

INSTITUTO DE HIGIENE DE SÃO PAULO
ESCOLA DE HIGIENE E SAÚDE PÚBLICA DO ESTADO
CAIXA POSTAL, 2920 — S. PAULO — BRASIL
DIRETOR: PROF. G. H. DE PAULA SOUZA

BOLETIM N. 69

EFICÁCIA DA APLICAÇÃO DO MÉTODO
ESTATÍSTICO AOS FENÔMENOS DE VIDA
REFERENTES AO HOMEM

P. EGYDIO DE CARVALHO
E
F. BORGES VIEIRA
1.ºº ASSISTENTES

1940
IMPRENSA OFICIAL DO ESTADO
SÃO PAULO

INTRODUÇÃO

O presente trabalho representa mais uma tentativa por parte dos autores em demonstrar, em nosso meio, o valor que pode assumir o método estatístico, quando aplicado aos fenômenos de vida referentes ao homem.

Embora a parte material tenha sido principalmente elaborada por um de nós, (Pedro Egydio), as idéias que vão ser expendidas vêm sendo propugnadas pelos autores há já longo tempo e sistematicamente postas em prática em suas publicações.

EFICÁCIA DA APLICAÇÃO DO MÉTODO ESTATÍSTICO AOS FENÔMENOS DE VIDA REFEREN- TES AO HOMEM

A ciência, em qualquer de seus domínios, tem sempre por finalidade estabelecer um mais íntimo contacto entre o Homem e a Natureza. Para realizar êste “desideratum”, o Homem começa por observar, cuidadosamente, na Natureza, tudo quanto tenha realidade ou existência, inclusive os fatos que lhe são interiores, para submetê-los, a seguir, ao exercício de suas faculdades intelectuais; estas últimas, para a aquisição de conhecimentos, atuam, imprimindo inicialmente às cousas, caracteres ou qualidades, que nada mais são do que as reações das impressões colhidas pelos sentidos; a tixação de tais caracteres permite, a seguir, o estabelecimento de noções de semelhança e de dissemelhança, bem como a formação lógica de grupos ou universos, cujos componentes, embora tendo de apresentar uma ou mais propriedades comuns, prefixadas e escolhidas conforme o fim colimado, podem diferir grandemente uns dos outros, no que concerne às qualidades restantes. Formados os diversos grupos, surge, desde logo, a necessidade de caracterizar a sua extensão; a reunião das unidades dá origem ao número, por intermédio do qual, é possível uma noção quantitativa do conjunto; aliás, a noção quantitativa pode também surgir, quando os diversos componentes de um mesmo grupo possuem um caráter comum, suscetível de mensuração.

Para o estudo dos grupos formados segundo os mais diversos pontos de vista e dos problemas peculiares a cada qual, surgiram as diversas ciências.

O estudo evolutivo de qualquer das ciências põe em evidência o fato de que, no início, todas fazem intervir apenas noções de qualidade e que, à medida que reconhecem a insuficiência dêste meio, na aquisição da verdade, passam a cogitações

de natureza quantitativa. Numerosos exemplos poderiam aqui ser invocados; lembremos, apenas, o citado por Pearl (1) e que se relaciona com a química; a alquimia já reconhecera numerosas propriedades materiais dos corpos (noção qualitativa), mas a química só pôde constituir-se em verdadeira ciência, quando os observadores fizeram entrar em conta as noções de pêso e de volume das substâncias (noção quantitativa), o que acarretou o estabelecimento de leis numéricas que governam êste ramo do saber (leis de Proust, de Dalton, de Gay-Lussac, etc.).

Se o fato é geral, seria de todo extranhável que o mesmo não se desse com a biologia; a exceção, de fato, não se verificou, pois, a tendência atual é, cada vez mais, a sua transformação em biometria.

O início desta transformação pode ser fixado com Graunt, que, em 1662, publicou uma obra notável, subordinada ao título: "Natural and political observations mentioned in a following index and made upon the bills of mortality chiefly with reference to the government, religion, trade, growth, air, diseases, etc., of the city of London"; muitas das conclusões a que chegou Graunt são verdadeiramente notáveis para o seu tempo. A seguir, em 1671, Christian Hüygens ocupava-se, pela primeira vez, com o cálculo da duração da vida. Também o célebre astrônomo e matemático Halley estudou, sob o ponto de vista teórico, os dados relativos à natalidade e à mortalidade, coligidos de 1686 a 1691, por Kaspar Neumann, em Breslau. Uma nova era abriu-se, entretanto, com a publicação, em 1741, do tratado: "Die göttliche Ordnung in der Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung desselben erwiesen", de autoria do reverendo Johann Peter Süssmilch; das idéias do autor, pode-se ter uma nítida noção, pela primeira parte de sua obra, onde êle demonstra que, nos nascimentos, nos óbitos, na reprodução do gênero humano, etc., existe sempre uma ordem constante, universal, grande, bela e perfeita, a qual governa todas as cousas; êle reclama, para a demonstração de suas idéias — contribuição valiosíssima — a necessidade de grande número de observações; o que diferencia, sobretudo, Süssmilch de seus predecessores é o seu profundo espírito de investigação; êle não se contenta com o sim-

