedades que possuem as médias, e entre elas se destaca o de ser a média o valor mais provável entre todos os valores de uma série de dados que divirjam entre si por pequenas diferenças, ou, por outros termos, o erro provável da média de um grande número de dados ser menor do que cada um dos dados em particular.

BERTILLON⁴⁰ dá um exemplo frisante dessa propriedade das médias, citando o caso das 487 medições da ascensão reta da Estrela Polar, realizadas pelo Observatório de Greenwich, durante os quatro anos de 1836 e 1839. Extraída a média dessas 487 medições, e feita a comparação do número médio obtido com os 487 absolutos, verificou-se que 82 dêses números absolutos eram por assim dizer exatos, pois só diferiam do número médio de cerca de um quarto de segundo, para mais ou para menos, ao passo que 72 observações deram mais 0,5 grau e 73 menos 0,5 grau; 63 mais 1 grau e 61 menos 1 grau; 38 mais 1,5 graus e 36 menos 1,5 graus; 16 mais 2 graus e 21 menos 2 graus; 5 mais 2,5 graus e 12 menos 2,5 graus; 1 mais 3 graus e 6 menos 3 graus; e 1 menos 3,5 graus.

A essa disposição dos erros em torno da média, acrescenta BERTILLON, é perfeitamente comparável a repartição das bolas negras e brancas, tiradas de um saco que contenha um número elevado de bolas, na proporção de uma bola preta para outra bola branca. Feita em larga escala a retirada das bolas, encontrar-se-á cada côr na proporção de 500 por mil. Mas, contadas as côres em séries, por exemplo, de 9 bolas, notar-se-á que, dentre as várias proporções resultantes do acaso,

³⁶ F. VIRGILII — Ob. cit., pág. 56.

³⁷ FERNAND FAURE — Ob. cit., pág. 42.

³⁸ N. COLAJANNI — Ob. cit., pág. 67.

³⁹ M. BLOCK — Ob. cit., pág. 122.

⁴⁰ J. BERTILLON — Ob. cit., págs. 109 e

serão mais freqüentes as que se aproximarem da referida proporção de 500 por mil, dando-se isto mais acentuadamente à medida que for aumentando o número das bolas colhidas no sacco.

No caso das medidas da ascensão reta da Estrela Polar, foram tomadas as dimensões de uma só grandeza. Trata-se, portanto, de uma média semelhante à que os estatísticos, em geral, denominam de objetiva e os ingleses de "mean", e de que GABAGLIO⁴¹ dá como exemplo a média das medidas repetidas da altura de uma casa, ou da estatura de um homem.

A essa média contrapõem os mesmos autores a que é deduzida de várias grandezas homogêneas mas distintas ou dissemelhantes, chamando-a de subjetiva, por ser abstrata e compendiar numa só as várias impressões produzidas no espírito humano pelas diferenças individuais dos objetos medidos.

São exemplos de GABAGLIO, como sendo dessas médias, as extraídas das medições das alturas das casas de uma vila ou das estaturas de um certo número de pessoas.

Ensina-nos BENINI⁴² que, de dois ou mais números, podem resultar infinitos valores médios, demonstrando-se essa possibilidade pela fórmula:

$$\sqrt[m]{\frac{a_1^m + a_2^m + a_3^m \dots + a_n^m}{n}}$$

em que $a_1, a_2, a_3, \dots, a_n$, são os termos de uma série, n o número desses termos, e m os valores que se queiram atribuir aos termos e à raiz.

Exemplificando o mesmo autor cita a média

$$\sqrt{\frac{a_1^2 + a_2^2 + a_3^2 \dots + a_n^2}{n}}$$

muito usada no cálculo dos erros de observação, e na qual m foi igualada a 2.

Acrescenta, todavia, BENINI, e com êle, todos os tratadistas, que, na prática, são apenas três as médias mais empregadas, isto é, a aritmética, a geométrica e a harmônica, as quais JULIN⁴³ coloca na seguinte ordem de importância e utilidade: 1.º, a aritmética; 2.º, a geométrica; 3.º, a harmônica. É mister acentuar que, das três, a primeira deixa a perder de vista as duas outras, cuja aplicação só tem sido feita em estudos de grande rigor técnico e relativos a casos muito especiais.

Média aritmética

A média aritmética decorre da soma dos termos dividida pelo número destes, e exprime-se pela fórmula:

$$\frac{a_1 + a_2 + a_3 \dots + a_n}{n}$$

ou, mais simplesmente, considerando apenas dois termos,

$$\frac{a + b}{2}$$

Média geométrica

A média geométrica é a raiz do grau n do produto dos n termos, ou seja,

$$\sqrt[n]{a_1 \times a_2 \times a_3 \dots \times a_n}$$

que se pode reduzir, também em relação a 2 termos, à fórmula

$$\sqrt{ab}$$

assim traduzida por VIRGLII⁴⁴: "raiz (de grau igual ao número de observações) do produto dos resultados das mesmas observações".

Média harmônica

A média harmônica é o valor recíproco da média aritmética das recíprocas. Essa definição não se compreende tão facilmente como as duas anteriores, e necessita de certo esclarecimento, que nos é dado, entre outros, por BENINI e JULIN, ao resumirem o trabalho de MESSEDAGLIA, "Il calcolo dei valori medi", publicado no 4.º fascículo do 5.º ano (1880), da revista *Archivio di Statistica*.

