

Uma alternativa metodológica para captação de quesitos estigmatizantes

Rebeca de Souza e Silva*

Antecedentes

Qualquer pesquisa que contemple a aplicação de questionários defronta-se com problemas de resistência por parte dos entrevistados, os quais via de regra, encontram-se estritamente relacionados com o caráter da questão formulada. Assim é que as perguntas inócuas tendem a receber boas respostas, enquanto que as que suscitam respostas comprometedoras ou altamente pessoais - por vergonha, medo ou desconfiança - tendem a ser evitadas.

Com o intuito de solucionar esse impasse, ou mais rigorosamente, com o objetivo de reduzir ou até mesmo eliminar o viés de respostas evasivas, Warner (1965) propôs um engenhoso procedimento de intervenção, que consiste, basicamente, em formular-se a questão de interesse com uma dada probabilidade "p" menor que a unidade. Trata-se, na verdade, de oferecer ao entrevistado a possibilidade de responder a uma de duas questões formuladas, sem que o entrevistador tenha conhecimento de qual das questões está sendo contestada.

Em essência, a utilização deste artifício possibilita que questões "sensíveis" sejam respondidas sem constrangimentos, posto que a privacidade do entrevistado com respeito à característica sensível é mantida. Tal procedimento foi denominado *Randomized Response Technique* e denotado, na literatura internacional, por RRT; em nosso meio foi traduzido por *Técnica de Resposta ao Azar* e, conseqüentemente, conhecido por TRA.

Assumindo-se, então, que o interesse de uma pesquisa reside em estimar a proporção de pessoas da população que apresenta uma determinada característica estigmatizante - uso de drogas, homossexualismo, provocação de um aborto, etc. - , tal como no procedimento convencional, extrai-se da população uma amostra representativa de "n" pessoas, tomando-se as devidas precauções para garantir que cada pessoa seja entrevistada. Para garantir que a pessoa coopere, isto é, que responda verdadeiramente, Warner idealizou a seguinte alternativa: para realizar-se a pesquisa, cada entrevistador é provido de um pião que assinala a letra "A" com probabilidade "p" e a letra "B"; com probabilidade (1-p), orientando os entrevistados que ao possuir o atributo de interesse ele deve identificar-se com a letra "A" e, em caso contrário, com a letra "B", em cada entrevista, pede-se ao entrevistado que, sem ser observado pelo entrevistador, rode o pião e declare se este assinala ou não o grupo ao qual pertence. Salienta-se que o entrevistado não declara ao entrevistador a letra indicada pelo pião, apenas contesta "sim" ou "não".

Cada resposta "sim" isolada não tem utilidade alguma, contudo, no total, permite estimar a verdadeira proporção - W - de pessoas que possuem a característica "A" na população. Com efeito, denotando-se por K a possibilidade de ocorrência de respostas afirmativas, é fácil equacionar que:

$$K = Wp + (1 - W)(1 - p) \quad (1)$$

O valor de "p" é fixado, portanto, conhecido. A proporção de pessoas da amostra que respondem "sim" constitui-se em uma estimativa de K. Conseqüentemente, percebe-se que o único termo desconhecido da equação (1) é W. Pode-se demonstrar que o valor de W, obtido com a resolução desta equação, constitui-se em uma estima-

* Doutoranda da Faculdade de Saúde Pública/USP.

tiva de máxima verossimilhança da verdadeira proporção. Assim sendo, para $p \neq 0,5$, tem-se:

$$\hat{W} = \frac{p - 1}{2p - 1} + \frac{K}{2p - 1} \quad (2)$$

com

$$E(\hat{W}) = W \quad (3)$$

e

$$\begin{aligned} \text{VAR}(\hat{W}) = & \frac{W(1 - W)}{n} + \\ & + \frac{p(1 - p)}{n(2p - 1)^2} \end{aligned} \quad (4)$$

Desde que \hat{W} é a máxima probabilidade estimada de um estimador não viciado, pode-se considerá-lo como sendo normalmente distribuído ao redor de W . Assim o intervalo de confiança usual pode ser facilmente estabelecido.