(1) Pearl. R. "Medical Biometry and Statistics". 1923.

ples fato observado, indo mais longe, à procura de suas causas; não é pois, sem razão, que Rümelin afirma encontrarem-se, nessa obra, os fundamentos da biologia social e que Knapp considera o seu autor como o fundador da demografia. Justamente ao findar do século XVIII e nos primeiros anos do século XIX, Gauss estabelece a curva dos erros e Laplace constrói o monumento que até hoje perdura; o trabalho de Laplace, publicado em 1812, sob o título "Theorie analytique des probabilités", pode, sem exagero, ser considerado como uma das mais notáveis expressões do gênio humano. Leibnitz, por exemplo, considera a teoria das probabilidades como "o mais feliz suplemento à ignorância e à fraqueza do espírito humano". Estava assim preparado o terreno, que tão admiravelmente seria cultivado pelo grande cientista belga, Adolphe Quetelet (1796-1874); de seus inúmeros trabalhos, são particularmente importantes, do ponto de vista biológico, os dois seguintes: "Sur l'homme et le developpement de ses facultés ou essai de physique sociale", publicado em 1835 e "Anthropometrie ou mesure des diferentes facultés de l'homme", dado à publicidade em 1870. Outro grande vulto, nos domínios médicos, é o de William Farr (1807-1883). Entretanto, apesar dos esforços até então realizados, muito havia ainda por fazer; a Francis Galton (1822-1907), cabe a glória de ter levado a têrmo uma boa parte da tarefa, graças aos seus trabalhos relativos à teoria das variáveis; baseada na teoria da evolução de seu primo Darwin, a obra de Galton compreende, além de grande número de artigos, dois tratados de extraordinário valor: "Hereditary genius", datando de 1869 e "Natural inheritance", de 1889. Foi Galton, sem dúvida, o verdadeiro fundador da biometria; mas, se esta pode atingir a um tão elevado grau de progresso, devemos-lo á extraordinária figura do Prof. Karl Pearson, a quem Pearl se refere nos seguintes têrmos: "o grande mestre... cuja pura força intelectual raramente tem sido igualada em toda a história da ciência". (1)

Seria fastidioso citar a opinião de todos os autores, encarecendo a importância desta "quantificação" das ciências; é já clássico o provérbio de que, enquanto não há medida, não há realmente ciência. Um exemplo, tomado no campo médico, vai

(1) Esta parte histórica foi extraída, quasi totalmente, do livro de um dos autores: "Metodologia estatística", vol. I, 1936. P. Egydio de Carvalho e W. S. Leser.

ressaltar bem êste ponto de vista: imaginemo-nos antes da era termométrica; diante de um doente, o médico, ao apreciar um elemento diagnóstico tão importante quanto a temperatura, estaria apenas capacitado a pronunciar (e assim mesmo de forma imperfeita), se o enfermo estava quente ou não quente. Como seria, então, possível, apreciar a evolução da febre? Todas estas dúvidas, entretanto, vêm a desaparecer por completo, uma vez introduzido êste aparelho tão simples, que é o termômetro, por intermédio do qual é possível a obtenção de uma noção quantitativa do fenômeno.

Para estabelecer estas noções quantitativas, o Homem tem, entretanto, de vencer obstáculos derivados de certas limitações, tanto nas suas faculdades físicas, como psíquicas. As primeiras, que, naturalmente, são devidas às imperfeições dos nossos sentidos, tentam-se corrigir por meio de instrumentos adequados; assim, procura-se sanar a limitação da vista, em extensão e profundidade, por intermédio do telescópio e do microscópio. Aliás, por mais perfeitos que sejam os aparelhos de medida, êles não podem remediar, de forma absoluta, estas imperfeições; há sempre, para qualquer grandeza mensurável, um certo grau de aproximação, que a técnica procura tornar cada vez maior; por mais aperfeiçoados, porém, que sejam os seus recursos, ela jamais poderá alcançar o ideal máximo, ou seja, a determinação do real valor de uma grandeza. Quanto às limitações de ordem psíquica, são elas derivadas do fato de que “existe uma deficiência nas nossas faculdades, que malogra nosso poder de associação e nossa mente luta com uma dificuldade insuperável, ao querer sintetizar, em relações quantitativas, os fatos observados; o intelecto humano não sabe condensar os resultados das observações dos vários fatos, individuais; se os percebe individualmente, não sabe valorizá-los no seu conjunto”. (1)

Quantos fenômenos de vida não se encontram nesta situação? É facilmente apreciável, por exemplo, a proporção de uma determinada moléstia transmissível, em cada uma das famílias de nossos clientes; entretanto, a diversidade que se constata, ao passar de uma para outra, torna já muito difícil a apreciação da verdadeira proporção do fenômeno, na coletividade a que pertencem tais famílias.

