

A ORGANIZAÇÃO DE UM PLANO AMOSTRAL

Luiz Caetano Grecco Teixeira
Centro de Estudos Superiores de Londrina

INTRODUÇÃO

Nestes anos em que tenho sido o responsável pelas disciplinas de Estatística no CESULON venho percebendo que os estudantes mostram certas dificuldades em aplicar os conhecimentos adquiridos na disciplina em situações concretas de pesquisa acadêmica. A análise detalhada de tais dificuldades demonstra que elas não se referem aos conceitos teóricos, mas principalmente se concentram no planejamento da amostra que será utilizada.

Também, e com grande frequência, sou procurado por estudantes que estão iniciando a pós-graduação, que me apresentam "problemas" decorrentes do mal planejamento da amostra, e que só se fazem sentir por ocasião da tabulação ou da análise dos resultados. Geralmente, nesse momento, pouco se pode fazer para salvar o imenso trabalho realizado na coleta dos dados...

A grande maioria dos trabalhos de pesquisa na graduação e no início da pós-graduação se referem a levantamentos simples, onde os dados são colhidos através de amostra simples ao acaso. O problema enfrentado pela maioria dos estudantes não é o tratamento estatístico dos dados, mas se resume em como planejar uma amostra de forma que os dados coletados possam ser efetivamente utilizados em função dos objetivos da pesquisa, sem os riscos de se coletarem dados desnecessários ou de, no momento da análise, perceber-se que os dados colhidos são incompletos.

O objetivo deste trabalho é fornecer um roteiro de itens que devem ser definidos pelo investigador no momento em que planeja a pesquisa. O roteiro que será apresentado é aplicável em levantamentos simples cujos objetivos sejam meramente descritivos, sendo os dados obtidos por amostragem probabilística.

Na elaboração do texto supomos que alguns conceitos da Teoria da Amostragem e da Estimação sejam do conhecimento do leitor; todavia, ao final sugerimos algumas leituras complementares que ajudarão a elucidar os conceitos envolvidos.

Roteiro para a elaboração de um Plano Amostral

1. Definição do Problema.

Ao contrário do que muitos pensam, um pesquisador, antes de se lançar efetivamente na coleta de dados, deve ter conhecimento amplo do problema que vai pesquisar e da população objeto da pesquisa. Este conhecimento é adquirido através daquilo que normalmente denomina-se revisão de literatura, a qual fornecerá os subsídios para responder às seguintes questões:

1.1. Qual o universo que será investigado?

Significa responder "o que?", "quem?", "quando?" e "onde?" será pesquisado. Não só delimita o problema, como também determina claramente a população de onde será coletada a amostra. A definição clara da população permitirá a adequada seleção da amostra e dará o contexto exato a que se referem as conclusões.

1.2. Quais os conceitos e características que serão levantadas?

Com base na revisão de literatura, o pesquisador deve deixar bem claro os conceitos e as características do que será pesquisado. Isso é importante para delimitar o que se busca na amostra e facilitará a definição da unidade de resposta (vide adiante). Por exemplo, ao se pesquisar sobre "aluguéis", é necessário caracterizar o que é um imóvel alugado para efeito da pesquisa (supõe ou não a existência de um contrato explícito), se refere a aluguel comercial ou residencial (e definir as características de cada um), etc. Da mesma forma, uma pesquisa sobre hábitos alimentares de uma população, ao medir o consumo de alimentos gordurosos, deve conceituar o que é um "alimento gorduroso", de forma que o entrevistado (e, posteriormente, o leitor) saiba decidir se "fei-

jão" é ou não gorduroso.

1.3. Que parâmetros serão estimados?

Determinar com clareza quais as medidas da população que se deseja estimar: se a média populacional, se a proporção, etc. Isso condicionará algumas perguntas a serem incluídas no instrumento de pesquisa (questionário ou roteiro de entrevista). Também deve o investigador ter claro se tais parâmetros desejados podem ser efetivamente estimados a partir das estatísticas usuais ou se será necessário o desenvolvimento de estimadores específicos (veja 3.1 e 3.2 adiante).

1.4. Serão produzidos estimadores para partes específicas da população (estratos)?

A resposta a este item determinará se o plano amostral deverá ou não prever estratos na amostra. No item 2.1, adiante, o conceito de estrato será melhor tratado.

1.5. Qual o sistema de referência a ser utilizado?

Aqui trata-se de definir o sistema de referência que será utilizado para o sorteio da amostra. Por exemplo: é disponível uma listagem de toda a população? há possibilidade de uma listagem que permita o uso de conglomerados? há um arquivo disponível que contenha todos os elementos da população? etc. Não existindo, o pesquisador deverá criar o sistema de referência e adotar um esquema amostral compatível. Em último caso, o pesquisador pode decidir por uma pesquisa do tipo "gall-up", apesar dos inconvenientes que tal amostragem apresenta na produção de estimativas.

Um sistema de referência muito utilizado em pesquisas urbanas é a planta da cidade, tomando-se os quarteirões como base do conglomerado, ou mesmo fazendo amostragem sistemática dos domicílios através de roteiros aleatórios.

