Normalidade estatística

Leonidas Hegenberg

HEGENBERG, L. *Doença*: um estudo filosófico [online]. Rio de Janeiro: Editora FIOCRUZ, 1998. 137 p. ISBN: 85-85676-44-2. Available from SciELO Books <http://books.scielo.org>.

8

NORMALIDADE ESTATÍSTICA

(Ad medicorum usum)

Resumo. Na linguagem clínica, as doenças, ou condições patológicas, são consideradas "anormais", ao passo que as condições saudáveis são vistas como "normais". A par disso, é comum falar em *valores normais* de variáveis clínicas (e.g., peso, altura, pressão sangüínea, etc.). Subjacente, há uma intuição fundamental: *a pessoa comum é normal* – sem o que, dificilmente se deixaria claro como entender, nesse contexto, a palavra 'normal'. Este "adendo" destina-se a deixar explícitos os significados de alguns termos da Estatística – básicos para a adequada consideração do *normal* em Medicina. [O material aqui reunido é bem conhecido. Novo, talvez, para leitores brasileiros, apenas o que se fixa na primeira seção.]

Norma e normal

A palavra 'norma' é empregada em diversos contextos. Remete, por exemplo, a "meios que conduzem a um fim". E' o caso, digamos, das normas técnicas. Sob as vestes de instruções, devem ser adotadas por alguém que pretenda alcançar um dado objetivo. Tomam a forma de diretrizes. São sentenças condicionais em que o antecedente menciona o que deve (ou não deve) ser feito. Exemplificando, "Se o senhor deseja eliminar esse mal-estar, deve ...". As bulas, nos medicamentos, dão exemplos de instruções, ou normas, neste particular sentido.

'Norma' remete, ainda, a "regras". Os jogos (atividade lúdica) são realizados de acordo com certas regras. A rigor, as regras *determinam* o jogo, fixando lances lícitos (permitidos, ou corretos), bem como lances ilícitos, (proibidos, ou incorretos), entendendo-se que, em dadas condições, se há apenas um lance permitido, ele será obrigatório. Em Medicina, a cirurgia se compara ao jogo. Há "lances" permitidos e proibidos e até lances obrigatórios. O exame clínico também se processa de acordo com regras e não será difícil imaginar lances corretos e incorretos.

'Norma' lembra, enfim, regulamentos, ou prescrições. Regulamentos são fixados por alguém (uma autoridade) com vistas a um destinatário (o subordinado). Expressam condutas que se deseja ver adotadas. Um tipo de regulamentos desse gênero são as Leis de Estado; outro, os comandos emitidos por autoridades militares ou eclesiásticas; outro, ainda, as proibições que pais impõem, para controle do comportamento dos filhos. Há um caráter prescritivo em regulamentos, usualmente associado a uma punição (ou ameaça de punição): quem não segue as normas está sujeito a um castigo. Em Medicina, talvez coubesse cogitar das receitas. O médico seria a autoridade, o paciente seria o subordinado; a receita seria o regulamento promulgado e a punição estaria implícita – deixando de seguir as instruções, o paciente pode não se curar.

Ao lado das diretrizes, das regras de um jogo e dos regulamentos, há normas "intermediárias", misto dessas formas principais. Os costumes, por exemplo, comparam-se a regras, pois determinam certos padrões de conduta; também se comparam a prescrições, pois exercem "pressão" sobre os elementos da comunidade, quase intimados a se ajustarem a tais padrões. Os princípios morais, por sua vez, se equiparam a prescrições (são vistos até como os mandamentos de Deus) e, não raro, se equiparam a normas técnicas (orientadas para alcançar um dado alvo).

Pelo prisma dessa discussão, *normal* é a *pessoa que se ajusta às normas*. E' normal quem, pretendendo atingir um objetivo, siga as instruções para alcançá-lo; quem obedeça às regras de um jogo e não execute lances proibidos; quem acompanhe os regulamentos. De outra parte, *anormal* é deixar de seguir as diretrizes, é executar lances ilícitos, é ignorar regulamentos, é fugir dos costumes, é dar as costas aos princípios morais.

Vale a pena insistir sobre o aspecto social das normas.

