

Maria Cecília Goi Porto Alves

Instituto de Saúde. Secretaria de Estado da Saúde. São Paulo, SP, Brasil

**Correspondência | Correspondence:**

Maria Cecília Goi Porto Alves  
Instituto de Saúde  
R. Santo Antônio, 590  
01314-000 São Paulo, SP, Brasil  
E-mail: cecilia@isaude.sp.gov.br

# Comentário: Criatividade na proposição de um plano de amostragem

## Creativity in proposing a sampling plan

Senti-me imensamente honrada pelo convite para comentar um trabalho da Prof<sup>a</sup> Eunice Pinho de Castro Silva, minha primeira referência em amostragem e por quem sempre tive profunda admiração. Conheci a Prof<sup>a</sup> Eunice quando, ainda estudante de Estatística, participei de um inquérito domiciliar sobre transtornos mentais no Município de São Paulo, executando, sob sua orientação, as tarefas de sorteio.

Formada em Matemática pela Universidade de São Paulo, especializou-se em amostragem na Universidade de Michigan, Estados Unidos, no *Summer Institute for Social Research Techniques*. Foi uma das primeiras brasileiras a participar do *Sampling Program for Foreign Statisticians*, formalizado por Leslie Kish em 1961. Esse programa tinha por objetivo formar “amostristas” e difundir o uso de amostragem, especialmente em países em desenvolvimento, e por onde, em 30 anos, passaram 400 estatísticos de 95 diferentes países (Kish,<sup>3</sup> 1996). Tornou-se o principal nome em amostragem na área de saúde pública, desempenhando as atividades de docente e pesquisadora do Departamento de Epidemiologia da Faculdade de Saúde Pública da USP, no período de 1960 a 1985. Nessa escola, a linha de pesquisa que instituiu e na qual criou os primeiros curso e publicação sobre amostragem na área de saúde pública vem sendo mantida até hoje, em seqüência ao seu trabalho pioneiro.

Algumas das qualidades do trabalho “Plano de amostragem utilizado no estudo de reprodução humana no Distrito de São Paulo” são características presentes em toda a produção da Prof<sup>a</sup> Eunice. Destacam-se o rigor e a competência em relação à aplicação da teoria de amostragem, refletidos na definição precisa e detalhada dos parâmetros a serem investigados e seus estimadores, na escolha adequada dos procedimentos de sorteio, coleta e estimação de erros de amostragem e na descrição sempre muito detalhada e

precisa desses procedimentos. Nada é subentendido, tudo é minuciosamente explicado e explicitado em expressões algébricas, claramente definidas e rigorosamente escritas.

No artigo, é notável a criatividade utilizada pela Prof<sup>a</sup> Eunice para contornar a ausência de um sistema de referência adequado. A existência de um cadastro da população de estudo, além de viabilizar a operacionalização do sorteio, é condição necessária para que um dos pressupostos da amostragem probabilística seja atendido: a de que a probabilidade de pertencer à amostra de todos os elementos da população seja conhecida.

O cadastro hoje utilizado na imensa maioria das pesquisas domiciliares realizadas no Brasil é a Malha de Setores Censitários, a base operacional dos censos. Esse sistema, embora existente desde a década de 50, não era facilmente acessado pelos pesquisadores à época da elaboração do trabalho. Os dados, processados em computadores de grande porte, eram disponibilizados a custos muito elevados e em prazos muito longos. A partir da década de 80, principalmente em função do desenvolvimento da tecnologia da informática, iniciativas adotadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) aumentaram o acesso às informações, a ponto de ser possível, atualmente, no Estado de São Paulo, consultar os mapas digitalizados dos setores censitários, sem custo algum, na sede regional desse Instituto.\* Assim sendo, a utilização dos setores censitários no planejamento de amostras tornou-se regra nos inquéritos domiciliares atuais.

Frente à escassez de dados para o planejamento da amostra, a Prof<sup>a</sup> Eunice propôs uma técnica alternativa para o sorteio das unidades de amostragem, exequível no prazo de tempo e orçamento disponíveis, e por meio da qual, mapas dos distritos do

\*Informações fornecidas pelos técnicos do Instituto Brasileiro de Geografia e Estatística, Antonio Luiz Carvalho Leme e Fanny Moore.

Município de São Paulo foram usados em substituição à relação de unidades da população de estudo. Nesse sentido, seu trabalho poderia ser considerado um precursor dos métodos simplificados de amostragem para levantamentos epidemiológicos, que surgiram na segunda metade da década de 60 do século passado na esteira da discussão sobre avaliação de cobertura vacinal do *Expanded Programme on Immunization* (EPI). Tais métodos justificavam-se pela necessidade de realização de estudos em contextos de escassez de recursos e de infra-estrutura para dimensionamento e execução de planos de amostragem probabilísticos.

No início da década de 80, a adoção de metodologias simplificadas de amostragem para levantamentos epidemiológicos em países em desenvolvimento passou a ser estimulada pela Organização Mundial de Saúde e o método do EPI tornou-se referência obrigatória nessa discussão. A crítica mais freqüente à sua proposta de amostragem refere-se à impossibilidade de se conhecer a probabilidade de seleção das crianças do estudo. Em artigo de 1986, a Prof<sup>a</sup> Eunice analisa de forma extremamente clara em que condições o método poderia gerar uma amostra probabilística. Com base nessa análise, propôs alterações considerando a realidade dos municípios brasileiros no que respeita a disponibilidade de dados para o sorteio e identificação de endereços da amostra (Silva,<sup>7</sup> 1986). A técnica criada pela Prof<sup>a</sup> Eunice, de sorteio por meio de mapas, não rompe com o princípio da equi-probabilidade da amostra e essa diferença em relação ao método do EPI merece ser destacada.