Há, portanto, fenômenos e entre êstes se encontram inúmeros fenômenos de vida, em que a simples observação não

(1) C. Gini. — “Curso de Estatística”. 1935.

pode sanar as limitações de ordem psíquica, atrás apontadas, e que necessitam, para o seu estudo, de um grande número de observações particulares de outros fenômenos mais simples.

Mas, poder-se-ia perguntar, para a resolução de tais casos, não seria possível invocar, sempre e exclusivamente, o auxílio do clássico método experimental?

É fato bem conhecido, que o método experimental só pode ser aplicado, “quando possível regular as circunstâncias essenciais da observação, de tal forma que seja lícito torná-las fixas ou fazê-las variar a nosso bel prazer, em todo o sentido e numa quantidade tão pequena quanto se desejar (1). A continuidade e a reversibilidade da observação constituem o característico essencial do método experimental.” (2)

A importância desta variedade do método indutivo é enorme, nos fenômenos de vida; tenha-se em vista, por exemplo, os inestimáveis serviços prestados à biologia e à medicina, por um Claude Bernard, ou por um Pasteur, utilizando-se dêste processo de investigação.

Não obstante, na maior parte dos fenômenos de vida, ou não se conhecem todas as circunstâncias, ou, quando elas são conhecidas, não se pode fazê-las variar à vontade, seja por motivos inerentes às próprias circunstâncias, seja por seu número extraordinariamente elevado; adicione-se o fato de que, no Homem, não se pode operar como em “corpore vili” e convencer-se-á da impossibilidade de resolução de todos os casos relativos aos fenômenos de vida, pelo método experimental. Considerem-se, por exemplo, os fenômenos de herança ou os fenômenos demográficos (mortalidade, natalidade, nupcialidade, etc.).

Para solucionar, quantitativamente, “sobretudo” (3) êstes casos, em que não se é senhor das circunstâncias primordiais da observação, em que há interferência de uma multiplicidade de causas, em que as observações individuais apresentam resultados os mais diversos e em que, portanto, o fenômeno coletivo exige, para o seu estudo, grande número de observações particulares de outros fenômenos mais simples, é que se invoca o auxílio do método estatístico.

(1) Seria melhor dizer: ... numa quantidade tão pequena quanto a que for compatível com a natureza das circunstâncias, cujas influências procuram-se estudar.

(2) L. March — “Les principes de la méthode statistique” — 1930.

(3) Dissemos “sobretudo”, porque há grande número de casos, em que os próprios resultados experimentais não podem prescindir da análise estatística, tendendo-se, modernamente, a ampliar, cada vez mais, êste campo de ação.

Assim compreendida, a estatística pode ser definida como uma forma particular de indução, destinada ao estudo quantitativo da “maior parte” dos fenômenos coletivos, isto é, daqueles fenômenos que resultam da reunião de um grande número de fenômenos mais simples — fenômenos individuais ou particulares.

Introduzimos, no conceito de estatística, a expressão “maior parte” dos fenômenos coletivos, para significar apenas os atípicos, isto é, aqueles fenômenos, dos quais não se pode ter conhecimento exato, apenas por observação de um único caso, quando mais não seja, devido à multidão das manifestações individuais.

Entretanto, para os fenômenos coletivos típicos, como por exemplo, as partes principais do organismo de uma espécie animal, a segurança do resultado, pela observação de um único caso, ou, ao menos, de reduzido número de exemplares, torna supérflua a aplicabilidade da estatística.

A atipicidade dos fenômenos coletivos é, em grande parte, de natureza intrínseca e resultado de um grande número de causas, que operam sobre os mesmos, mas ela também pode ser extrínseca, devido à imperfeição dos nossos sentidos ou dos nossos instrumentos de medida; é o que se observa, por exemplo, quando se toma, com a maior precisão possível, o peso de um objeto.

Considerando-se agora, como faz notar Livi (1), “que nas ciências biológicas, a atipicidade dos fenômenos coletivos é a regra, não só quando se estudam os diversos órgãos nas suas funções ativas, mas também quando se examina a estrutura dos mesmos, que atípico ao máximo é o organismo humano, mesmo quando considerado sob o ponto de vista anatômico, que, quando se passa do homem para os fenômenos de sua vida natural (natalidade, desenvolvimento do organismo no tempo, enfermidade, etc.), ou social (nupcialidade, profissão, etc.), as oscilações de caso a caso são ainda mais profundas e graves”, convencer-nos-emos da procedência da tese a que este trabalho se propôs.

Um exemplo concreto: suponhamos que se pretenda estudar a ação de um determinado tratamento sobre uma certa afecção. Afim de que as possíveis conclusões não fiquem inteiramente sujeitas à ação do acaso, é claro que o investigador não se poderá contentar com um número reduzido de observações:

(1) Livio Livi — “Elementi di statistica”. 1937.