Há certos comportamentos que se repetem. São os costumes. Por outro lado, há comportamentos que não se tornam repetitivos ou que só atingem minorias – como os pseudocostumes ou, quando passageiros, os modismos. Costumes variam com o tempo e variam de uma para outra sociedade. Num dado grupo social, porém, costumes imperantes exercem pressão sobre os indivíduos, compelidos a adotá-los. Costumes controlados pela sociedade são as normas sociais. A norma social é, pois, um tipo de comportamento que a sociedade, por um motivo qualquer, aprecia ver adotado – a ponto de punir quem deixa de se comportar da maneira desejada.

Temos, assim, no âmbito social, as definições usuais de 'regra' e de 'norma'. A regra é um procedimento aceito, um costume ou hábito que se repete e é valorado pelo grupo social; é um princípio que rege interações sociais. Adquire, então, a característica de traço típico, usual, habitual, natural.

A sociedade raramente se preocupa em exercer controle sobre a salivação, o bocejo, o rubor e a curiosidade. Pouco se preocupa em controlar a tosse, o espirro, a maledicência, a prostituição e os tipos de adornos que os indivíduos preferem. Controla, freqüentemente, o consumo de narcóticos, o roubo, o aborto, o homicídio. Em geral, atenta para a alimentação e as relações sexuais e dá algum realce às trocas de informações. Existem, na sociedade moderna, algumas normas altamente valoradas (a monogamia, p. ex.) para as quais existem "códigos de boa conduta". A norma jurídica, ponto máximo desse controle social do comportamento, assume a forma de um par de enunciados que se complementam e completam: (1) se um ser humano *H* se filia a certo grupo *G*, deve manter um comportamento C; e (2) se H não mantém C, deve sofrer uma punição P. Essa P destina-se a preservar o valor de C (a conseqüência objetivada pela norma), algo que se encara como legítimo e digno de apreciação.

Nesse quadro de referência, *normal* é a pessoa que se submete à pressão das normas, que procede como se espera e cujas ações não conflitam com os ditames das normas. Já a palavra 'anormal' parece inadequada nesse quadro. De fato, quem foge às normas, quem se recusa a proceder de acordo com os costumes, não é exatamente "anormal" – é uma pessoa *diferente, excêntrica*. Talvez coubesse usar 'anômalo' – que significa, precisamente, "aquilo que foge à norma". (A valoração que parece presente em 'anormal' deixa de atuar quando se usa 'anômalo'.)

Passemos para um quadro de referência mais amplo. Pensemos na alimentação, biologicamente imperativa. O corpo necessita de carbohidratos, proteínas, cálcio, fósforo, ferro e outros minerais. O apetite, entre animais, é bom indício das necessidades do corpo. Alguns experimentos revelam que certos animais ingerem alimentos que lhes dão os nutrientes necessários e fazem crer que os humanos também ingerem alimentos de acordo com suas necessidades. Há, portanto, uma espécie de "sabedoria do corpo", atuando como guia para dietas adequadas. Quase todos os animais de uma dada espécie (racionais inclusive) têm aspecto exterior mais ou menos fácil de reconhecer, "comum", digamos; e têm órgãos de tamanhos "apropriados" que executam algumas atividades "corriqueiras", segundo padrões uniformes, ou "usuais". São, pois, *normais* – entendendo que o vocábulo indica usual, corriqueiro, costumeiro, uma espécie de "média".

Assim posta a questão, *anômalo* é o animal cujo aspecto está "fora da média", cujos órgãos se apresentam deformados, deixando de executar as atividades corriqueiras ou executando-as segundo padrões inusitados.

Acontece que os seres humanos fogem, muitas vezes, dos padrões que a "sabedoria do corpo" aprecia ver seguidos. Exemplificando, notemos que os chineses adotam o arroz beneficiado em vez do comum; a carne de cavalo é apreciada em algumas regiões da Europa, mas pouco apreciada na América; o leite de vaca é visto, em certos países, como algo repulsivo; índios brasileiros comem insetos que "paladares educados" jamais aceitariam; o fumo, o café, o chá, o álcool, por exemplo, conflitam com a "sabedoria do corpo". A cultura desenha, em boa medida, certos costumes não obrigatoriamente condizentes com o que, hoje, se imaginaria *saudável*. Índios deformam lábios e põem tintas sobre o corpo; chineses costumavam deformar os pés de suas donzelas; senhoras usavam espartilhos e ainda agora não abandonam sapatos de saltos altos, mesmo quando provocam tensões dolorosas nos tornozelos.