Com respeito à fixação do valor de "p", deve-se ressaltar que esta assume importância fundamental no emprego da TRA. Como aquela probabilidade associa-se à questão de interesse, ao pesquisador revela-se conveniente que assuma um valor tão próximo de 1 quanto possível; todavia, ao entrevistado pede-se menos quanto maior for sua chance de selecionar a letra "B", isto é, quanto menor for o valor de "p". Nesse sentido, "p" pode ser pensado como descrevendo a natureza da cooperação entre pesquisador e entrevistado (Warner, 1965).

Por outra parte, a questão do tamanho amostral requerido para um dado nível de precisão, também depende do parâmetro "p". Com efeito, a expressão (4) permite visualizar que quanto mais próximo do valor unitário estiver "p", menor será o tamanho amostral requerido.

O interesse por tal procedimento de intervenção direcionou alguns pesquisadores a buscarem alternativas que viessem a

aprimorá-lo. Todavia, um balanço realizado nos trabalhos sobre aborto provocado, que propõem, ou até mesmo empregam, os diversos tipos de jogos - TRA - desenvolvidos a partir da idéia de Warner, permitiu constatar que aquelas possuem com respeito a essa o inconveniente de requererem o conhecimento de técnicas estatísticas complexas, o que acabam por acarretar maior dificuldade no entendimento do "jogo".

Acreditando-se, contudo, na vantagem potencial da TRA, e buscando minimizar as dificuldades que permeiam seu uso apresenta-se, a seguir, uma proposta alternativa que visa facilitar e estimular o emprego desta técnica.

Proposta alternativa

Atendo-se à expressão (4), observa-se que a variância estimada derivada da TRA pode ser escrita como sendo a soma da variância devida ao procedimento binomial mais a da devida ao artifício de aleatorização. Este componente, portanto, pode ser interpretado como sendo o custo que se paga pela introdução do artifício. Em definitivo, é por conta deste custo que o tamanho amostral requerido para se utilizar a TRA revela-se alto.

Por outra parte, em decorrência, pode-se demonstrar que sob a hipótese das mulheres não oferecerem resistência em responder à questão sensível diretamente, o modelo convencional - binomial - revela-se mais eficiente que a TRA. Greenberg (1969), apresenta uma simulação que confirma esse fato.

Estes argumentos, aliados às dificuldades que permeiam a utilização da TRA, sugerem uma nova alternativa de empregar-se esta técnica. Esta alternativa consiste, simplesmente, em requerer que a entrevistada responda reservadamente à questão sensível. Em termos práticos, significa atribuir a "p" o seu valor máximo. Para tanto, pode-se idealizar a seguinte situação: a

cada entrevista, dois cartões são entregues à entrevistada, em um deles imprime-se a palavra "sim" e estampa-se um desenho alusivo à questão sensível; no outro, imprime-se a palavra "não" e estampa-se um desenho que de maneira alguma deve aludir à questão de interesse. Esclarecendo-se, então a entrevistada sobre a importância de sua cooperação, pede-se que deposite o cartão referente à sua situação em um envelope, sem ser observada pelo entrevistador, e o deposite lacrado em uma envelope transparente, preferencialmente, vedada.

Para não despertar a desconfiança das primeiras mulheres entrevistadas, sugere-se que cada entrevistador deposite uma quantidade pré-estabelecida de envelopes na urna antes do início das entrevistas.

Outra questão de fundamental importância reside em identificar-se cada resposta ao perfil característico da respondente. Nesse sentido, sugere-se que o questionário seja impresso em uma única folha, contendo o menor número possível de informações, preferencialmente pré-codificadas. Sugere-se, ainda, que o nome da entrevistada seja omitido do rol de informações, com a finalidade de garantir-se uma maior privacidade. Uma vez recolhidas as informações, entrega-se a folha de questões à entrevistada, pedindo-lhe que a envelope juntamente com o cartão referente à sua situação. Esta atitude, além de necessária, assegurará à entrevistada que não há "trapaça" nesta forma de se recolher a informação sensível.

Vale a pena salientar que, tal como foi preconizado pelo grupo de pesquisadores da Universidade John Hopkins, o fato da entrevistada sequer verbalizar um "sim" torna a resposta requerida menos comprometedora. Mais ainda, acredita-se que há vantagens na utilização desta alternativa em populações com baixa instrução, posto que por um lado as entrevistadas deverão identificar-se simplesmente com um desenho e que, por outro, estão familiarizadas com o processo de votação.