Para essa média, lança-se mão do valor recíproco, isto é, do valor resultante da divisão da unidade pelo termo.

A recíproca de a é $\frac{1}{a}$, donde a média

aritmética das recíprocas dos termos $a_1, a_2, a_3, \dots, a_n$ deverá ser

$$\frac{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} \dots + \frac{1}{a_n}}{n}$$

Portanto, a recíproca dessa média será

$$\frac{1}{\frac{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} \dots + \frac{1}{a_n}}{n}}$$

ou, multiplicando-se o numerador e o denominador por n ,

$$\frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} \dots + \frac{1}{a_n}}$$

Destas fórmulas, não será difícil chegar-se à comumente citada como represen-

⁴¹ A. GABAGLIO — Ob. cit., II, pág. 203.

⁴² R. BENINI — Ob. cit., pág. 92.

⁴³ A. JULIN — Ob. cit., I, pág. 361.

⁴⁴ F. VIRGLII — Ob. cit., pág. 56.

tativa da média harmônica: $\frac{2ab}{a+b}$, a qual é obtida, tomando-se em apreço dois termos, a e b , cuja média harmônica, então, será

$$\frac{2}{\frac{1}{a} + \frac{1}{b}}$$

ou, por outra,

$$\frac{2}{\frac{a+b}{ab}}$$

donde resulta a fórmula

$$\frac{2ab}{a+b}$$

Pela visível complexidade da média harmônica, é fácil de avaliar o dédalo de cálculos a percorrer para o estabelecimento de outras médias referidas pelos autores, como sejam as médias anti-harmônicas e quadráticas, cujo conhecimento será, todavia, acessível a quem quiser adquiri-lo pela sempre altamente instrutiva leitura do tratado de JULIN.

Propriedade das médias

A tódas as médias cabem as seguintes propriedades, mencionadas por BOSCO⁴⁵.

a) a média biparte a série considerada, quer dizer, cai entre os extremos da própria série, embora a maior ou menor distância de um ou de outro extremo;

b) se tódas as quantidades, de que foi deduzida a média, forem iguais entre si, a média também será igual às quantidades consideradas;

c) a média conservar-se-á a mesma, qualquer que seja a ordem em que estiverem alinhadas as quantidades de que foi extraída.

A média aritmética, por seu turno, tem propriedades especiais, que lhe são intrínsecas, e se condensam no seguinte:

a) a soma dos afastamentos positivos, entre os termos e a média, é igual à soma dos afastamentos negativos entre os mesmos termos e a média. Ou, mais sinteticamente, a soma algébrica de todos os afastamentos é igual a zero;

b) a soma dos quadrados dos afastamentos existentes entre os termos e a média é um "minimum" em relação à soma dos quadrados dos afastamentos verificados entre os termos e outro qualquer valor diferente da média aritmética.

É o que se evidencia no seguinte exemplo escolhido por BENINI:

| Têrmos | Afastamentos positivos (+) | Quadrados dos afastamentos |
|----------|----------------------------|----------------------------|
| | Afastamentos negativos (-) | |
| 25 | - 4 | 16 |
| 32 | + 3 | 9 |
| 28 | - 1 | 1 |
| 26 | - 3 | 9 |
| 19 | - 10 | 100 |
| 33 | + 4 | 16 |
| 40 | + 11 | 121 |
| Média 29 | Soma 0 | Soma 272 |

As qualidades privativas da média aritmética forçaram o severo escritor que é BENINI a confessar que, sendo ainda essa média de cálculo muito fácil e cômodo, tudo isso concorre para admitir-se o seu emprêgo mesmo nos casos de observações discordantes ou em série dinâmica, o que vale dizer ser a média aritmética, praticamente, a única aplicada em todos os ramos da Estatística.

Média ponderada

Um aperfeiçoamento dessa média veio, ainda, torná-la mais significativa e aproximada. Observando-se que, em certas estatísticas, sobretudo nas concernentes às questões sociais e econômicas, os números coletados e apurados eram de importância assaz variável, conforme se reportavam a êsse ou àquele fenômeno, impôs-se a necessidade ou, ao menos, a conveniência de, na extração da média aritmética, levar em conta essa importância, à qual se deu a denominação técnica de pêso.

Passou-se, então, a calcular a média aritmética da seguinte forma, conhecida sob o nome de média composta ou ponderada: Os termos, cuja média se pretende extrair, serão multiplicados pelos respectivos pesos, isto é, pelas quantidades a que se relacionam, dividindo-se a soma desses produtos pela soma dos pesos considerados.

Como exemplo, apresenta JULIN⁴⁶ o cálculo do valor médio do hectolitro de cerveja produzida numa fábrica.

Considerando-se a produção anual de X hectolitros, dos quais as quantidades X_1 , X_2 e X_3 foram vendidas, respectivamente, aos preços de V_1 , V_2 e V_3 , a média ponderada desses preços será:

$$\frac{X_1 \times V_1 + X_2 \times V_2 + X_3 \times V_3}{X_1 + X_2 + X_3}$$

Para a organização dos "index-numbers", ou números-índices, foi, em alguns países, adotada essa média ponderada, cujos resultados, embora não se diferenciando, extraordinariamente, dos produzidos pela média aritmética simples, não deixam, todavia, de apresentar um grau superior de aproximação, justificando, perfeitamente, a sua preferência.