Diante de todas essas diferentes maneiras de reagir face à "sabedoria do corpo", outra idéia de "normalidade" se apresenta: *normal* é o *comportamento efetivo, real, adotado pela maioria*, ainda quando esse comportamento, por diversos prismas, deixa de ser "aceitável".

Eis, pois, itens a reter, a propósito de normas. Em primeiro lugar, correspondem a um admissível padrão de comportamento. E' nesse sentido que se fala de máquinas cujo funcionamento é "normal", de órgãos que funcionam "normalmente", ou de pessoas "normais" (lembrando que merecem censura social as pessoas que ignorem as normas).

Em segundo lugar, normas se associam, em certas circunstâncias, ao comportamento "da maioria". Aludindo a pessoas, parece "normal", entre senhoras de boa educação, usar sapatos de saltos altos e pintar as faces; como é "normal", para a maioria dos brasileiros, gostar de Carnaval e de futebol; ou, entre adolescentes, fumar e dançar.

Por um terceiro ângulo, as normas fixam índices de perfeição, ou excelência. É "normal", em tal caso, tentar alcançar, p. ex., elevados padrões de capacidade atlética, mesmo que isso contribua, digamos, para deformar o coração (deixando-o maior do que o usual – entendendo 'usual' de acordo com os dois prismas anteriores).

O termo 'norma' admite, pois, vários significados – que se transportam, muitas vezes, para 'normal'. Há, no entanto, um denominador comum na base desses significados. A *norma* remete, em geral, a uma regra, ou a um conjunto de regras: regras de um jogo; regras que conduzem a um objetivo; prescrições (a que se associam regras de comportamento ditadas pelos costumes). Em alguns casos, norma lembra lei. Vale a pena sublinhar que 'lei' corresponde, no Grego, a 'nomos' (de onde deriva 'anômalo' = o que não está conforme a norma). O termo 'anômalo' é descritivo; 'anormal', porém, ao lado de sua carga descritiva, admite, ainda, uma carga valorativa.

Em Medicina, o que foge à norma se indica por meio de prefixos como 'hipo' (hipocondria) ou 'hiper' (hipercrinia). Indispensável, porém, é entender com mais clareza o que seja 'normal', sob pena de a expressão 'fugir à norma' perder sentido. Valendo-se de muitos índices mensuráveis (sobretudo biofísicos e bioquímicos), examinando minuciosamente numerosos atributos biológicos, os especialistas, em trabalho conjunto, conseguiram fixar a normalidade em termos estatísticos – quase sempre acoplados às noções de "comum" e "anômalo". O assunto merece atenção.

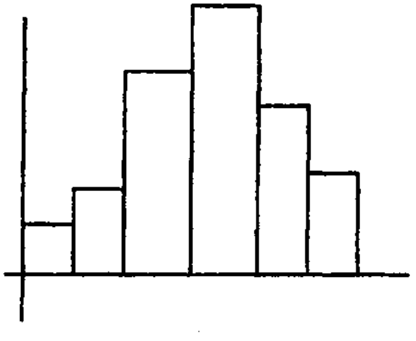
Normalidade estatística

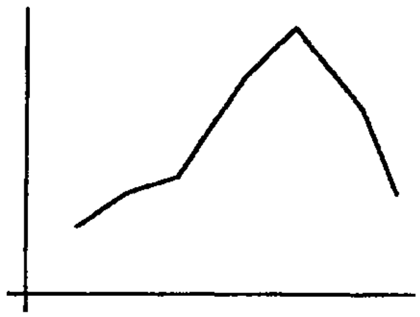
Cogitaremos da normalidade em termos daquilo que é comum, ou usual. Admitiremos que os atributos de interesse possam ser medidos, autorizando, pois, comparações e análises numéricas. Mergulhamos, desse modo, nos domínios da Estatística. Fundamentalmente, a palavra 'estatística' é empregada para fazer alusão a certas inferências feitas *em situações de incerteza*, inferências em que nos valemos de números a fim de determinar alguns aspectos da natureza e da experiência.

