

## Estatística e experimentação pedagógica

JACYR MAIA

Técnico de Educação  
chefe da Secção de Orientação  
e Seleção Profissional  
do I. N. E. P.

### 1. A investigação sobre fatos sociais

As investigações de cunho científico realizadas sobre os fenômenos sociais apresentam-se com características diferentes das que tenham por objeto os fatos físico-químicos. A interpretação dos resultados, em ambos os casos, é sempre feita em termos de probabilidades. Mas, a circunstância de ser a variabilidade dos fenômenos sociais muito maior que a dos físicos (e tanto que por vezes nos dão a ilusão de invariáveis) leva-nos à necessidade da aplicação de métodos tais que nos permitam conhecer até que ponto as observações possam ser tomadas como válidas.

Há ainda a considerar não permitirem os fenômenos sociais o que se poderia chamar de "experimentação provocada". O físico demonstrará, quantas vezes o desejar, que o calor dilata os corpos ou que o som não se propaga no vácuo. Repetirá a experiência, em situações idênticas, duas, três, dez, cem vezes... O mesmo já não acontece com o sociólogo, o psicólogo ou o educador. Esses terão que se utilizar de cada experiência em particular para chegar às suas conclusões e para generalizá-las, quando possível.

Por isso mesmo, as suas investigações terão que ser efetuadas dentro de condições especiais. E, muitas vezes, embora efetuadas com todo o rigor e a melhor das intenções, é possível que os resultados a que se chegue não possam ser utilizados com caráter de previsão senão com erro muito grande. Será, contudo, sempre uma informação que poderá orientar o pesquisador, de ma-

neira mais ou menos segura. Dará sempre um conhecimento sobre o grupo observado.

\* \* \*

Assim acontece com todas as experimentações no campo da educação. Por esse motivo, muitas das experiências levadas a efeito aqui no Brasil têm sido acoimadas de inúteis. A verdade porém, é que, na grande maioria dos casos, a afirmação acima não corresponde aos fatos. O que parece haver, quasi sempre, é que a essência dos métodos usados para análise, a interpretação dos resultados dessas investigações, e o alcance do uso que se poderá fazer daqueles resultados colhidos, são ainda mal conhecidos.

Tivemos a oportunidade de trabalhar, durante cerca de cinco anos, no Instituto de Pesquisas Educacionais da S. G. E. C. do Distrito Federal. Ai tivemos a ocasião de acompanhar numerosas pesquisas sobre diversos assuntos de educação e pudemos, nós próprios, testemunhar as dificuldades de toda ordem que surgem em cada caso.

Uma das primeiras conclusões a que nos levou a prática foi a de que o experimentador tem necessidade de se contentar, inicialmente, com o imperfeito. A paciência foi, também, outra condição que a experiência nos fez supor indispensável, tão necessária quanto os conhecimentos técnicos do assunto.

Recordamo-nos, com alguma satisfação, da nossa luta para o aperfeiçoamento dos testes de aproveitamento, tão cheios de falhas no princípio

e quasi bons ao fim de alguns anos... Do trabalho imenso que foi necessário para que pudéssemos demonstrar aos professores — isso ao cabo de quatro anos — que a seleção dos analfabetos deveria ser feita, realmente, por uma prova que revelasse a capacidade para a aprendizagem das técnicas fundamentais de leitura e escrita. Muitos outros exemplos poderíamos citar... Se nos perguntassem de que depende o êxito das pesquisas educacionais é possível que não hesitássemos em afirmar que, grande parte, provém de dois fatores: conhecimento seguro do método estatístico e... paciência.

Aquele é a poderosa e perigosa arma indispensável ao investigador. Sem ele nada se fará; só com ele também nada se conseguirá. Na interpretação dos resultados deverá estar presente essa coisa tão simples que se chama bom senso. O método estatístico, aí, representa apenas o bom senso, expresso em números. A paciência é a qualidade básica para que possa ser levada a cabo a investigação.

Mas esse mesmo método estatístico, que tanto pode esclarecer o investigador, pode também, por outro lado, iludí-lo de forma desconcertante. Não raras vezes se chega a conclusões falsas, por interpretações errôneas, motivadas pelo emprego de processos que se destinam a fornecer outras informações. A escolha da medida apropriada ou, mais claramente, a escolha da fórmula matemática adequada é, portanto, das partes mais delicadas nas investigações.

## 2. Significativo e representativo

Confusão muito frequente pode ser observada entre as idéias de significativo e representativo.