é evidente também, que, para a apreciação da eficácia ou não do processo empregado, êle deverá ter sempre em mente, um grande número de circunstâncias que poderiam interferir sôbre cada resultado, podendo-se citar, entre outras, como principais: a benignidade dos casos, o tempo decorrido entre o aparecimento da moléstia e o início do tratamento, a forma clínica da doença, o sexo e a idade dos pacientes, etc. Um outro exemplo ainda: como seria possível dissecar, com pequeno número de casos, o fenómeno da mortalidade infantil, quando, sôbre cada resultado, pode interferir um complexo de fatores, como: alimentação natural ou artificial da criança, o padrão de vida e nível intelectual da família, a saúde das pessoas que conviveram com a criança, o grau de civilização do país e muitos outros? Como vemos, em ambos os exemplos, os resultados se referem a um fenómeno coletivo, afetado por uma mutiplicidade de causas.

Demonstrada, assim, a necessidade do estudo da maioria dos fenómenos de vida, por intermédio do método estatístico, veja mos agora os elementos que dele emanam, para os fenómenos coletivos em geral, e, particularmente, para os fenómenos de vida referentes ao Homem.

Pode-se reconhecer, no método estatístico, duas funções primordiais, denominadas pelo professor Tschuprow, de ideográfica e de nomográfica.

Pela primeira, o método estatístico nos ensina como se pode descrever a extensão, a forma e a qualidade de um fenómeno coletivo; o exemplo mais típico, que se pode dar, neste particular, é o da população de uma localidade: é o método estatístico que nos pode fornecer, em termos precisos, qual o número de unidades que compõem essa população, como se distribuem essas pessoas por sexo, idade, profissão, etc. A estatística é também ideográfica, quando procura reduzir as grandes massas de resultados de observações, a poucos, mas eficazes valores, que se podem denominar de "características"; ela pretende, assim, apreender, na variedade infinita das manifestações individuais, os caracteres mais frequentes, as tendências mais comuns, averiguar o absoluto entre o relativo, o típico no acidental, o permanente no passageiro, o geral no particular. A importância dessa função é bem patenteada por Pearl, quando, ao estudar as aplicações do método estatístico à biologia, afirma que a estatística, "em última análise, é um método descritivo, não dos indivíduos ou casos singulares, mas do grupo, no seu conjunto, o que

realiza por meio de características constantes, ou seja, de valores que, no caso particular das pesquisas biológicas, seriam chamadas de constantes biométricas”.

A realização dêste fito é conseguida por intermédio das diversas médias (aritmética, geométrica, harmônica, etc.), desvio quadrático médio, diferença média, momentos de ordens diversas, números proporcionais, números índices, índices de homofilia, de correlação, de contingência, de atração, etc. Tais valores têm, cada qual, um escopo bem determinado, uns servindo para a fixação da intensidade do fenômeno, outros para o estabelecimento do grau de variabilidade dos valores individuais, outros para a determinação do grau de assimetria da distribuição dos valores, e, finalmente, outros ainda tendo intuito comparativo; a finalidade comparativa pode, aliás, ser atribuída de qualquer das características supra-apontadas, desde que se considerem duas distribuições de valores.

Fato digno de nota é que determinado aspecto de um conjunto pode ser fixado, quantitativamente, por formas diversas; o estudo, porém, das propriedades de cada uma das variedades de características destinadas a um mesmo fim e o alvo da própria pesquisa constituem elementos, para tornar cada vez mais objetiva a escolha das mesmas.

Estas características são de todo imprescindíveis para a economia do pensamento científico. Tomemos um exemplo que, embora grosseiro, fixa bem esta função sintética da estatística: o recenseamento paulista de 1934 acusou o total de unidades de 6.433.327; o statisticista, após o levantamento, teve de transferir as unidades coletadas para cartões adequados; em seguida, pela contagem, transformou o montão de cartões em resultados tabelares, e, finalmente, calculou as características principais. Se o conjunto das 6.433.327 pessoas pesava certamente mais de 200 milhões de quilos, e, o montão de cartões, não menos de 60 mil quilos, os números absolutos e relativos mais importantes podem, sem dúvida, ser colocados em uma ou duas páginas comuns, pesando assim, aproximadamente, 15 gramas.

Além da função ideográfica, a estatística, como já salientamos acima, é também invocada para se encarregar de algumas finalidades nomográficas. Estas últimas ganham, hodiernamente, cada vez maior importância, tendendo assim a ocupar o primeiro plano.

Antes de tudo, certos elementos nomográficos penetram na estatística, pelo fato de que as nossas observações não possuem sempre o caráter exauriente. Em algumas ocasiões, o caráter completo das observações é impossível de ser conseguido; seria absurdo o pretender-se determinar a fórmula leucocitária de um indivíduo, pela retirada de todo o seu sangue. Em outras circunstâncias, a realização do levantamento completo é perfeitamente razoável, mas deixa êle de ser levado a cabo, devido ao tempo, dinheiro e trabalho necessários ao mesmo. O conhecimento, por exemplo, do metabolismo basal normal dos paulistas, em idade adulta, deveria ser teoricamente obtido, considerando-se todos os indivíduos normais, que se encontrassem nas circunstâncias prefixadas; na prática, entretanto, tal tarefa seria inexequível.