Com poucas e triviais exceções, os eventos e fenômenos são muito numerosos, muito variados e muito complexos para que se possa imaginar uma "observação completa". Não é possível, por exemplo, experimentar uma droga em todas as pessoas e nem mesmo em todos os doentes de um país. De um ponto fixo na Terra não é viável examinar a Lua em todos os momentos. Não podemos conhecer a opinião de todos os moradores de uma grande cidade, relativamente a um tema controverso qualquer. Precisamos, por isso, das *amostras*. Medidas que aparecem em qualquer experimento científico são amostras de ilimitada seqüência de medidas que resultaria de repetição indefinida desse experimento, vez após outra. O total de medidas seria, nesse caso, uma *população*.

De início, a Estatística enfrenta os problemas da *formação* e da clara *descrição* das amostras. O melhor meio de construí-las, como se sabe, é o da aleatoriedade: amostras são "boas" ("não viciadas", "não tendenciosas") quando resultam de escolhas feitas ao acaso. A descrição, por sua vez, pode ser feita com auxílio de tabelas, histogramas, curvas de freqüência – termos familiares, mas cujos significados serão brevemente relembrados a seguir.

Efetuando medidas, obtemos listas de números. Podemos dispô-los em ordem (crescente ou decrescente), o que permite verificar existência de itens repetidos ou de concentração de itens em torno de certos valores privilegiados. Em geral, a lista é transformada em tabela de freqüência, em que os dados se mostram de modo mais compacto. Quando as medidas não são discretas, é mais interessante considerar *intervalos*. Por exemplo, tratando-se do peso de seres humanos adultos, envolvendo quilos e gramas, vale a pena formar intervalos que abrangem, digamos, pesos entre 45 e 50 quilos; entre 50 e 55 quilos; entre 55 e 60 quilos, e assim por diante. Cada intervalo contém, pois, grupos de itens. O número de intervalos varia conforme a situação analisada. Em geral, tomam-se de 5 a 15 intervalos. A freqüência, nesse caso, é o número de itens em cada intervalo. Formam-se, dessa maneira, os *histogramas* e, a partir deles, os *polígonos de freqüência*. As figuras são familiares e não será preciso acrescentar maiores explicações.





Quando as amostras contêm crescente número de elementos, passa-se dos polígonos para as *curvas de freqüências* (que acompanham o desenho dos polígonos).

Comumente, uma curva de freqüências tem a forma de *sino* – a conhecida *curva de Gauss*. Medidas que se distribuem segundo a curva de Gauss revelam o que se chama *tendência central*: os valores se acumulam em torno de um ponto, mais ou menos próximo do centro, e a freqüência diminui à medida que nos afastamos desse ponto, para a direita ou para a esquerda.

São muitos os aspectos biológicos em que a distribuição de freqüências toma a forma de uma curva de Gauss. Esse fato despertou a curiosidade de estudiosos preocupados em explicar porque os atributos em causa haveriam de se distribuir desse modo. Pensando em altura, peso, envergadura, capacidade pulmonar, número de batimentos cardíacos, etc., em seres vivos (particularmente humanos), esses estudiosos concluíram que há distribuição gaussiana em virtude do risco de vida que pesa sobre os "extremos", um risco sensivelmente menor para os valores "intermediários". Notando que o peso de recém-nascidos também se distribui gaussianamente, os estudiosos imaginaram que a hereditariedade também é fator ponderável a considerar. A par disso, constatou-se que o comportamento das pessoas, fruto da ação ambiental, também leva a variáveis em que há distribuição gaussiana.

Exemplo curioso e ilustrativo é o do comportamento de motoristas, diante de um aviso "Pare". Ora o aviso é ignorado (um por cento dos motoristas), ora é levado "ao pé da letra" (um por cento deles); ora é considerado (21% que reduzem sensivelmente a velocidade); ora pouco obedecido (11% que mantêm a velocidade anterior); ora, enfim, é levado a sério (61% dos motoristas – que reduzem, mas não muito, as velocidades em que mantinham seus veículos.

Seja como for (risco de vida, hereditariedade, ação ambiental), o fato é que são muitos os atributos em que se apresenta a distribuição gaussiana. Nessas distribuições, temos as chamadas *medidas de tendência central*: média, mediana e moda. Elas são fundamentais para a adequada caracterização de *normalidade*. Embora muito conhecidos, não custa rememorar os significados desses termos.