Em estatística, diz-se que um resultado é significativo quando, com os dados obtidos experimentalmente, podemos afirmar (com probabilidade praticamente equivalente à certeza) que ele não ocorreu por mero acaso. As medidas que nos permitem essas conclusões são conseguidas por cálculos baseados na teoria dos erros. E nenhum resultado de qualquer observação deve ser tomado como válido sem essa verificação.

Diz-se que um grupo é representativo quando temos elementos para supor (e, em certos casos, até para demonstrar) que o mesmo possui características idênticas às do grupo total. É o problema das amostras de observações.

O problema da representatividade aparece, portanto, antes da investigação. Levado, na grande maioria dos casos, pela impossibilidade prática de trabalhar sobre todos os indivíduos de um grupo ou universo total, o experimentador lança mão de uma amostra. Para tanto é necessário que conheça a técnica da obtenção de uma amostra que possua as características do grupo total que ela irá representar. Assim sendo, ela, a amostra, será colhida "a priori". O que acontece é que os resultados que se obtêm vêm, naturalmente, afetados de erros, que podemos chamar de "erros de flutuação de amostras". Um erro maior ou menor não indica, neste caso, de maneira alguma, que o grupo sobre o qual se investiga, não é representativo. E esse erro varia em função da extensão da amostra e da variabilidade do fenômeno observado. Na prática, obtém-se uma amostra representativa tomando-se, rigorosamente ao acaso, elementos do grupo total. Convém notar que não importa, para a representatividade, que o grupo do qual se tirou a amostra seja homogêneo ou não, selecionado ou não. O importante é que se possa admitir que a amostra possui as características do grupo total.

Quanto à extensão da amostra, está em função da exatidão ou precisão que exigirmos para o nosso trabalho.

"Qual o número mínimo de observações a serem feitas para que uma amostra seja "significativa"? Eis uma pergunta tão ingênua quanto frequente entre pessoas que se dedicam às pesquisas no campo bio-social.

Esta pergunta deverá sempre incluir também a precisão que se quiser exigir dos dados experimentais.

## 3. Equivalência de grupos

Sendo as observações dos fenômenos sociais de caráter todo especial, não susceptíveis de repetição em situações rigorosamente iguais, apresentam ao investigador um sério problema: cada vez que ele quiser se certificar das consequências de determinado fator em estudo, terá que apelar para novos grupos de indivíduos. Uma nova variável portanto, que influirá fatalmente nos resultados finais.

Exemplifiquemos. Imagine-se que o educador queira experimentar um novo processo de ensino. Para isso utiliza duas turmas de um mesmo nível escolar. Em uma aplica o novo processo; a outra continuará o aprendizado pelo processo em uso.

Imagine-se agora que, ao cabo de algum tempo, ele tenha dados experimentais e objetivos que permitam afirmar que os alunos submetidos ao novo processo apresentaram, de fato, rendimento superior aos outros.

Admita-se, ainda, para simplificação do exemplo, a hipótese de serem os professores de capacidades iguais. Será lícito, nessa situação concluir que o novo processo seja superior? Absolutamente não. Há, aí, um problema fundamental, básico, ao qual toda e qualquer conclusão está sempre subordinada: "os grupos que serviram a experiência precisam ser equivalentes".

Diz-se que dois ou mais grupos são equivalentes quando apresentam, em seu conjunto, forças resultantes idênticas, isto é: deverão estes grupos encerrar, em si, constantes de concentração e dispersão semelhantes, sem diferenças significativas. São, assim, grupos que se podem substituir mutuamente, grupos que, para os efeitos práticos da nossa observação vão permitir, indiretamente, a repetição da experiência. Permitirão o juízo sobre a importância relativa do fator que se observa, como o novo processo de ensino, no exemplo dado.

Referindo-se em seu trabalho, "A experimentação pedagógica e o ante-projeto de decreto-lei sobre formação do magistério primário" (1), o técnico de educação Rvdmo. Pe. Helder Câmara faz um "depoimento público a respeito da experimentação em educação", afirmando que a mais "ampla experimentação pedagógica realizada no Brasil (no Distrito Federal)" *falhou*. Para tanto, examinando a documentação existente, chega a conclusão de que todo o "esforço fora praticamente inútil, pois haviam faltado à experimentação requisitos indispensáveis para justificar qualquer conclusão".

Um desses requisitos, segundo o autor, foi a *ausência de grupo representativo*. "Experimenta-se para verificar a conveniência ou inconveniência de estender a um campo mais largo o que foi experimentado em campo menor. O campo experimental, no entanto, precisa ser representativo, isto é, precisa ser tal que as conclusões nele obtidas possam, com segurança, ser interpretadas como conclusões do campo geral".