2. Construção da amostra.

Então, de posse das informações que definem o problema e a população, o investigador passa ao planejamento da amostra propriamente dita, através dos seguintes passos:

2.1. Serão utilizados estratos? quais? quantos? por que?

Aquí compete ao investigador decidir se o problema será estudado separadamente nos diversos setores que compõe a população. Tais setores devem ser definidos a priori, com base no conhecimento prévio que se tenha da população e de suas características, e constituirão os estratos da amostra, para cada um dos quais serão produzidas as estimativas.

É importante não confundir estrato com sub-população. Os estratos são definidos previamente, a partir das características da população; eles já aparecem no sistema de referência. Quando o investigador separa os "estratos" com base nos dados colhidos, na verdade ele separa amostras das sub-populações, pois não há garantias de que todos os estratos existentes foram amostrados. Na amostra estratificada, o pesquisador tem o controle do tamanho da amostra em cada estrato, ao passo que na sub-população o tamanho da amostra em cada parte é aleatório, isto é, não controlado pelo pesquisador. Essa distinção é importante para o tratamento estatístico dos dados obtidos e para a análise final dos resultados. No caso de estratificação, cada estrato é tratado em separado e produz estimativas específicas; no caso de sub-população, os dados devem ser tratados em conjunto.

2.2. Qual a unidade amostral?

A unidade amostral é exatamente o que será sorteado, isto é, a base do conglomerado se este esquema for o utilizado. Por exemplo, uma pesquisa deseja saber a opinião da dona-de-casa sobre determinados produtos; porém, para se chegar à dona-de-casa, se sorteia a residência, que é a unidade amostral. Já a dona-de-casa se constitui na respondente, que deve ser bem caracterizada como unidade de resposta.

2.3. Qual a unidade de resposta?

Não se trata do respondente, mas a definição das características que deve ter o respondente. Assim, no exemplo acima, se deve ter bem claro qual o conceito de "dona-de-casa" a ser considerado: se é a mãe de família, ou se é a filha mais velha já que esta cuida da decisão sobre que produtos comprar, etc.. De tal definição depende a escolha exata

do respondente, possibilitando consistência entre todos os elementos que constituem a amostra, permitindo-se assim eliminar da amostra as unidades sorteadas cujas características sejam diferentes das definidas (por exemplo, sendo sorteada uma pensão, a dona da pensão não é, necessariamente, "dona-de-casa", mas uma comerciante...). Note-se bem que esta definição depende diretamente das decisões tomadas em 1.1. e 1.2.

2.4. Em quantos estágios e quais as unidades em cada um?

Aqui devem ser caracterizados os estágios de sorteio, caso se opte por um esquema de estágios múltiplos. Por exemplo, sorteia-se os quarteirões a partir do sistema de referência criado através da planta da cidade (1º estágio); em cada quarteirão sorteado, sorteia-se oito residências (2º estágio); em cada residência sorteada se entrevista a "dona-de-casa" (3º estágio).

2.5. Qual o tamanho da amostra?

O tamanho da amostra é diretamente proporcional ao custo da pesquisa. Não depende do tamanho da população! mas dos recursos disponíveis e da confiabilidade desejada. Todavia, um bom estatístico, a partir de um levantamento piloto e com base na dispersão obtida poderá determinar o tamanho ideal para se obter a confiabilidade desejada.

2.6. Qual a confiabilidade desejada?

Deve-se definir a priori qual o nível de significância desejado para os intervalos de confiança das estimativas e para os testes de hipótese. É importante que seja definido previamente, antes de se dispor dos dados para se evitar manipulações nos testes.

2.7. Qual o custo da pesquisa?

Estes três últimos itens estão correlacionados (tamanho da amostra, confiabilidade e custo). O pesquisador deve ter uma boa estimativa do custo total da pesquisa e deve conhecer os componentes desse custo, para evitar que a pesquisa seja interrompida por "falta de verba". Com base na confiabi-

lidade e nos recursos disponíveis, se poderá determinar o tamanho da amostra. Muitas vezes a solução é equacionar os três itens de forma a otimizá-los adequadamente. Dentre os componentes do custo, destacam-se: a remuneração dos entrevistadores e dos tabuladores, o material a ser utilizado, despesas de locomoção do pessoal envolvido, custos de impressão e duplicação do instrumento, etc.

3. Escolha dos estimadores.

3.1. Quais estimadores serão utilizados?

Os estimadores nem sempre são as estatísticas usualmente calculadas. Muitos problemas exigem que o pesquisador desenvolva estimadores específicos, os quais deverão ter as qualidades inerentes a um bom estimador: precisão, acurácia e não-viesado. Geralmente, este é um trabalho próprio de um estatístico que deverá estar assessorando o investigador principal. Todavia, para os parâmetros usuais, os compêndios mais elementares de Amostragem e Estimação apresentam os estimadores adequados.

3.2. Existem variáveis auxiliares à disposição?

A existência dessas variáveis é de grande auxílio na definição de estimadores específicos. Devem ser buscadas na revisão de literatura, especialmente em outras pesquisas sobre o mesmo assunto ou sobre a mesma população, bem como em registros oficiais como o Censo (IBGE).