⁴⁵ A. BOSCO — Ob. cit., pág. 446.

⁴⁶ A. JULIN — Ob. cit., I, pág. 365.

Uso das médias

F. VIRGILII encarece, com razão, o cuidado e a prudência com que se devem empregar esses meios de síntese estatística.

É rigorosamente necessário, diz ele, que a média seja extraída:

a) de elementos homogêneos, existindo essa qualidade ao menos em relação ao assunto, ao tempo ou ao espaço;

b) de elementos de igual precisão, usando-se a média ponderada quando o peso ou valor dos elementos for muito diverso;

c) antes de termos individuais do que de valores médios, pois, quase sempre, a média das médias não corresponde à média dos termos;

d) de um grande número de observações, pois a Estatística se baseia no estudo das massas;

e) de termos cujos valores absolutos não sejam muito diversos entre si, quer esses termos formem uma série de caráter estático, isto é, apresentando uma certa constância nas suas manifestações, quer constituam uma série de caráter dinâmico, isto é, revelando tendência para um movimento contínuo, caso em que será oportuno, quando for possível, reunir os elementos em grupos parciais, que manifestem um aumento regular.

Dizendo que segue o conselho dos ilustres mestres MESSEDAGLIA e GABAGLIO, o mesmo autor assim discrimina as ocasiões em que se deve recorrer a cada uma das espécies de médias:

À média aritmética: a) nas séries de caráter estático, isto é, quando os elementos se mantêm quase constantes, variando entre limites restritos; b) nas séries de caráter dinâmico, cujas variações seguem, mais ou menos regularmente, as leis de uma progressão aritmética; c) nos fenômenos que variam irregularmente, para mais ou para menos, contanto que se distingam claramente as variações positivas das negativas e se tirem as médias respectivas.

À média geométrica: nas séries de caráter dinâmico, cujos elementos variam segundo uma progressão geométrica, e nos fenômenos direta ou inversamente proporcionais a outros que, em suas manifestações, seguem o desenvolvimento de uma progressão geométrica.

À média harmônica: nas séries de termos que sejam recíprocos dos termos correspondentes de uma série aritmética.

Medianas e dominantes (modos)

Assim como não foi abordado, neste trabalho, voluntariamente elementar, o estudo das outras médias, que não as três, acima citadas, não nos deteremos no exame das medianas e dominantes (ou modos), a que aludem, com certo interesse, JULIN e BOWLEY, e outros notáveis tratadistas.

Seja-nos lícito dizer, apenas, como o primeiro autor citado⁴⁷, que a mediana é o

térmo que ocupa a posição central numa série de termos, de caráter variável, e dispostos segundo a respectiva grandeza.

Mais fácil e rápida de calcular do que a própria média aritmética, a mediana obtém-se pela colocação dos termos segundo a sua ordem de grandeza, e destacando-se dentre eles, o que estiver no meio, se a série for de termos em quantidade ímpar, ou a média dos dois que ocuparem essa posição, se a série for de termos em número par.

O próprio JULIN, que se manifesta tão favorável à mediana, faz diversas restrições quanto à possibilidade e conveniência do seu emprêgo, que, no entender do mesmo autor, só se apresenta como inteiramente aconselhável nas estatísticas dos rendimentos e dos salários.

Ainda de mais rara utilização é a dominante, posta em circulação, em 1895, pelo Professor PEARSON. Também cognominada de valor normal, de moda, ou de norma, consiste a dominante no valor que, numa determinada série, se apresenta com a frequência mais elevada, e em torno do qual se agrupam, mais densamente, os outros valores.

A dedução da dominante só pode ser feita por processos bem complexos, estabelecidos pelo próprio Professor PEARSON e por BOWLEY, ressentindo-se tais proporções de uma inevitável arbitrariedade, que ainda diminui a relativa utilidade desse cálculo.

EXPOSIÇÃO DOS DADOS

Exercitando o estatístico a sua atividade segundo as normas aludidas nos capítulos anteriores, será de esperar que logre satisfatório êxito nos seus esforços empregados com boa técnica, obtendo, assim, copiosa coleta de fatos, cuja representação numérica, após bem orientada crítica, será submetida a conscienciosa apuração, que redundará em tabelas ou quadros, formados por números absolutos e relativos.

Esses quadros constituem a exposição estatística dos dados coligidos e apurados, e poderão ser completados por diagramas de variadas formas ou figuras, englobados sob o nome genérico de gráficos.

Tabelas ou quadros

Quanto à constituição das molduras "sui generis" que são as tabelas estatísticas, já fizemos, no capítulo da coleta, ligeira descrição do que seja um quadro estatístico, visto poder apresentar esse formato o questionário ou boletim empregado no levantamento.

Aliás, o simples manuseio de qualquer anuário ou de publicações que contêm dados estatísticos, proporcionará aos estudiosos imediata compreensão da natureza desses quadros, cujo traçado bastará ser adaptado para nele serem lançados os números resultante da apuração do inquérito.