Esta condição, quando se trata de *observar a influência de um novo fator* (como no exemplo mesmo dos processos novos aplicados nas Escolas Experimentais), *nem sempre* se apresenta como necessária. Na experimentação em questão não existe, em última análise, o problema da representatividade. Se não vejamos.

Que se desejava quando foram experimentados novos processos didáticos naquelas escolas? Verificar se, realmente, o rendimento aumentaria.

Como poderia ser feita tal verificação?

Observemos antes que o que se deseja investigar é o valor de um fator externo, novo, que *atuaria* sobre o grupo de alunos: o novo processo de ensino. Portanto, precisamos de grupos de crianças *equivalentes* àqueles que serviram à observação. Como se vê, não interessa, no caso, que os grupos possam ou não *representar* o grupo total do sistema escolar. O importante e essencial é que a comparação seja feita com outros grupos *estatisticamente idênticos* aos observados. E eles terão que ser equivalentes de um ponto de vista global (meio social, idade, nível mental, etc.). Os resultados dessa comparação *poderão* então permitir as conclusões:

- a) superioridade do novo método;
- b) inferioridade do novo método;
- c) impossibilidade de se afirmar uma das duas hipóteses acima.

Essas conclusões estão condicionadas à aproximação ou precisão com que se desejar a informação que, por sua vez, dependerá do número de casos observados e da variação do fenômeno; no caso, a aprendizagem escolar.

Verificação dessa ordem, aliás, foi levada a efeito em 1937. Na falta de grupos equivalentes, os resultados das Escolas Experimentais, nas provas (2) aplicadas em dezembro, foram comparados com o de outras escolas situadas próximo a cada uma delas. Não queremos dizer que os alunos dessas escolas formassem grupos equivalentes

(2) Sabemos perfeitamente que essa prova não estava aferida. A validade das provas de aproveitamento é, contudo, quasi sempre bem maior do que comumente se pensa. Além disso, sendo a prova igual para todos os alunos, o erro obtido terá, para os efeitos da nossa comparação o mesmo sinal, positivo ou negativo e, assim, não perturbará sensivelmente a investigação. Não cabe aqui, porém, a demonstração estatística deste fato.

(1) In "Formação" — N. 32, março, 1941.

aos das experimentais. Chamamos até a atenção para o fato de ser essa comparação muito grosseira. Poderia, contudo, fornecer qualquer orientação. Como trabalhávamos, naquela época, na Secção de Medidas e Eficiência Escolares do Instituto de Pesquisas Educacionais possuímos, ainda, cópia de alguns dados, com os quais nos foi possível organizar a tabela que se segue:

**A MEDIDA DA CONFIANÇA DAS MÉDIAS E DA SIGNIFICÂNCIA DAS DIFERENÇAS DAS MÉDIAS DAS DIVERSAS ESCOLAS PERMITIU A SEGUINTE CLASSIFICAÇÃO :**

GRAU DE APROVEITAMENTO	1ª SÉRIE	2ª SÉRIE	3ª SÉRIE	4ª SÉRIE	5ª SÉRIE
1º Plano	Inst. Educação 1ª. Exp. = Ba. Ot.	2ª. Exp. = M. Bomf. 1ª. Exp. = Ba. Ot. Inst. Educação	Inst. Educação	Inst. Educação	6- 3 = N. Peçanha
2º Plano	3ª. Exp. = Argent. 4- 5 = Est. Sá	4-11 = Parêto 1- 7 = J. Nab. 5- 3 = H. Melo	5- 3 = H. Melo	2-14 = Sta. Catar. 5ª. Exp. = México 6- 3 = N. Peçanha 3ª. Exp. = Argent.	Inst. Educação
3º Plano	2ª. Exp. = M. Bomf. 5ª. Exp. = México 4- 3 = Fr. Caneca 5- 1 = Ben. Ot. 4-11 = Parêto	5ª. Exp. = México 4-13 = S. Pereira	1- 7 = J. Nabuco 5ª. Exp. = México 1- 9 = M. Viana 5- 1 = Be. Otoni 4- 5 = Est. Sá 6- 3 = N. Peçanha 4-11 = Parêto 4-13 = S. Pereira 3ª. Exp. = Argent.	1ª. Exp. = Ba. Otoni 2ª. Exp. = M. Bomf. 1-12 = M. Gerais	1ª. Exp. = Ba. Ot.
4º Plano	1- 7 = J. Nabuco 4-13 = S. Pereira 1- 9 = M. Viana 5- 3 = H. Melo 5-12 = Equador 6- 3 = M. Peçanha 4ª. Exp. = E. Unidos	6- 3 = N. Peçanha 6-12 = Equador 3ª. Ex. Argent. 4- 3 = Fr. Caneca 1- 9 = M. Viana	5-12 = Equador 2ª. Exp. M. Bomf. 4- 3 = Fr. Caneca 4ª. Exp. = E. Unidos 1ª. Exp. = Ba. Ot.	4-14 = A. P. Alegre 5- 2 = B. Menezes 4E = E. Unidos	5ª. Exp. = México
5º Plano		5- 1 = Be. Otoni 4- 5 = Est. Sá 4ª. Exp. = E. Unidos			2ª. Exp. M. Bomf. 2-14 = Sta. Catar.
6º Plano					3ª. E — Argent. 1-12 — M. Gerais 4ª. E = E. Unidos 5- 2 = B. Menezes 4-14 = A. P. Alegre