Note bem que estes dois itens estão intimamente relacionados com o que vimos em 1.3 e 1.4.

4. Alguns cuidados adicionais na coleta e tabulação dos dados.

4.1. Decidir se a coleta é por método direto ou in direto.

4.2. Construção e testagem do instrumento de coleta.

4.3. Construção e testagem dos instrumentos de registro da informação (mapas de tabulação e de cálculos).

4.4. Seleção e treinamento dos entrevistadores.

Este é um item da maior importância, pois se deve garantir que todos os entrevistadores atuarão de maneira o mais uniforme possível reduzindo-se assim o risco da introdução de uma variável não controlada que se denomina "viés de entrevistador".

4.5. Decidir como tratar ausentes e não respondentes.

Muitas vezes a unidade sorteada não é encontrada ou se recusa a colaborar. Assim, é necessário um procedimento padrão para se tratar estes casos. Só em última circunstância se admitirá a substituição de uma unidade sorteada, e quando isso acontecer sempre será usado o mesmo critério de escolha da substituta.

4.6. Codificação de respostas.

Um sistema de códigos bem elaborado tornará mais eficiente e produtivo o trabalho de tabulação e processamento dos dados obtidos.

4.7. Exame de qualidade dos resultados.

É importante que se verifique a consistência das informações obtidas, cruzando-se respostas, comparando-se os totais de respostas, etc. Sempre que possível, cotejar os resultados com dados extraídos do Censo (IBGE), ou de outras fontes, seja sobre a mesma população, seja sobre o mesmo problema em outra população; as diferenças devem ser adequadamente justificadas ou explicadas a partir do contexto da pesquisa e não do método amostral utilizado.

4.8. Composição do relatório.

Deve-se prever o nível de detalhe requerido pelo relatório final para evitar-se trabalho inútil; deve haver consistência com os objetivos da pesquisa e com os itens da

definição do problema.

CONCLUSÃO:

Esta lista de ações não esgota o tema. Um pesquisador experiente poderá ampliar ou reduzir os quesitos aqui mencionados em função de seus objetivos. Todavia, é sempre bom que o pesquisador elabore seu plano amostral com o auxílio e assessoramento de um estatístico, pois ele poderá apresentar sugestões que facilitarão todo o trabalho.

O estatístico deve participar da pesquisa desde o seu planejamento, para garantir a qualidade dos resultados. Muitas vezes, o estatístico se vê em situações difíceis justamente quando chamado a colaborar em uma pesquisa após a coleta de dados e percebe que o pesquisador não trabalhou dirigido por um plano amostral e cometeu erros de coleta e de amostragem que comprometem definitivamente os resultados. O estatístico não pode "ajeitar as coisas" em tal situação, pois além de ser anti-ético, ele estaria ferindo os procedimentos técnicos de sua área. A Estatística não é um conhecimento falacioso, parte do reino da "magia" ou da manipulação de informação; pelo contrário, é um método científico rigoroso e que deve ser aplicado dentro de critérios metodológicos próprios.

Por outro lado, um plano amostral bem elaborado, permitirá a um estatístico se envolver na pesquisa em qualquer momento, pois poderá tomar conhecimento de todas as ações já desenvolvidas e a desenvolver.

Finalmente, não se deve confundir o roteiro aqui estabelecido com a metodologia da pesquisa. O plano amostral é apenas um item do planejamento da pesquisa, e é desenvolvido tendo em vista os objetivos "estatísticos" da pesquisa. Se determinado problema não vai ser abordado estatisticamente, o plano amostral é dispensável; porém, se o tratamento estatístico é necessário, só poderá apresentar resultados seguros mediante o desenvolvimento de um plano amostral.

Leituras Complementares

1. BLALOCK, H.M., Junior-Introdução à Pesquisa Social. 2ª ed. Rio de Janeiro, Zahar Ed. 1976 (especialmente o capítulo 3)
2. BRUYNE, Paul et alii - Dinâmica da Pesquisa em Ciências Sociais: os pólos da prática metodológica. Rio de Janeiro, Francisco Alves ed., 1977. (especialmente o capítulo 7)
3. GATTI, Bernadete A. et FERES, Nagib L. - Estatística Básica para Ciências Humanas. 3ª ed. rev. e aum. São Paulo, Alfa-Omega, 1978.
4. GOODE, William J. et HATT, Paul K. - Métodos em Pesquisa Social. 7ª ed. São Paulo. Ed. Nacional, 1979.
5. LEVIN, Jack - Estatística Aplicada a Ciências Humanas. São Paulo, Ed. Harbra, (1978).
6. UNESCO - Manual de Pesquisa Social nas Zonas Urbanas. São Paulo, Liv. Pioneira ed., (1978). (especialmente a primeira parte)
7. WONNACOTT, Thomas H. et WONNACOTT, Ronald J. - Introdução à Estatística. Rio de Janeiro, Livros Técnicos e Científicos, 1980.