A *moda* é o valor que se apresenta com maior freqüência. E' de interesse, por exemplo, para um fabricante de fardas – ao fixar quantas fardas, de cada tamanho, será preciso manter em estoque. Para esse fabricante, a moda é o "normal". Em casos desse tipo, 'moda' e 'normal' são vistos como termos sinônimos.

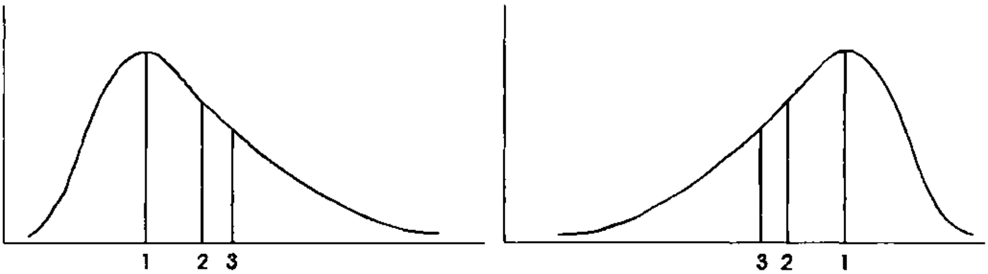
A palavra 'média' associa-se, em geral, à média aritmética. Como elemento de informação, a média não é muito satisfatória.

Vejamos um curioso exemplo. Um jovem deseja passar uns dias na praia. Recebe a seguinte informação: na praia *P* estão 14 pessoas do sexo feminino, cujas idades levam à média de 19 anos; na *Q* estão outras 14, cujas idades levam à média de 31 anos. Delibera, pois, ficar em *P*. Erro sério. Explica-se. Na praia *P* há pessoas com idades 2, 2, 2, 4, 5, 7, 10, 11, 34, 35, 35, 50 e 58 (média 19) – ou seja, as vovós com as netinhas. Na praia *Q*, há pessoas com idades 18, 19, 19, 19, 19, 19, 20, 20, 45, 45, 46, 47, 48 e 50 (média 31) – ou seja, as mamães e oito filhas cuja companhia era desejada.

As médias raramente são indicativas de normalidade. O "normal" é, antes, o "típico", o "paradigmático", não o "médio". Mesmo em expressões como "brasileiro médio", "aluno médio" etc., o que se tem em mente é um representante mais ou menos típico – não alguém que esteja "na média".

A *mediana*, por sua vez, é um valor que ocupa a posição central, em uma lista de valores, deixando 50% deles "para cima" e outros 50% "para baixo". Dificilmente a mediana retrata, por algum prisma, o que se diria "normal", numa distribuição qualquer.

Nas distribuições gaussianas (a curva de freqüências toma a conhecida forma de sino, com simetria em relação ao "pico"), as modas, as médias e as medianas se confundem. Pensando em normalidade em termos de *o mais comum,* ou seja, o item que se apresenta com maior freqüência, parece razoável usar a média aritmética para identificar esse item mais freqüente, "normal". Se, no entanto, as curvas de freqüência são esconsas, a moda difere da média e deixa de parecer razoável supor que esta última indique normalidade. E' o que se ilustra, sem mais comentários, com as duas figuras seguintes.



1-Moda
2-Mediana
3-Média

1-Moda
2-Mediana
3-Média

Uma distribuição pode apresentar duas ou mais modas, isto é, dois ou mais valores que se apresentam com freqüências relativamente elevadas. A distribuição se diz, então, *multimodal* e, nesse caso, torna-se ainda menos razoável equiparar o normal à média ou à moda. Voltando ao exemplo das praias, um freqüentador da praia *P* não seria "normal" por ter 19 anos (média) ou 2 anos (uma das modas).

Mesmo em distribuições gaussianas (quando é tolerável entender "normal" em termos de "média"), há desvios perfeitamente aceitáveis, com respeito à média. Exemplo: um aluno tem média 8 em um ano letivo; isso não o impede de ter obtido um 4, por exemplo, nota "equilibrada" por três notas altas, 9, 9 e 10. A par disso, numa turma de bom rendimento – digamos média 75 – que sentido teria dizer que um aluno cuja nota foi 73 é "mais normal" do que seu colega que obteve 71?