Há casos, portanto, em que temos de julgar das características de uma certa coletividade, não pelo estudo da própria coletividade, mas por intermédio das características de amostras das mesmas. Quando o número de observações, que fazem parte da amostra, é suficientemente grande em relação ao total dos elementos de que se compõe a coletividade de ordem superior, demonstra-se matematicamente que, qualquer que seja a técnica empregada na escolha dos elementos da amostra, é sempre possível, partindo-se das características encontradas, obter limites, dentro dos quais, com grandes probabilidades, devem cair as características incógnitas da coletividade de ordem superior considerada. Se a circunstância que vimos de apontar não está preenchida, somos então obrigados a executar certas regras de escolha, pois, só desta maneira, será possível a obtenção de resultados satisfatórios; a execução de tais princípios pode então, fornecer os chamados levantamentos representativos (ou de provas ao acaso) de Kiaer, que se podem admitir como quasi equivalentes aos levantamentos completos.

A melhor maneira de se proceder, neste particular, é dividir a coletividade de ordem superior em um certo número de coletividades parciais homogêneas, realizar misturas adequadas nestas coletividades e, em seguida, tomar um certo número de elementos de cada uma delas. Como, entretanto, tal processo exigiria um conhecimento prévio do verdadeiro âmbito das diversas coletividades parciais, o que nem sempre é possível, na prática, o processo utilizado é a realização de uma mistura fundamental de todas as unidades de ordem superior, como se faz nas questões técnicas usuais; diga-se de passagem que, se por vezes, esta mistura essencial é realizada por si mesma, o julgamento, nestas circunstâncias, da representatividade ou não de uma amostra, mesmo nos casos em que se acham afastados,

aprioristicamente, os fatores evidenciáveis por si mesmos, constitue, certamente, um dos problemas mais delicados da estatística. Critérios vários, como por exemplo, o estudo da variação do comportamento do fenômeno em sub-amostras, adequadamente constituídas dentro da própria amostra a analisar, só podem dar indícios probabilísticos e jamais uma certeza incontestável; assim, consideremos o caso em que, pretendendo determinar a frequência de positividade da reação de Von Pirquet, entre os escolares paulistanos, examinássemos uma em cada cinco escolas da nossa Capital; mesmo que a variação na percentagem de positividade das amostras das diversas escolas examinadas se mantivesse dentro de limites previstos pelo que chamamos de flutuações de amostras simples, isto é, mesmo que as variações estivessem dentro dos limites que, em probabilidade, demonstram-se ser devidos ao “acaso”, não estaríamos autorizados, teoricamente, a generalizar a todos os escolares paulistanos, os resultados da amostra constituída pela reunião das diversas escolas efetivamente examinadas, pelo simples motivo de que poderia ter escapado uma escola, em que o fenômeno se apresentasse de forma aberrante; é claro, porém, que o fato constitue um ótimo elemento para que tal generalização fosse bastante provável.

Em muitas circunstâncias, porém, a estatística fornece meios de se transformar levantamentos não representativos, por causas sistemáticas, em representativos e, desta forma, preencher a finalidade já apontada. Este estudo, que é relativamente recente, já recebeu aplicações no campo biológico, como, por exemplo, a verificação estatística das leis de Mendel, para certos caracteres transmissíveis, como a epilepsia, por intermédio do chamado processo de Weinberg e Lundborg.

É esta função da estatística, de poder dar o meio de se julgar de uma coletividade de ordem superior, por intermédio de uma amostra da mesma, o traço fundamental que a liga ao cálculo de probabilidades.

A segunda função nomográfica, imputável à estatística, é a pesquisa de possíveis leis, tomando este vocábulo na sua mais ampla acepção, significando por tal, não apenas uma relação constante entre dois fenômenos de natureza diferente (relação entre a idade da mãe e o peso do recém-nascido), mas também qualquer regularidade na aparente irregularidade das manifestações individuais (distribuição dos indivíduos segundo um certo caráter antropológico).

Um exemplo vai ilustrar esta finalidade da estatística: suponhamos haver medido a altura de um certo número de paulistas adultos, de cor branca; a diversidade dos resultados das observações individuais faz com que possamos repetir com o filósofo: “tudo aparece, à primeira vista, espontâneo, livre e caprichoso como o vento que sopra; todavia, como o vento que sopra, tudo obedece a condições precisas e a leis fixas. Seria cousa útil o evidenciá-las” (1). Se, entretanto, calcularmos a média aritmética dos resultados (que, subjetivamente, significa o valor que as observações teriam, caso fossem iguais e que, objetivamente, fixa o tipo da raça) e dispusermos em um gráfico a frequência dos diversos desvios, contados a partir da mesma, começaremos a notar que a distribuição tende a tomar uma forma regular, simétrica, que lembra imediatamente a conhecida curva de Gauss-Laplace. Uma conveniente interpolação desta curva demonstra, efetivamente, a sua adaptabilidade à distribuição efetiva; e, desde que se conhece a expressão matemática da curva de Gauss-Laplace, somos levados a fixar que esta também é a lei que rege a distribuição das alturas, uma vez que o fato se repete nos resultados de outros conjuntos.