De forma retangular, com linhas verticais e horizontais, não mais fechadas à esquerda e à direita, como se praticava antiga-

⁴⁷ A. JULIN — Ob. cit., I, pág. 405.

mente, sendo os traços grossos ou duplos, e finos, conforme tenham de separar as variações dos fenômenos, no tempo e no espaço, tais quadros podem apresentar as mais simples ou engenhosas combinações, cuja perfeição ou cabimento dependem, principalmente, da ciência, e mesmo, do senso artístico, de seu organizador.

Os títulos e subtítulos nos cabeçalhos, ou na margem esquerda, devendo exprimir com toda a exatidão o conteúdo das linhas abaixo ou em frente, exigem, para a sua escolha e redação, o máximo cuidado por parte do seu autor, que, para tanto, deverá possuir a qualidade de poder condensar em vocábulos absolutamente apropriados e inteligíveis os característicos dos números a que se reportam.

Recomendam os autores sejam evitados os quadros confrontantes, ocupando duas páginas da publicação, e, se indispensáveis para traduzir toda a gama de variações dos fenômenos, lembrem sejam impressos de maneira a haver perfeita correspondência entre as linhas horizontais de ambas as páginas.

É hábito de nossos estatísticos escrever com maiúsculas, integralmente, ou, ao menos, quanto à letra inicial, todos os vocábulos das principais divisões dos cabeçalhos, variando os tamanhos tipográficos conforme o grau de dependência ou de subdivisão, que apresentem.

Prática, outrossim, merecedora de imitação é manter, uniformemente, horizontais, os textos dos cabeçalhos, cuja leitura será destarte, muito mais rápida e agradável, do que a das palavras, ora verticais, ora horizontais, não raro assim impressas até em tipografias oficiais.

No abalizado entender de JULIN, êsses quadros deverão ser projetados, na sua maioria, por ocasião do estabelecimento dos quesitos do formulário de coleta dos fatos a serem estudados, a fim de saber-se, desde logo, qual o destino das respostas aos quesitos, e de limitar-se o número dêstes ao estrito necessário para o preenchimento dos quadros de apuração.

Todavia, realizadas as operações de coleta, crítica e apuração, ter-se-á um critério mais seguro para organizar os quadros definitivos da exposição de todo o acervo de elementos de que possa dispor o organizador, e, evidentemente, será, aí, o momento de planejar as ilustrações gráficas tendentes a realçar os aspectos mais sugestivos de todo o inquérito.

Ilustrações gráficas

Implicando conhecimentos e prática de desenho, pintura e geometria, a confecção dêsses gráficos incumbe a especialistas, afeitos à cartografia, cabendo ao estatístico proporcionar ao cartógrafo os quadros a serem convertidos em diagramas, cartogramas ou estereogramas, e, quando muito, colaborar na escolha da maneira mais interessante e apropriada de reproduzir, gráficamente, os resultados finais do inquérito.

Desde as suas primeiras publicações anuais, a Diretoria de Estatística Comercial tem intercalado nas tabelas do comércio exterior bem elaborados diagramas da importação e exportação do Brasil, não sendo menos notáveis os gráficos, de variadas espécies, com que vem, de há muito, enriquecendo os seus magníficos trabalhos a Diretoria-Geral de Estatística.

O exame de um gráfico não provoca nunca, no leitor, a sensação de aridez, que, às vezes, é produzida pelos melhores quadros, ao primeiro e superficial exame. Substituindo visões concretas às evocações abstratas dos algarismos, revelando, de chofre, o movimento dos fenômenos, com as curvas ou linhas de fácil apreensão, e cujas interdependências também se evidenciam de relance, tais desenhos ou pinturas, completados, discretamente, com os principais números que representam, têm, ainda, a seu favor, a qualidade de servir de elementos de controle da veracidade dos números apurados, e de guia, ou ponto de partida, para o descobrimento dos princípios ou leis que regem os fatos estudados.

Esse processo de exposição dos dados é de alta valia no que diz respeito às comparações, tão em uso na Geografia Econômica, sendo obrigatório o seu emprêgo em qualquer manual dessa ciência.

Dividem-se os gráficos em diagramas propriamente ditos, isto é, figuras geométricas que não sejam representativas de sólidos; estereogramas, ou sejam figuras geométricas reproduzindo sólidos; e cartogramas, isto é, cartas geográficas coloridas, ou ornadas de figuras, que também podem ser outros diagramas.

Diagramas e estereogramas

Os diagramas baseiam-se, em geral, no sistema cartesiano, pelo qual as coordenadas são duas linhas retas que se cruzam num ponto, em ângulo reto ou oblíquo, denominando-se abscissa a linha horizontal e a outra, ordenada. Se as linhas se cruzarem em ângulo reto, o diagrama será ortogonal; se, de outra maneira, será oblíquângulo.

Dentre os múltiplos feitios que se poderão dar aos diagramas, estão, ainda, consagrados, pela sua aplicação e serventia, os de pontos e linhas, e os de superfície, sendo êstes retangulares, triangulares, quadrados, circulares, etc.

Nessas figuras geométricas, o movimento dos termos é graduado na largura ou na altura, ou, se forem circulares, em setores proporcionais ao total dos números considerados. Por êsse motivo, os gráficos aludidos são mais apropriados para representar relações dinâmicas, cabendo aos cartogramas a expressão de relações estáticas.