As dificuldades apontadas levaram os estudiosos a considerar *medidas de dispersão*. Quanto podemos nos afastar da média sem perder normalidade? Entre as medidas de dispersão, a de maior interesse prático tem sido o *desvio padrão*. A noção pode ser facilmente compreendida mediante exemplo. Sejam dados os valores 12, 13, 16 e 19. A média é 15. Os "afastamentos", com respeito à média, são

12 - 15 = -3; 13 - 15 = -2; 16 - 15 = +1; 19 - 15 = +4.

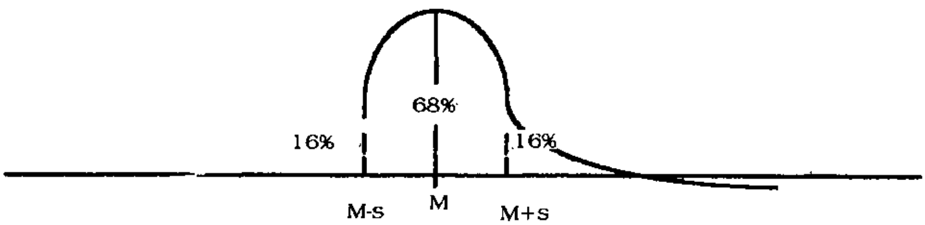
A soma algébrica dos afastamentos é (-3)+(-2)+(+1)+(+4) = 0 (propriedade que se verifica sempre). Desconsiderando os sinais, porém, a soma 3+2+1+4 = 10 difere de zero. A média aritmética dos afastamentos (em valor absoluto, ou seja, sem os sinais) é o que se denomina *desvio médio*. No exemplo, 10/4 = 2,5. Há outra maneira de desconsiderar os sinais. Tomam-se os quadrados dos afastamentos e, a seguir, a média aritmética. Resulta a chamada *variância*. No exemplo, os quadrados são 9, 4, 1 e 16. Média: [(9+4+1+16)/4] = 7,5 . Daí se obtém o *desvio padrão*, a raiz quadrada da variância. No caso, a variância, *s*, é a raiz quadrada de 7,5 – ou seja, *s* = 2,77 .

Variância e desvio padrão mostram-se muito úteis medidas de dispersão, pois relativamente pequenos quando as medidas originais se agrupam em torno da média e, em oposição, relativamente grandes quando estas se afastam amplamente da média.

Pensando ainda em termos de médias, *normal* seria o que não "foge demais" da média; *anormal*, o que dela apreciavelmente se aparta. A noção de "afastar-se pouco" torna-se precisa mediante uso do desvio padrão. Tem-se:

normal = o que permanece no intervalo (M-s , M+s)

ou seja, aquilo que não se afasta mais do que *s* (desvio padrão) unidades da média *M*. Empiricamente, comprova-se que "normais", assim entendidos, são cerca de 68% dos itens considerados (*desde que a distribuição original seja gaussiana)*. Portanto, 32% dos itens são "anormais", 16 "para mais", outros 16 "para menos". A figura (bem conhecida) ilustra o que ficou dito:



Alguns autores preferem considerar o dobro do desvio padrão *s*, entendendo que o intervalo de normalidade seria (M-2s , M+2s). Nesse caso, aproximadamente 95% dos itens são "normais", com 2,5% "para menos" e 2,5% "para mais", fora do intervalo. Temos, pois, como entender a *normalidade estatística*.

A Estatística tem por objetivo fundamentar certas inferências realizadas em condições de incerteza. De hábito, não podemos conhecer todos os objetos de uma vasta coleção (a "população"). Trabalhamos com amostras. Através delas, cogitamos das propriedades dos objetos da coleção. Depois de coligir amostra e de descrevê-la de modo apropriado, o estatístico enfrenta duas questões importantes. Primeira: partindo da evidência de que se dispõe nesta particular amostra, quais as melhores conclusões a estabelecer, relativas à população? Segunda: quão dignas de confiança podem ser tais conclusões?