Os cálculos permitiram uma classificação das escolas em seis níveis, segundo o aproveitamento revelado nas provas.

Vejamos a 1.ª série :

A 1.ª Experimental, que havia sido comparada com as escolas Benedito Otoni, Nilo Peçanha e Escola Primária do Instituto de Educação (3)

(3) A localização dessas quatro escolas é a seguinte:  
1.ª Experimental — Rua Senador Furtado, 80.  
Benedito Otoni — Rua Senador Furtado, 90.  
Nilo Peçanha — Av. Pedro II, 398 ( a uns 500 metros da 1.ª Experimental).

Escola Primária do Instituto de Educação — Rua Mariz e Barros, 273 (cerca de 200 metros da 1.ª Experimental).

apresentou-se no 1.º nível juntamente com a do Instituto de Educação, enquanto a escola Benedito Otoni se localizou no 3.º e a escola Nilo Peçanha no 4.º nível.

Poderíamos então, de modo geral, supor maior o rendimento na 1.ª Experimental, em relação àquelas duas outras escolas.

A comparação feita desse modo dá apenas uma orientação geral, grosseira, em face da qual não se poderá ainda afirmar que o processo empregado na 1.ª escola experimental (sistema de projetos) tenha sido o fator determinante da diferença favorável encontrada.

\* \* \*

Comentando a falta de representatividade do grupo observado, o autor citado afirma que "certeza de que um grupo é representativo da massa geral só é possível ter a *posteriori*". O contrário exatamente do que se verifica, pois o que a Estatística nos ensina é justamente formar grupos

representativos *a priori*; aliás, só assim esses grupos poderão ser uteis ao pesquisador. Para a verificação da representatividade, aplica ele a fórmula :

$$N = \left( \frac{3 \sigma}{0.01 M} \right) + 1$$

onde:

N = Número procurado

$\sigma$  = desvio padrão da distribuição

M = a média aritmética

3 = número de sigmas que permite praticamente a certeza

0,01 = precisão desejada, expressa em função da média.

Aquí cabe um esclarecimento.

A precisão de uma média de observações pode ser dada em função do seu erro padrão, cuja fórmula é a seguinte :

$$\sigma_M = \frac{\text{D. P. distrib.}}{\sqrt{N}}$$

O valor de  $\sigma_M$  fornecerá, em termos de probabilidades, os limites dentro dos quais poderá oscilar a média. Assim é que, se somarmos e subtrairmos à média 1 vez o valor de  $\sigma_M$  teremos cerca de 2 probabilidades contra 1 de estar ali compreendido o valor real da média. À medida que aumentarmos esse limite, crescerá a probabilidade de ser encontrada a média naquela região. Se quisermos então jogar com probabilidade praticamente igual à certeza, poderemos tomar 3 erros padrões para cada lado da média, o que nos dará 99,73% de probabilidades, ou cerca de 369 contra 1, de serem aqueles pontos os limites dentro dos quais poderá oscilar fortuitamente a média encontrada na experiência.

Invertamos agora a questão. Tomemos o exemplo da 1.<sup>a</sup> escola experimental da 1.<sup>a</sup> série, 1939, e suponhamos que se queira uma precisão de 95% para que a nossa média tenha um erro igual ou menor que uma unidade (aproximação exagerada e inútil para o caso). Teríamos que determinar o valor de N a partir dessa base. Vejamos qual seria esse valor.

Sendo

$$\sigma_M = \frac{\text{D. P.}}{\sqrt{N}}$$

sabemos que há praticamente certeza (probabilidade favorável de 99,73%) de que a média só poderá oscilar fortuitamente entre

$$M \pm \frac{3 \text{ D. P.}}{\sqrt{N}}$$

Para o nosso exemplo — 95% de probabilidade para que o erro de M seja igual ou menor que 1 — teríamos :

$$1,96 \sigma_M = 1$$

$$\sigma_M = \frac{1}{1,96}$$

$$\frac{1}{1,96} = \frac{22.995}{\sqrt{N}}$$

$$N = (1,96 \times 22.995)^2 = 2031 \text{ casos (4)}$$

O cálculo acima nos diz que : se quisermos uma precisão *exagerada* de um erro igual ou menor do que 1, numa escala centesimal, teremos que trabalhar sobre 2.031 casos. Uma turma com 2.031 alunos!... Evidentemente se se exigirem precisões dessa ordem, bem poucas vezes poderiam ser feitas pesquisas no campo da educação...