Os diagramas mais usados são os ortogonais. Os de pontos são muito singelos para atrair, devidamente, a curiosidade do leitor. Já os de linhas e os de superfície, refletindo as variações no tempo e no espaço, numa multiplicidade de aspectos tão sugestivos quanto verdadeiros, graças às combinações

de coloridos nas figuras, que se podem encaixar umas nas outras, despertam, poderosamente, o interesse de quem os tem à vista, de tal forma que existem alguns álbuns estatísticos compostos, unicamente, de diagramas e cartogramas.

As dimensões das áreas de superfície, que se queiram comparar por meio de diagramas, devem ser calculadas em relação às respectivas raízes quadradas, se se tratar de superfícies, ou segundo as raízes cúbicas, se de volumes⁴⁸.

Os estereogramas, conforme referimos, representam figuras sólidas, isto é, com três dimensões, e permitindo, assim, a exposição combinada de três elementos. São unânimes os autores em considerar os estereogramas de rara aplicação, pela dificuldade de imaginá-los e realizá-los com a indispensável clareza, — requisito "sine qua non" de qualquer gráfico.

Cartogramas

Essa qualidade é encontrada nos cartogramas, sobretudo nos de superfície, hipótese em que, segundo GABAGLIO⁴⁹, eles não são mais do que cartas topográficas comuns, sobre as quais, mediante simples diagramas de superfície, são figuradas relações estatísticas na respectiva situação geográfica e conforme as diferenças de grandeza.

Mais populares, ainda, do que êsses cartogramas, são os de côres graduadas, isto é, constituídos por mapas geográficos, cujas regiões, alcançadas pelo quadro estatístico que se quer evocar, serão coloridas, com maior ou menor intensidade, de acordo com as convenções adotadas em relação às tintas a empregar.

Muito menos usados do que êsses cartogramas são os de pontos, faixas, e curvas de nível, citando-os os autores, apenas, de passagem, e não lhes atribuindo vantagens que os recomendem, particularmente.

INTERPRETAÇÃO DOS DADOS

Destino dos números estatísticos

Os esclarecimentos, sob a forma de números, obtidos por meio dos inquéritos estatísticos, por si só, já causam grande satisfação ao espírito humano, sempre ávido e curioso de informações acerca de tudo quanto seja suscetível de percepção e análise. Interpretando-os, todavia, à luz de outros conhecimentos já adquiridos pela observação, ou pelo estudo, será possível, e profundamente interessante e proveitoso, descobrir a sua razão de ser, as leis que os regem, as causas que os motivaram.

Dar vida aos algarismos, friamente alinhados nas colunas de um quadro estatístico, e, sem lhes atribuir significação diversa da que realmente têm, fazê-los, por assim dizer, falar tudo quanto representam em si próprio

ou em relação a outros fatos, expressos ou não por outros números, eis a verdadeira interpretação que se impõe, a bem do progresso do saber humano, e de todas as ciências que o constituem.

Quem deve interpretar os dados

Ressentindo-se, ainda, muitos países civilizados, da falta de boas estatísticas sobre as suas principais condições sociais, parece-nos que, aos órgãos oficialmente incumbidos de elaborar estatísticas, incumbirá, primordialmente, o encargo de efetuá-las, aproveitando todas as oportunidades e divulgando os respectivos resultados sem a preocupação de lhes dar, imediatamente, minuciosa e aprofundada interpretação, sempre suscetível de encampar explicações ou pontos de vista, que, logo após, se verificarem estarem inquinados de erros ou lacunas.

Mas, por outro lado, é mister assinalar o fato de, para uma apreciação sumária das principais conclusões que ressaltem dos números obtidos em cada inquérito, ninguém está mais habilitado do que o estatístico, sob cujas vistas hajam passado todas as fases desses trabalhos, comprovando essa asserção os comentários que têm acompanhado as exposições dos resultados dos numerosos inquéritos procedidos e publicados pela Diretoria-Geral de Estatística e Diretoria de Estatística Comercial, de nosso País.

Essa interpretação, entretanto, não será nunca privilégio dos próprios autores das estatísticas, cujo material poderá ser aproveitado, não raro magistralmente, pelos competentes e versados nos assuntos de que tratam as respectivas tabelas, e que nelas poderão destacar relações ou leis sugestivas e importantes, na mesma ocasião em que, por outros, sejam acoidadas de inexpressivas ou insuficientes.

Preceitos de interpretação

Da sagacidade do interpretador dependerá, principalmente, o êxito das análises e sínteses que empreender a respeito dos resultados parciais ou totais do inquérito; mas, além dos preceitos que mencionamos na parte relativa à crítica dos dados, deverão ser respeitados outros princípios de lógica estatística, enunciados por QUETELET e completados por MESSE DAGLIA, para a devida apreciação dos dados, e, no tocante à pesquisa das causas dos fenômenos numericamente representados, os processos de lógica indutiva, formuladas por STUART MILL.

Seria, aliás, absurdo, contestar a necessidade de serem tomadas as seguintes precauções, para o julgamento do valor e do alcance dos números contidos nas tabelas:

1.º) Não ter idéias preconcebidas acerca do resultado final: preceito de equanimidade estatística;

2.º) não eliminar, arbitrariamente, algum dado, isto é, não desprezar os números que contradigam o resultado esperado: preceito de imparcialidade;

⁴⁸ A. JULIN — Ob. cit., I, pág. 573.