Notar que aludir a uma população é fazer referência a itens não observados. Daí a pergunta: com que "certeza" as ilações (obtidas via amostra) se aplicam a um particular item não observado, aleatoriamente escolhido? O problema do caso individual é um dos mais delicados da Estatística e tudo leva a supor que não encontrará, por ora, solução satisfatória. Em verdade, o *indivíduo psicológico* (uma pessoa particular) não se importa com a conclusão de que 90% dos pacientes anteriormente submetidos a um tipo de cirurgia se recuperaram em dois ou três dias; deseja saber o que acontecerá a ele, *como caso específico*. Embora a Medicina trabalhe com o *indivíduo epistêmico*, o fato é que, para o indivíduo psicológico, a pergunta sobre uma recuperação, após a cirurgia, pede resposta categórica, em termos de "sim" ou "não". Para a Estatística, no entanto, o indivíduo psicológico é apenas um dos elementos de certa população e ele – como, de resto, qualquer outro, nessa população – está com 90% de "chances" de rápida recuperação, pois foi exatamente isso que se fixou em termos de indivíduo epistêmico.

Felizmente, as populações em que há distribuições gaussianas de freqüências têm uma importante propriedade que autoriza contemplar com otimismo a dificuldade do *caso individual*. A propriedade em causa é a seguinte: *as amostras, em populações gaussianas,* herdam *a distribuição.*

O assunto merece atenção. Imagine-se ter certa população ("população-mãe"), em que a distribuição é gaussiana, com dada média e determinado desvio padrão. Da população obtêm-se amostras aleatórias, com certo número fixo (digamos n) de elementos. Calcula-se a média de cada amostra. Verifica-se que *a população das médias* também se distribui gaussianamente, tal qual a população-mãe. Em vista do processo de "balanceamento" (cálculo das médias), os novos dados acham-se, agora, mais aglomerados do que os dados iniciais, relativos à população-mãe.

A par disso, o desvio padrão é determinado dividindo o desvio padrão da população-mãe pela raiz quadrada de n, onde n é número de elementos de cada amostra.

Vejamos um exemplo concreto. Imagine-se população-mãe com vários milhares de itens. Formam-se algumas dezenas de amostras, cada qual delas com 64 itens. Calculam-se as médias de tais amostras. Essas médias se distribuem gaussianamente. O desvio padrão das médias dessas amostras será um oitavo (8 é a raiz quadrada de 64) do desvio padrão da população-mãe.

Em outras palavras, a distribuição gaussiana tem aspecto "reprodutivo": as "filhas" herdam a propriedade relevante, ou seja, continuam gaussianas.

Inferências estatísticas

Há um ponto importante a acrescentar ao registrado anteriormente: é muito confortador saber que as médias de amostras amplas, recolhidas de uma população qualquer, ***não necessariamente gaussiana,*** têm distribuição aproximadamente gaussiana. Isso, naturalmente, simplifica apreciavelmente numerosos cálculos.

As informações estatísticas permitem "boas" respostas para as questões básicas que nos importam: (1) quais as melhores ilações relativas à população e (2) com que confiança tais ilações podem ser recebidas?

Ressaltemos os aspectos notáveis utilizando exemplo concreto. Imagine-se que um dado grupo de operários tenha sido examinado ao longo de alguns anos e que o tempo de trabalho de cada servidor tenha sido próximo de 33 anos e alguns meses, com desvio padrão da ordem de dois anos – mais precisamente, 10 mil dias, com desvio padrão de 800 dias.

Forme-se amostra com 64 servidores, sujeitos a tratamento especial (p. ex., melhor alimentação, atenção às reivindicações, ambiente não poluído, ou algo do gênero). Submetendo a teste essa amostra, verifica-se que os trabalhadores que a integram trabalharam 10.200 dias, prolongando, assim, por 200 dias, o período útil "habitual". E' possível que a diferença não seja "efetiva": a amostra recolhida foi, casualmente, um pouco mais "favorável" do que seriam outras amostras eventualmente recolhidas. O que importa, naturalmente, é saber se o prolongamento do período ativo é real ou meramente casual.

Para saber se a melhoria foi efetiva, formam-se diversas amostras com 64 itens. O desvio padrão das médias de amostras de 64 itens é um oitavo do desvio padrão da população original. Um oitavo de 800 é igual a 100. O aperfeiçoamento, na particular amostra, foi de 200 – o dobro do desvio padrão das médias.

A teoria das probabilidades afirma que uma diferença dessa magnitude só se deve ao acaso uma vez em cada vinte. Dito de outro modo, o aperfeiçoamento constatado tem 19 possibilidades de ser "real" e apenas uma (em cada vinte) de ser "casual". Parece razoável admitir, portanto, que a observação não foi excepcional, mas, ao contrário, que a amostra seria fruto de uma população cujo tempo de serviço útil se teria, de fato, tornado maior do que o esperado. Os estudiosos concluem que aquele tratamento especial prolonga, sem dúvida, o tempo de trabalho dos servidores.