Muito mais complexo se torna o pensamento, quando se pretende determinar relações de causalidade. É esta uma das tarefas mais importantes da estatística e pode-se mesmo dizer que a pesquisa das causas domina todo o trabalho estatístico — levantamento dos dados, elaboração dos mesmos e pesquisa das causas. A complexidade a que nos aludimos se refere, em primeiro lugar, ao fato de que, não se sendo senhor das circunstâncias que governam o fenômeno e, portanto, não se podendo separar ou fazer variar à vontade as complexas influências suscetíveis de ação sobre o mesmo, tem-se, quasi sempre, de se contentar com as suas manifestações espontâneas, nas quais, entre as circunstâncias que o acompanham, figuram, com intensidades diversas, a suposta ou as supostas causas. Além disto, a causa de um fenômeno residindo, como faz notar Stuart Mill, no conjunto de seus antecedentes, é um ato, em grande parte arbitrário, o atribuir-se a um ou alguns de seus antecedentes um papel decisivo, fato especialmente notável em biologia. O critério fundamental, que se utiliza em estatística, para sanar, em parte, tal arbitrariedade, no indício de intervenção de uma causa, é a constatação do fato dos resultados esta-

(1) Taine — “Philosophie de l’art”. 1906.

tísticos se afastarem do comportamento que deveriam seguir, se nenhuma circunstância operasse sobre os mesmos, sendo sintomas mais importantes da intervenção de uma causa, a descontinuidade dos fenômenos estatísticos e a divergência entre o desenvolvimento efetivo do fenômeno e seu desenvolvimento teórico.

Um dos exemplos mais característicos, que se pode fornecer, para evidenciar a importância da descontinuidade do fenômeno, como indício de intervenção de uma causa, é o da febre tifóide. Estudando a evolução daquela moléstia na cidade de São Paulo, durante um longo período de tempo, um de nós (1) verificou, na série histórica formada, que, a partir de um determinado ano, a mortalidade por aquela enfermidade sofreu uma franca diminuição; buscando os motivos dessa descontinuidade, conseguiu aquele autor demonstrar que a mesma era causada pela cloração das águas de abastecimento.

Quanto à discrepância entre o desenvolvimento efetivo do fenômeno e seu desenvolvimento teórico, o exemplo clássico é o referente à distribuição das alturas, quando, no levantamento, se estuda um grupo misto, fusão de dois ou mais grupos, cujas constantes típicas são diversas; a curva efetiva apresentar-se-á, então, com dois ou mais máximos (curva bi ou pluri-modal), contrariamente ao que deveria suceder, de acordo com a curva teórica do fenômeno, que, como já se salientou, é a de Gauss-Laplace (curva unimodal).

Aliás, a causa determinante de um fenômeno é apenas uma concepção do espírito, uma noção limite, de que a realidade raramente se aproxima bastante, uma hipótese inteiramente indispensável à vida prática, para a qual, entretanto, nunca serão excessivos os controles e verificações.

As formas de se operar, no estabelecimento destas relações causais, são as mesmas empregadas na indução experimental e que Stuart Mill distingue em quatro variedades: o método das variações concomitantes ou correlativas, o método das diferenças, o método das concordâncias e o método dos resíduos; entre êles, o mais importante é, sem dúvida, o primeiro, havendo mesmo alguns autores que são de opinião de que este processo pode absorver todos os outros, razão pela qual, ocupar-nos-emos unicamente dele, embora de forma sumária.

(1) Borges Vieira, F. — “Epidemiologia da febre tifóide em São Paulo”, *Acción Medica*, 1934.

Consiste o processo das variações concomitantes no seguinte: ao se verificar que um fenômeno varia, constantemente, em uma direção e medida determinadas, quando outro fenômeno varia também, em uma certa direção e medida, pode-se concluir que, ou existe entre os dois uma relação como a de causa e efeito, ou ambos dependem de uma mesma causa. Para a aplicação deste processo, são imprescindíveis as noções de grandeza das variações, sua correspondência, grau de acôrdo ou desacôrdo, noções estas que são fixadas por características especiais. Tomando-se, por exemplo, a capacidade respiratória de crianças em diferentes idades, constata-se que, de uma forma geral, ela aumenta à medida que cresce a idade, o que pode ser fixado, quantitativamente, quer pelo coeficiente de correlação de Pearson, quer pelo índice de homofilia de Gini; conclue-se, então que as circunstâncias, que determinam aumento da capacidade vital, dependem diretamente da duração da vida da criança.