⁴⁹ A. GABAGLIO — Ob. cit., II, pág. 435.

3.º) enumerar, completamente, as causas, e não atribuir a uma só delas o que é efeito do concurso de diversas: preceito de causalidade;

4.º) não comparar elementos que sejam incomparáveis (na acepção negativa de impossibilidade de comparar): preceito de comparabilidade;

5.º) tirar conclusões na medida dos fatos observados: preceito de positividade;

6.º) nas deduções numéricas, não aspirar a um grau de precisão maior do que o permitido pelos dados com que se opera: preceito de contentabilidade;

7.º) quanto aos resultados gerais, só confiar em observações numerosas, segundo as condições do caso e o grau de certeza a que se pode aspirar: lei dos grandes números.

A clareza meridiana destas regras parece dispensar a repetição, por outras palavras ou frases, dos termos em que se acham articulados tais preceitos, os quais, além disso, se completam e esclarecem mutuamente, não se podendo dizer qual seja o mais digno de acatamento, tão imperiosa se nos afigura a necessidade de respeitar, integralmente, êsses admiráveis cânones de interpretação lógica e estatística.

Lei dos grandes números

Para COLAJANNI⁵⁰, o último, concernente à lei dos grandes números, é fundamental, o que talvez indique considerá-lo o referido autor como o mais importante de todos. Todavia, pouco adiante, COLAJANNI faz restrições à generalidade da aplicação desse princípio, notando que, em muitos casos, especialmente nas estatísticas morais, os resultados serão mais verdadeiros quando derivados de pequenos números. Aconselha, porém, sempre que possível, cotejar os resultados assim obtidos com outros, procedentes de grandes números.

Assim denominada por POISSON⁵¹, a lei dos grandes números consiste no fato de, qualquer que seja a amplitude das oscilações, ou o grau dos afastamentos, estar-se tanto mais próximo da realidade quanto maior fôr o número das observações realizadas.

Embora haja sido batizada com êsse nome, pelo referido autor, essa lei já era conhecida de BERNOULLI, que dela deduziu o seu teorema, nos seguintes termos:

"Repetindo-se indefinidamente o número de provas, a probabilidade de um acontecimento aproxima-se da certeza; à medida que se multiplicam as provas, tem-se uma probabilidade, sempre crescente, de sorte que a relação entre o número de acontecimentos favoráveis e o dos contrários não se afastará da relação existente entre as respectivas probabilidades além de um dado limite."⁵²

Acrescenta o mesmo autor que a precisão (aproximação ou probabilidade) dos resultados cresce com o aumento do número de observações, porém em proporção um tanto

menor. Teoricamente, a precisão só cresce segundo a raiz quadrada do número das observações.

Judiciosamente, declara JULIN⁵³ que a demonstração desse teorema pertence ao Cálculo das Probabilidades, cabendo somente ao estatístico reportar-se aos tratados que se ocupam do assunto.

VIRGILII, analisando, com alguma minúcia, tanto a lei dos grandes números, como o teorema de BERNOULLI, reflete que esse teorema deriva um ensinamento de grande importância, qual seja o de adquirirmos a convicção e a confiança de que certos fatos se reproduzirão dentro de determinados limites, após a repetição dos mesmos fatos, embora desconhecendo as causas que os reproduzem.

"Suponhamos, continua êle, ter observado que em certo número A de indivíduos de idade m , A' atingiram a idade M . Numa segunda observação notamos que em B indivíduos da mesma idade m , B' atingiram a idade M . E, por meio de terceira observação, sabemos que C' dentre C indivíduos de idade m , atingiram a idade M , e assim sucessivamente.

Feitas as relações

$$\frac{A'}{A} \cdot \frac{B'}{B} \cdot \frac{C'}{C} \dots$$

e reconhecido que elas pouco diferem uma das outras, se concluirá que, para o futuro, dado um grupo X de indivíduos de idade m , a probabilidade de que X' indivíduos do mesmo grupo cheguem à idade M está numa relação compreendida entre os limites das obtidas precedentemente.

O teorema de BERNOULLI é, pois, aplicável, tanto ao jogo de dados como à mortalidade humana, tanto à extração duma esfera de urna como à determinação da lei da criminalidade de um povo."

Métodos de raciocínio para investigação das causas

Os quatro métodos de raciocínio indutivo, a saber, da concordância, da diferença, dos resíduos e das variações concomitantes, estabelecidos por STUART MILL, são apenas, na opinião de BOSCO⁵⁴, processos mais rigorosos e precisos do que os, comumente, utilizados em quaisquer investigações estatísticas, e, também, não diferem muito daqueles a que recorreremos, até inconscientemente, nas contingências práticas de nossa vida.

Método de concordância

Servindo-nos das próprias definições de STUART MILL, temos que o método de concordância será aplicável na seguinte hipótese:

"Se dois ou mais casos, em que o fenômeno se manifesta, acontecem em circunstâncias tôdas diferentes entre si, exceto uma, essa única circunstância comum, em que todos os casos são acordes, é a causa ou parte necessária da causa do mesmo fenômeno."