A moderna Estatística opera com alguns parâmetros mais complicados, como o *intervalo de confiança* e os *coeficientes de confiabilidade.* O estatístico não pretende alcançar conclusões irretorquíveis. Suas análises conduzem, de hábito, a um par de asserções, assim resumidas: (1) "A melhor estimativa cabível é a de que ... "; e (2) "O grau de confiança com que tal estimativa há-de ser acolhida é de ..." .

Em geral, o estatístico principia escolhendo um número – o coeficiente de confiabilidade. Imaginemos que tenha sido 0,95. Significa: adotar-se-á procedimento que estará correto, em média, 95% das vezes. Tem-se, então, um índice de confiabilidade dos resultados obtidos. Escolhido o coeficiente de confiabilidade, a Estatística teórica nos dá a amplitude do intervalo de confiabilidade – um intervalo cujo ponto central é a média das amostras.

Voltando ao exemplo anterior, uma série de cálculos efetuados por especialistas permite fixar o índice 195. Isso determina o intervalo de confiabilidade, (10.200 - 195 ,10.200 + 195), ou seja, o intervalo (10.005 , 10.395).

De posse de tais informações, eis como reage o estatístico:

> *Avalio que a média da população dos tempos de trabalho dos servidores submetidos a tratamento especial é maior do que 10.005 e menor do que 10.395 dias. Não posso garantir que esteja correto. Mas, em ampla coleção de afirmativas desse gênero, estarei certo 95% das vezes. Considero que o intervalo de 10.005 até 10.395 está acima do tempo de trabalho anteriormente constatado. Concluo que o tratamento especial tem efeito: prolonga, de fato, o período de atividade dos servidores.*

Se o estudioso escolhesse outro número para atuar como coeficiente de confiabilidade (p. ex., 98, em vez de 95), o intervalo de confiabilidade também seria alterado. Aumentando o coeficiente, diz-se algo menos preciso, porém mais digno de confiança. Em oposição, diminuindo o coeficiente, diz-se algo menos confiável, porém mais preciso.

As idéias subjacentes são facilmente compreensíveis quando situadas no contexto de exemplos concretos. Assim, quem diz "Vai chover" afirma algo muito provável e pouco preciso – "Vai chover em algum lugar, em algum momento" é provável, mas pouco informativo. Quem diz "Vai chover aqui onde estamos" afirma algo mais preciso e menos provável. Dizendo "Vai chover aqui, às 16 horas", aumenta a precisão, diminuindo a probabilidade. Cada afirmação está, para a anterior, na mesma relação em que se acham as correspondentes afirmações a respeito de precisão e confiabilidade: um desses atributos cresce em detrimento do outro.

Essas, afinal, as noções fundamentais que permitem caracterizar *normalidade*, em termos estatísticos. E' a partir delas que se delimita a noção de *normalidade* em Medicina.

Indicações bibliográficas

Comentários em torno de 'norma' e 'normalidade' estão assentados no que diz von Wright, em *Norm and Action* (1963).

A normalidade estatística tem sido amplamente discutida em numerosos livros e artigos. Particularmente bem feito e de fácil acesso é o artigo "Statistics", de Weaver, publicado em *Scientific American*, 1952. O artigo foi incluído, com vários outros, igualmente interessantes e importantes, em *Mathematics in the Modern World*, antologia organizada por Kline (1958; várias edições posteriores). Ver, também, *The World of Mathematics*, antologia organizada por Newman (1956), especialmente volume 2, partes VI e VII.

Noções a respeito de médias, medianas, medidas de tendência central, desvio padrão, normalidade, etc., encontram-se em qualquer bom livro de Estatística. Entre muitos, ver, p. ex., *Statistical Analysis*, de Edwards (1958, ed. rev.) ou *Concepts of Statistical Inference*, de Guenther (1953). O assunto foi divulgado em nosso idioma em várias obras. Ver, p. ex., *Fundamentos da Estatística*, de Nick e Kellner (1980).

Talvez seja oportuno dialogar com especialistas, a fim de atualizar conhecimentos – chegando a livros mais recentes. Lembrar, porém, que o assunto aqui examinado não sofreu alterações dignas de nota.