Suponhamos agora que, procurando estar mais perto da realidade, nos contentemos com um erro de 5. Admitamos que se exija 85% (grande precisão e perfeitamente satisfatória) de probabilidades para a verificação do erro.

Teríamos :

$$1,44 \sigma_M = 5$$

$$\sigma_M = \frac{5}{1,44}$$

$$\frac{5}{1,44} = \frac{\text{D. P.}}{\sqrt{N}}$$

(4) Tomámos, propositadamente exagerado nesse exemplo, erro de 5 e  $p = 99,73\%$ . Se bem que seja ainda cerca de 10 vezes menor que o exigido pelo autor, o número necessário desceu de 17.607 para 2.031.

$$\frac{D. P. 5}{\sqrt{N}} = \frac{5}{1,44} \quad 5 \sqrt{N} = 1,44 D. P.$$

$$N = \left( \frac{1,44 \times 22,955}{5} \right)^2 = 43,85 \text{ ou } 44$$

Por aí se vê que, com erro mais que satisfatório para o tipo de pesquisa em apreço, poderíamos perfeitamente nos contentar com 44 casos, *para efeito da segurança dos dados obtidos*. É o problema da *confiança* que oferecem os resultados da observação. Note-se que não nos estamos referindo à capacidade de representatividade do grupo observado. Mesmo porque a fórmula empregada não se destina a essa verificação.

Por esse motivo estranhamos o emprego da fórmula:

$$N = \left( \frac{3 \sigma}{0,01 M} \right)^2 + 1 \quad (5)$$

como indicação da representatividade do grupo. Por ela teremos noção apenas da significação dos valores.

Se fizermos os cálculos naquela base para todas as escolas experimentais da 1.<sup>a</sup> série, e para o conjunto delas, teremos:

1. <sup>a</sup> experimental	= N = 44
2. <sup>a</sup> "	= N = 32
3. <sup>a</sup> "	= N = 42
4. <sup>a</sup> "	= N = 48
Conjunto	= N = 45

Como se verifica, a confiança desses dados e até das mais satisfatórias, em relação à 1.<sup>a</sup> série,

(5) É a mesma fórmula usada aqui. Apenas a aproximação exigida aí foi de 1 centésimo da média, jogando-se ainda com certeza prática para a determinação de N. O 1 que aparece na fórmula é perfeitamente desprezível.

pois em todos os casos trabalhou-se com número suficiente para as informações com a precisão acima estipulada.

Aliás, a simples observação dos resultados a que chegou o autor nos deixa seriamente confusos. Pois seriam necessários mais de 70 mil alunos só na 1.<sup>a</sup> série... Esse resultado estranho provem do emprego de uma fórmula que se destina a fornecer informações de outro tipo. Nada tem a ver com a representatividade.

#### CONCLUSÕES

a) a experimentação social se apresenta sob condições especiais e exige, por isso, métodos especiais que permitam avaliar até que ponto são válidas as observações;

b) a confusão por vezes observada entre alguns conceitos básicos da Estatística, tem levado investigadores a interpretações falsas;

c) essa confusão se apresenta particularmente frequente em relação aos conceitos de *representativo* e *significativo*;

d) o conceito de *equivalência de grupos* não raras vezes tem sido abandonado ou confundido com o de *grupos representativos*, nas pesquisas sociais;

e) é essencialmente indispensável ajustar a precisão dos valores obtidos com os objetivos da pesquisa que se realiza;

f) quando se trata de observar a influência de um fator novo, sobre o grupo, nem sempre estamos diante de problema de representatividade;

g) o emprego de uma fórmula destinada a fornecer informações sobre a confiança dos valores, com o fim de se avaliar da capacidade de representatividade dos grupos de alunos das escolas experimentais, levou a resultados absurdos;

h) se se desejar sempre precisão da ordem da que imaginam necessária certos críticos, bem poucas vezes poderiam elas ser realizadas.

CONCORRA PARA O SILÊNCIO DO RECINTO EM QUE  
TRABALHA: O BARULHO E A CONVERSA A TODOS  
PREJUDICAM E MAIS AINDA AO SERVIÇO