Outro elemento de planos simplificados de amostragem que pode ser identificado nesse artigo clássico da Revista, diz respeito à seleção no campo do domicílio a ser pesquisado. Em planos de amostragem que utilizam a metodologia padrão, de sorteio de conglomerados em vários estágios, está prevista a elaboração de uma lista de endereços, em ida ao campo anterior ao sorteio de domicílios. Evitar o retorno ao campo para a realização das entrevistas traz agilidade ao processo de amostragem, sendo um aspecto importante a ser considerado em desenhos simplificados. Essa possibilidade mostrou-se eficiente em planos simplificados, tendo sido testada sob diversas formas. Pela utilização de um roteiro que orienta o percurso dos entrevistadores que, ao mesmo tempo em que registram a localização do domicílio em formulário próprio, identificam aqueles sorteados para a composição da amostra e realização das entrevistas (Silva & Sundefeld,<sup>8</sup> 1999). Ou ainda, em função do sorteio de conglomerados em somente um estágio, tomando-se todas as unidades de amostragem neles contidas (Alves,<sup>1</sup> 2002).

É importante, ainda, salientar que os procedimentos propostos no trabalho clássico visavam impedir que o entrevistador exercesse julgamento pessoal no processo de seleção. A teoria estatística não pode ser aplicada a processos intencionais, que impedem a mensurabilidade da amostra. E retirando a possibilidade de dimensionamento dos erros de amostragem, fica inviabilizada a avaliação da qualidade das estimativas obtidas. A despeito da grande escassez de dados, a Prof<sup>a</sup> Eunice conseguiu arquitetar, de forma engenhosa, esquema que prescindia de sorteio no campo, mas não deixava margem para escolhas e decisões pessoais por parte do entrevistador, evitando, assim, a introdução de vícios.

Há no trabalho em avaliação um aspecto que, à luz dos conhecimentos hoje existentes, poderia ser mais simples do que o proposto. Refere-se ao sorteio de mulheres no segundo estágio de seleção. Embora o sorteio dentro do domicílio impeça a equi-probabilidade da amostra, impondo a introdução de pesos e conseqüente diminuição da precisão das estimativas, há situações em que o pesquisador deseja que somente um morador de cada domicílio seja incluído na amostra, quando há vários elegíveis.

O método probabilístico de referência para a seleção de respondentes nos domicílios é o desenvolvido por Kish<sup>2</sup> (1949). Assim como a alternativa proposta pela Prof<sup>a</sup> Eunice, requer a elaboração de uma listagem, ordenada por sexo e idade, de todos os moradores do domicílio e o uso de tabelas que orientam o sorteio. Trabalhos de pesquisa por telefone sugeriram que esse método causa desconforto às pessoas, resultando em informantes que desligam quando solicitados a especificar os moradores do domicílio, em especial, pessoas que não querem revelar a identidade de algum dos residentes ou que moram sozinhas.

Nesse contexto, diversos métodos alternativos foram desenvolvidos nos últimos 40 anos, entre eles o do último e o do próximo aniversariante, que prevêm a seleção da pessoa elegível que possui a data de aniversário mais próxima do dia da entrevista, antes ou depois dela, respectivamente. Baseiam-se na ausência de correlação entre o mês de aniversário e as características de interesse das pessoas elegíveis e a experiência tem mostrado que essa é uma pressuposição razoável (Salmon & Nichols,<sup>6</sup> 1983). Estudos indicam que a seleção baseada nas datas de aniversário gera amostras representativas, e em comparação ao método de Kish, produz amostras semelhantes em relação a características demográficas, tendo sido observadas diferenças muito pequenas referentes a raça e idade em estudos específicos (Oldendick et al,<sup>5</sup> 1988; O'Rourke & Blair,<sup>4</sup> 1983). Além disso, os

métodos do aniversariante levam à obtenção de taxas de resposta ligeiramente superiores, diminuem o tempo de entrevista e são preferidas pelos entrevistadores por serem mais fáceis de operacionalizar, dispensando o arrolamento dos nomes e das idades de todos os moradores e o manuseio de tabelas.

Retornando aos comentários sobre o artigo, quero finalizar considerando que a possibilidade da Prof<sup>a</sup> Eunice de criar alternativas a técnicas consagradas de sorteio originou-se no conhecimento com profundidade da teoria de amostragem. E foram alternativas engenhosas, propostas com criatividade e competência.

## REFERÊNCIAS

1. Alves MCGP. Simplificação do método de estimação da densidade larvária de *Aedes aegypti* no estado de São Paulo. *Rev Saúde Pública*. 2001;35:467-73.
2. Kish L. A procedure for objective respondent selection within the household. *J Am Stat Assoc*. 1949;44:380-7.
3. Kish L. Developing samplers for developing countries. *Int Stat Rev*. 1996;64:143-62.
4. O'Rourke D, Blair J. Improving random respondent selection in telephone surveys. *J Mark Res*. 1983;20:428-32.
5. Oldendick RW, Bishop GF, Sorenson SB, Tuchfarber AJ. A comparison of the Kish and last birthday methods of respondent selection in telephone surveys. *J Off Stat*. 1988;4:307-18.
6. Salmon CT, Nichols JS. The next-birthday method of respondent selection. *Public Opin Q*. 1983;47:270-6.
7. Silva EPCS. Métodos de amostragem para estimação da cobertura vacinal. *Rev Saúde Pública*. 1986;20:377-84.
8. Silva NN, Sundefeld MLM. Simplified sampling plans for epidemiological surveys in metropolitan areas in Brazil. *Surv Stat*. 1999;40:13-8.