Cuidados especiais devem ser tomados na aplicação deste processo, as conclusões devendo ser formuladas, somente após exame de um grande número de casos e indagação da possível interferência de outros fatores, sôbre o fenômeno considerado.

A influência do número de casos, para a interpretação de um determinado resultado, afim de se julgar se o mesmo é ou não significativo ou seja, para se aquilatar a influência que o "acaso" poderia ter sôbre êle, é determinada, em estatística, por meio da chamada teoria das flutuações de amostras; constitui esta, sem dúvida, um dos mais importantes alicerces, em que se funda o estudo crítico dos resultados, evitando-se muitas conclusões erradas e colocando-se, no campo do "acaso", fatos que teriam, sem ela, sido atribuídos a causas de natureza definida. Supostos os resultados obtidos de uma maneira idêntica a um dos esquemas clássicos do cálculo de probabilidades, é possível determinar limites de variação dos mesmos, função do número de casos estudados e compatíveis com a simples ação do "acaso".

Quanto à eliminação da influência de fatores perturbadores, é ela, naturalmente executada, fazendo-se com que os mesmos permaneçam constantes. Dois são os processos principais, que se empregam com esta finalidade: o método dos grupos escolhidos e o processo da população tipo.

De acôrdo com o primeiro, divide-se o fenômeno que se quer estudar, em tantos grupos parciais, quantas são as modalidades do fenômeno que opera como causa perturbadora; assim, se no exemplo anterior, supõe-se que o sexo pode ter influência

sobre a relação verificada, pode-se, muito simplesmente, afastar este fator, estudando-se, separadamente, a relação entre idade e capacidade respiratória nos meninos e nas meninas.

O método da população tipo que, contrariamente ao que seu nome faz supor, tem um caráter geral, consiste no seguinte: estabelecida a ação de uma causa, em dois grupos, em que a mesma opera com intensidades diversas, se se objeta que o fato é devido a um fator estranho, o mesmo pode ser eliminado, verificando-se como a suposta causa operaria nos dois grupos, caso ambos tivessem, em relação ao fenômeno perturbador, uma composição idêntica à de um padrão predeterminado; se se verifica, por exemplo, que a natalidade é maior nas classes sociais menos elevadas e se se objeta que tal fato é devido a uma composição etária diferente nos dois grupos analisados, compreende-se que a resposta a tal dúvida pode ser facilmente formulada, estudando-se como se comportaria o fenômeno da natalidade, nos mesmos dois grupos, caso ambos tivessem uma composição por idades idêntica à de um padrão conhecido.

É claro que, toda a regularidade ou permanência de uma relação sugere a possibilidade de um laço entre os dois termos que dela fazem parte; a demonstração de sua existência real só é feita, em estatística, quando a análise das circunstâncias adjacentes estabelece a sua necessidade; a não observância destes princípios já fez com que, muitos autores incidissem em erros grosseiros de interpretação.

No exemplo relativo à relação entre capacidade vital e idade das crianças, a multiplicidade, a concordância das possíveis observações e os meios de que se pode lançar mão, afim de evitar a influência de circunstâncias perturbadoras, fazem com que a conclusão tenha grandes probabilidades de ser verdadeira. Esta experiência indireta, de comparação das variações das causas com as variações sofridas pelo fenômeno em estudo, é realizada sob variadas formas: ora através de longos períodos de tempo, ora por meio de lugares diversos, ora por intermédio de grupos distintos de elementos.

A grande extensão deste processo explica o porquê a estatística é um método prevalentemente comparativo.

É claro que os resultados estatísticos não comportam a precisão dos resultados experimentais, uma vez que, enquanto nestes, pode-se proceder a verificações com continuidade, naqueles a observação é esporádica. Esta descontinuidade das variações constitui um sério óbice, especialmente na interpretação de certos fatos submetidos a tendências gerais comuns, em uma certa direção, tendências estas que podem dar uma

ilusória uniformidade nas mudanças observadas. Deve-se notar, entretanto, que o raciocínio, no método estatístico, apresenta um espírito idêntico ao do método experimental, ambos obedecendo a idênticos princípios de lógica.

Fato importante a ser levado em consideração, para a interpretação das relações de causalidade, é que elas, como aliás acontece com qualquer conhecimento estatístico, não se referem a cada uma das unidades observadas, mas sempre a certas coletividades das mesmas; assim, quando se estabelece que o coeficiente de natalidade de uma determinada região alcança a cifra de 27 por mil habitantes, isto não significa, absolutamente, que um homem tenha a chance, ainda que mínima, de dar à luz uma criança.

Se a estatística descobre leis, no amplo sentido a que já nos referimos, não é, de forma alguma, de sua alçada, determinar o mecanismo atribuível à produção das mesmas; assim, o porquê da constância da proporção de homens e mulheres, ao nascer, deve ser decidido por médicos e pelos estudiosos dos problemas de herança. Esta afirmativa não significa, de forma alguma, a não possibilidade de existência de um íntimo contacto entre o estatístico e o especialista que se aproveita dos resultados estatísticos. A cooperação mútua, muito ao contrário, só pode trazer vantagens recíprocas, devendo-se mesmo considerá-la como imprescindível; nunca é demais insistir sobre tal fato, visto como, na prática, o mesmo não se realiza com a frequência que seria de desejar-se.