Como exemplo, cita-nos COLAJANNI⁵⁵ o de furtos, freqüentes em dois países, cuja re-

⁵⁰ N. COLAJANNI — Ob. cit., pág. 105.

⁵¹ M. BLOCK — Ob. cit., pág. 131.

⁵² Segundo F. VIRGILII — Ob. cit., pág. 51.

⁵³ A. JULIN — Ob. cit., I, pág. 640.

⁵⁴ A. BOSCO — Ob. cit., pág. 659.

⁵⁵ N. COLAJANNI — Ob. cit., pág. 122.

ligião, organização política, clima, raça, grau de cultura, etc., sejam diferentes, notando-se, apenas, como desgraça comum, a miséria. Será esta a causa da frequência dos furtos.

Métodos de diferença

A fórmula do método de diferença é a seguinte:

"Se os casos, em que o fenômeno se manifesta, se realizam em circunstâncias comuns a todos, exceto uma, essa única circunstância é a causa ou parte necessária da causa do fenômeno."

Este método procede de maneira inversa da do anterior, mas ambos recorrem à eliminação, neste último, das circunstâncias iguais, e no primeiro, das diferentes, de modo que, afinal, só restará uma causa plausível do fenômeno estudado.

O mesmo exemplo do furto é dado, pelo autor há pouco mencionado, para ilustrar esse segundo método. Aqui se consideram no mesmo nível de clima, religião, raça, desenvolvimento intelectual e constituição política, os dois países, num dos quais são frequentes os furtos, e no outro, quase raros. Numa circunstância, todavia, se desassemelham esses dois países, isto é, na diferença de riquezas, distribuídas em proporções diversas pelas respectivas classes.

Nesse fenômeno, residirá a causa da maior ou menor frequência de tais crimes nos dois países.

Método dos resíduos

STUART MILL deu ao método dos resíduos este enunciado:

"Excluindo-se de um fenômeno a parte que, em virtude de induções anteriores, sabemos ser o efeito de certos antecedentes conhecidos, o resíduo será o efeito de antecedentes, a determinar."

Acompanhando o autor, que nos proporcionou os dois exemplos anteriores, lembraremos o caso do suicídio, que é mais frequente nos tedescos do que nos eslavos, nas pessoas cultas do que nas incultas. Entre os tedescos, são mais comuns os suicídios de protestantes; logo, a diferença de religião é, aí, o resíduo causal.

Método das variações concomitantes

Caracteriza-se o método das variações concomitantes pelo seguinte:

"Se um fenômeno varia, constantemente, numa determinada direção e medida desde que um outro varie noutra determinada direção e medida, o segundo fenômeno será causa ou efeito do primeiro, ou, então, ambos se dirão dependentes duma causa comum, que os fará variar de maneira harmônica."

Se a variação é no mesmo sentido, chamar-se-á de direta; se, em sentido contrário, de indireta.

COLAJANNI dá os dois exemplos, a) aumento de instrução e maior número de suicídios: variação direta; b) nível mais alto de cultura e menor número de homicídio: variação indireta.

GABAGLIO, por esse mesmo processo, julga que, ocorrendo, concomitantemente com a elevação ou o decréscimo dos preços das subsistências, maior ou menor frequência de óbitos, se poderá concluir que o preço dos víveres influi sobre a mortalidade.

Os mesmos autores e VIRGILII referem que o próprio STUART MILL não depositava grande confiança nos resultados da aplicação desses métodos ao estudo dos fatos sociais, considerando-os de real valia somente nas pesquisas das ciências naturais.

Insurgem-se esses estatísticos contra o pessimismo do grande lógico inglês, e afirmam que, manejados com tôdas as cautelas e suficiente habilidade, hão de prestar, os quatro métodos, excelentes serviços, muito concorrendo para o bom emprêgo do quarto processo os diagramas de linhas e os cartogramas coloridos.

Classificações das causas

Não contrariando, antes, tendo sempre em mira respeitá-los com todo o escrúpulo, os princípios estatuidos por MESSEADAGLIA e QUETELET, e pondo em contribuição os métodos que acabam de ser mencionados, será possível ao atilado observador discriminar os antecedentes, isolados ou em grupos, dos fenômenos estudados, e, verificando a constante e incondicional dependência dos fenômenos aos antecedentes, postos em foco por esse exame, também lhe será lícito considerar estes últimos como a causa dos fatos coletados pelo processo estatístico.

Em GABAGLIO⁶⁰ se louvam, em geral, os autores, quanto à classificação das causas que, influenciando no livre arbítrio dos homens, dão aos fatos individuais e sociais uma determinada direção ou lhes traçam os limites.

Tais causas ou circunstâncias ou são particulares, isto é, têm a sua sede no próprio ser humano, ou são gerais, e se encontram fora dêle.

As primeiras ainda se podem subdividir em físicas ou espirituais, conforme dizem respeito à natureza corpórea ou espiritual do homem.

Causas peculiares ao homem, físicas e espirituais

Causas ou circunstâncias físicas são as propensões animais ou dos sentidos, o sexo, a idade, a constituição física, a sanidade e as enfermidades do corpo, o temperamento, etc.