Além destas vantagens, esta alternativa vislumbra as seguintes vantagens:

- eliminar o custo pago pela introdução do artifício de aleatorização;
- requerer, conseqüentemente, um tamanho amostral equivalente ao requerido pelo emprego do modelo binomial;
- maior facilidade de compreensão por parte da entrevistada;
- estimular a utilização da técnica, inclusive, por pesquisadores leigos em estatística;
- possibilitar o emprego da técnica em pesquisas nacionais, como o Censo, por exemplo.

Matematicamente, admitindo-se que todas as pessoas de uma população possuem ou não a característica de interesse e que deverá ser estimado por pesquisa a proporção da população que possui esta característica, estabelece-se que:

$$W = \begin{cases} \text{a verdadeira probabilidade da característica de interesse;} \\ \end{cases}$$

$$X_i = \begin{cases} 1 \text{ se o } i\text{-ésimo elemento da amostra diz "sim"} \\ 0 \text{ se o } i\text{-ésimo elemento da amostra diz "não"} \end{cases}$$

Então:

$$P(X_i = 1) = W$$

$$P(X_i = 0) = (1 - W)$$

Supondo que de uma amostra de tamanho n - extraída de forma a garantir sua representatividade -, n_1 elementos declarem "sim", e conseqüentemente, $(n - n_1)$ declarem "não", a probabilidade da amostra, ou seja, a função de verossimilhança, é dada por:

$$L = W^{n_1} (1 - W)^{(n - n_1)}$$

O logaritmo desta função, dado por:

$$\log L = n_1 \log W + (n - n_1) \log(1 - W)$$

Verifica-se que:

$$W = \frac{n1}{n}$$

Como todas as variáveis X_1, X_2, \dots, X_n , assumem os mesmos valores - 0 ou 1-, com as mesmas probabilidades, isto é, $P(X_i = 1) = W$ e $P(X_i = 0) = 1 - W$, tem-se que:

$$E(X_i) = 1 \times W + 0 \times (1 - W) = W$$

Donde depreende-se que:

$$E(\hat{W}) = W$$

$$\text{VAR}(\hat{W}) = \frac{W(1 - W)}{n}$$

Percebe-se, portanto, que \hat{W} é um estimador de máxima verossimilhança de W . Além disso, constata-se que a alternativa proposta, em termos matemáticos, equivale-se ao procedimento binomial.

Considerações finais

A autora teve a oportunidade, com o apoio financeiro da ABEP, de utilizar uma modalidade de TRA na captação da informação sobre aborto provocado no ano de 1987,

dentre 996 mulheres em idade fértil, residentes no sub-distrito de Vila Madalena, no município de São Paulo. Mais que isso, pôde contrastar a estimativa obtida por esta via com a obtida mediante levantamento daquela informação nos moldes convencionais. Como se esperava, os resultados foram animadores. Enquanto a abordagem direta permitiu observar que apenas 8 em 1000 mulheres recorreram à provocação de um aborto no ano de 1987, o uso da TRA diagnosticou que 42 mulheres em 1000, utilizaram-se de tal prática naquele período.

Nessa perspectiva, pretende-se num futuro próximo, tornar a se investir nesse tipo de metodologia para captação de questões sensíveis. O primeiro passo a ser trilhado, seguramente, deverá ser o de validar a técnica aqui proposta, concomitantemente à validação de uma das modalidades de TRA proposta por outros autores, com intuito de se eleger a mais apropriada. Para se proceder à validação tratar-se-á de aplicar a técnica a um grupo de mulheres cujas histórias reprodutivas relativas, especialmente, ao aborto provocado, sejam conhecidas a priori. Aliás, deve-se dizer de passagem, que o procedimento metodológico referente a tal passo, constitui-se em um dos objetivos de um projeto elaborado pela autora, destinado à compreensão das relações existentes entre aborto provocado, fecundidade e anti-concepção, enviado recentemente, à Fundação MacArthur.

Referências bibliográficas

- GREENBERG, B.G et alii - 1969. The unrelated question randomized response model: theoretical framework. *Journal of the American Statistical Association*, Washington, 64: 520-539.
- WARNER, S.L. - 1965. Randomized Response: a survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, Washington, 60: 63-69.

Recebido para publicação em 28/09/90.
Aprovado para publicação em 11/04/91.