Finalmente, pode-se reconhecer na estatística, uma função de previsão. Esta passagem do conhecido para o desconhecido, baseando-se nos resultados da experiência, ou na observação através do tempo, embora inteiramente indispensável à vida prática, constituindo mesmo o alvo final de todas as ciências, é de natureza incerta, os seus resultados sendo sujeitos a modificações, de acôrdo com a melhoria dos nossos conhecimentos.

A obtenção de um resultado de previsão é, fundamentalmente, baseada na hipótese de uma certa permanência nas condições em que o fenômeno se produz; verificada que seja uma certa regularidade, admite-se que as influências, suscetíveis de atuar de modo sensível sobre o mesmo (causas principais), tenham operado de uma forma constante, enquanto que as circunstâncias que, por ignorância, são rubricadas genericamente de "acaso" (causas secundárias), cuja ação é imperceptível, tenham

operado com irregularidades, ora num sentido, ora noutro, de forma altamente complexa, a ação das mesmas se neutralizando, praticamente, quando consideradas no seu conjunto; admitida, a seguir, a hipótese de que tal estado de cousas se mantenha no futuro, torna-se possível, então, deduzir da observação no passado o que deverá acontecer no futuro.

O estabelecimento da lei, que rege os resultados observados efetivamente, é feito, quasi sempre, pelos chamados processos de interpolação, cujas bases residem no cálculo de probalidades; as previsões possuem unicamente o caráter hipotético das extrapolações, e os resultados só podem ser expressos como prováveis estimativas.

É evidente que os resultados de previsão são tanto melhores quanto mais homogêneos e estáveis forem os grupos a que se referem, pois, desta maneira, as hipóteses imprecindíveis à sua determinação se coadunam mais com a realidade; esta especialização possui, não obstante, o inconveniente de restringir o número de elementos de cada grupo, de forma que a mesma pode ser levada a cabo, só até certo ponto.

Da importância desta finalidade da estatística nos fenômenos de vida, falam bem alto os métodos de estimativa de população, base de todos os trabalhos demográficos, bem como as tábuas de mortalidade, que muitos autores consideram como uma das pedras angulares da estatística.

Para terminar estas nossas considerações sobre a eficácia da aplicação do método estatístico aos fenômenos de vida referentes ao homem, queremos deixar aqui assinaladas as opiniões de dois autores: a primeira é a de Anderson (1), que assim se pronuncia: “aquele ideal orgulhoso do conhecimento humano, que foi pela primeira vez formulado, há cerca de 120 anos, por Laplace, no seu “*Essai philosophique sur les probabilités*”, (1814): une intelligence qui pour un instant donné connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome; rien ne serait incertain pour elle, et l'avenir comme le passé serait present a ses yeux, etc., e que Du Bois Reymond denominou de ideal astronômico, parece esmaecer cada vez mais e ser substituído por um ideal estatístico”.

(1) Anderson N. O. — “Einführung in die mathematische Statistik”. 1935.

A segunda opinião a que queremos nos aludir é a de um médico paulista, W. Leser, (1) segundo o qual: “os métodos estatísticos não são apenas aplicáveis, na medicina, mas sim, absolutamente indispensáveis; aí, onde pouco, ou mesmo nada, existe de absoluto, é quasi impossível a obtenção, com segurança, de qualquer conclusão, a não ser mediante a análise estatística; esta asserção receberá, da parte de muitos, um formal desmentido: — nunca conhecemos tais métodos e temos chegado a conclusões, sem deles fazer uso —; na verdade, houve aplicação, embora inconciente. É justamente, nesse fato, que se acha a falha e, ainda mais, o perigo; falha, porque nem todos os recursos terão sido empregados, acarretando a obtenção de resultados incompletos; perigo, porque não poucas conclusões falsas serão admitidas como verdadeiras. Seria infeliz a comparação de tais casos, aos dos curandeiros que, inconcientemente, fazem uso, algumas vezes, de processos terapêuticos verdadeiramente científicos?”

Subscrevemos, “in totum”, ambas as opiniões e não estranhemos mesmo o certo grau de cepticismo que ainda perdura para com este instrumento de pesquisa tão delicado, que é a estatística, porque não desconhecemos as dificuldades peculiares à aceitação de noções novas.

Não estaria, por ventura, se dando com a estatística, o que sucedeu com o termômetro, que na época do seu aparecimento, foi qualificado, por parte de dois illustres clínicos de então, Corvisart e Ledouble, como “inovação anti-clínica e anti-biológica”?

(1) Leser, W.S.P. — “Contribuição para o estudo dos métodos estatísticos aplicáveis à medicina e à higiene”. Tese de doutoramento. 1933.