As causas espirituais são as inclinações inatas da pessoa para o verdadeiro, o belo, o bom; as virtudes, os vícios, a ciência, a ignorância, o estado civil, as condições sociais, a profissão, etc.

Causas gerais, naturais e sociais

As causas gerais, isto é, dependentes da natureza externa e da sociedade humana, distinguem-se em naturais e sociais.

⁶⁰ A. GABAGLIO — Ob. cit., II, pág. 336.

As naturais podem classificar-se da seguinte maneira:

a) telúricas, posição geográfica do país, superfície e constituição geológica do solo, configuração e extensão do território; estado hidrográfico;

b) atmosféricas, temperatura, umidade, do ar, estado do céu, ventos, etc.;

c) telúrico-cósmicas, mudança das estações, dos dias e das noites; acontecimentos extraordinários (terremotos, inundações);

d) telúrico-agronômicas, abundância, ou escassez das colheitas, etc.;

e) higiênicas, estado sanitário da população dos animais e dos vegetais;

f) geográfico-sociais, distribuição territorial da população, qualidade dos locais habitados (cidades, vilas, aldeias, bairros, casas especiais).

As causas sociais se bipartem em sociais propriamente ditas e políticas.

As sociais, "stricto sensu", são:

a) econômicas, ramos de indústrias capitais, grande e pequena indústria, grande e pequena propriedade imóvel, meios de intercâmbio, de transporte e comunicação, instituições de crédito, relações entre o trabalho e o capital; quantidade e qualidade dos consumos, crises econômicas, etc.;

b) intelectuais, grau de cultura pública, científica, literária e artística; instituições concernentes a essas culturas, etc.;

c) morais, religião dominante, espírito religioso, instituições de beneficência, moralidade pública.

As causas políticas são o mecanismo administrativo, a polícia, as finanças, a justiça, o exército, a constituição política, a política externa, as perturbações políticas, etc.

Causas constantes, variáveis e acidentais

Conforme atuem de forma contínua e sempre no mesmo sentido; também continuamente, porém com intensidade e direção mutáveis; ou, ao contrário, de maneira imprevisível, ora num sentido ora noutro, sem se poder determinar os efeitos e a intensidade, serão as causas: a) constantes, b) variáveis, e c) acidentais.

Nas causas sociais, diz COLAJANNI⁶⁷, não se conhecem ainda exemplos de constantes, mas são numerosos os de variáveis, e não poucos os de acidentais ou fortuitas (revoluções, guerras, descobertas científicas, etc.).

Nas outras causas, há das três espécies, sendo, por exemplo, constantes, o sexo, o solo e o clima; variáveis, a idade e as estações; e acidentais, os terremotos e as doenças.

Nas causas variáveis, observa-se que podem influir segundo leis determinadas, ou, ao contrário, sem lei alguma aparente, e escapando, assim, a qualquer previsão.

Para caracterizar as primeiras, os autores acrescentaram ao termo "variáveis" o vocábulo "regulares", sendo, por exemplo, causas variáveis regulares, as épocas (dias, meses, ou anos), as estações e a idade.

Para as segundas, foi atribuído o adjetivo de "irregulares", sendo causas variáveis irregulares, ou perturbadoras, entre muitas outras, os anos de abundância, o de carestia, e todos os períodos calamitosos, em geral.

Se as causas são representadas por um só fenômeno, dizem-se simples, tais como, a idade, o sexo, etc.; se constituídas por uma multiplicidade de condições, dizem-se complexas, como seja a profissão, que implica uma certa capacidade, um grau de desenvolvimento intelectual, a condição social, e outras circunstâncias.

É também de COLAJANNI, na página por último citada, a asserção de que na fenomenologia social se não poderá nunca falar de fatores simples.

A evidente significação dos adjetivos dispensa qualquer comentário acerca das distinções entre causas diretas e indiretas, mediatas e imediatas, determinadas e indeterminadas.

Causas e efeitos

Frisa o mesmo autor, que estamos acompanhando, a ação e reação recíprocas entre causas e efeitos, nos fenômenos sociais, cujas manifestações poderão ser produzidas pelos efeitos, oriundos de causas, e tornados também causas, por seu turno.

Como exemplo dessas complicações, cita o autor o caso do alcoolismo engendrar a miséria, e desta provocar o alcoolismo.

Uma instrução mediana necessitará, para ser adquirida, de certo grau de recursos, mas a instrução permitirá a obtenção e o aumento da riqueza.

A miséria constringe alguns a emigrar; outros, a roubar; uns tantos, ao suicídio, e muitos, à mendicância.

Entre outros motivos, é óbvio que essa complexidade torna qualquer interpretação dos dados tarefa não só espinhosa, como ingrata, pela raridade com que são recompensados, com satisfatório êxito, os trabalhos empreendidos a êsse respeito, a fim de descobrir as leis que regem os fenômenos sociais.

Mas as dificuldades estimulam os persistentes e batalhadores, já existindo diversas leis demográficas, econômicas e financeiras, a par de inúmeras observações quase dignas desse qualificativo, formuladas por espíritos superiores, cujo engenho, forrado de paciência, logrou a insigne consagração de descobertas felizes, desde logo agregadas ao precioso acervo dos conhecimentos realmente científicos, que formam a substância da Sociologia.

⁶⁷ N. COLAJANNI — Ob. cit., pág. 118.