



## BREVE ESTUDO SOBRE AMOSTRAGEM E TAMANHO MÍNIMO DA AMOSTRA

Adelia Onay Ribeiro de Mello

### Resumo

Este artigo trata de uma fase no processo da pesquisa que é a determinação do tamanho da amostra. Para isso foi feita uma explanação do método estatístico visando um esclarecimento do uso da ferramenta estatística envolvida no processo da escolha dos elementos que comporão a amostra e a metodologia a ser utilizada.

### Palavras-chave:

Estatística descritiva; Amostra; Parâmetro; Probabilística; Inferência

### 1 Introdução

É comum no processo de pesquisa surgirem dúvidas envolvendo as ferramentas estatísticas. A maior hesitação ocorre na determinação do tamanho mínimo necessário para a amostragem e qual técnica a ser utilizada.

Esse estudo visa apresentar uma breve reflexão das etapas do método estatístico, amostragem e determinação do tamanho da amostra, buscando o esclarecimento ao leitor no momento em que utilizar esta ferramenta.

### 2 Desenvolvimento

#### 2.1 Conhecendo as fases do método estatístico:

- a) Definição do problema: Consiste em uma definição correta do problema a ser investigado. Saber exatamente o que se pretende pesquisar. Responde a pergunta “o que?”;
- b) Planejamento: Consiste em determinar o procedimento necessário para a solução da investigação. Como levantar informações sobre o assunto propriamente dito. Quais dados deverão ser obtidos? Como se deve obtê-los?;
- c) Coleta de dados: A coleta pode ser direta ou indireta. A coleta é direta quando é feita sobre elementos informativos de registro obrigatório

(nascimento, casamentos e óbitos, importação e exportação de mercadorias, entre outros), ou ainda, quando os dados são coletados pelo próprio pesquisador através de inquéritos ou questionários;

- d) A coleta é indireta quando inferida a partir dos elementos obtidos pela coleta direta ou através de conhecimento de outros fenômenos que, de algum modo, estejam relacionados com o fenômeno em questão;
- e) Apuração dos dados: É um tratamento dos dados antes da análise. Consiste em resumir os dados através de sua contagem e agrupamento;
- f) Apresentação dos dados: Usualmente feita através de gráficos e tabelas. Embora a apresentação tabular (tabelas) seja importante para facilitar a análise numérica dos dados, ela não permite ao analista obter uma visão rápida e fácil do fenômeno, e de sua variação como a conseguida através de um gráfico;
- g) Análise e interpretação dos dados: A análise dos dados estatísticos está essencialmente ligada ao cálculo de medidas, cuja finalidade principal é descrever o fenômeno (Estatística Descritiva). Feita a partir do conjunto de dados a serem analisados e pode ser expresso por números-resumos (estatísticas) que evidenciam características particulares deste conjunto (amostra).

Tem-se que a análise da totalidade do universo estatístico ou população pode se tornar cara e enfadonha. É comum e viável o estudo de uma parte da população, denominada amostra. A partir dos resultados nesta amostra serão obtidos valores aproximados ou estimativos para as características populacionais de interesse. Este tipo de pesquisa é chamada de *levantamento de amostragem*.

Definida a população, é preciso estabelecer a técnica de amostragem, ou seja, a escolha do procedimento que será adotado para a seleção dos elementos que compõem a amostra. Segundo ANDERSON (2007), características numéricas de uma população (como a média e o desvio padrão) são chamadas de *parâmetros*, geralmente desconhecidos, que serão estimados através da amostra. Esses resultados são estimativas numéricas dos valores populacionais.

Por fim, indicam-se métodos de amostragem os possibilitados pela amostragem probabilística, onde cada elemento da população possui determinada probabilidade de ser selecionado e a variabilidade dos resultados da amostra é obtida com precisão, sendo possível o cálculo do erro padrão. Ainda, é possível utilizar-se da amostragem não probabilística, onde não é possível calcular o erro padrão, não havendo garantia da representatividade da população.

## 2.2 Técnicas de amostragem probabilísticas

A seguir demonstraremos os tipos de amostragem:

- a) **Amostra aleatória simples:** Para este tipo de amostragem é necessário saber, *a priori*, o número total de elementos de uma população e consiste em selecionar a amostra através de um sorteio sem reposição. Na amostragem aleatória simples, qualquer elemento da população tem a mesma probabilidade de fazer parte da amostra.
- b) **Amostra Sistemática:** Quando os elementos da população já se encontram ordenados (prontuários médicos, os prédios de uma rua, as linhas de produção, etc), a seleção dos elementos que constituirão a amostra pode é feita por um sistema imposto pelo pesquisador e não ao acaso.
- c) **Amostra Estratificada:** Esta técnica consiste em dividir a população em subgrupos denominados *estratos*. Estes subgrupos devem ter características em comum (mais homogêneos que a população total), no que tange a variável em estudo.
- d) **Amostragem estratificada proporcional:** Caso particular da amostragem estratificada. A proporção do tamanho da amostra é mantida no estrato (estrato corresponde a 20% do tamanho da população, ela deve corresponder a 20% do tamanho da amostra).
- e) **Amostra por Conglomerado:** Nessa técnica, dividimos a área da população em seções (conglomerados); escolhemos algumas dessas seções e todos os elementos das seções escolhidas serão selecionados. Uma

diferença importante entre a amostragem por conglomerado e amostragem estratificada é que a amostragem por conglomerado utiliza todos os elementos dos conglomerados selecionados, enquanto a amostragem estratificada utiliza uma amostra de membros de cada estrato.

### 2.3 Técnicas de amostragem não probabilísticas

- a) **Amostragem por conveniência**: A amostra de conveniência é formada por elementos que o pesquisador reuniu simplesmente porque dispunha deles. Geralmente utilizada em pesquisa de opinião;
- b) **Amostragem por julgamento**: Os elementos escolhidos são aqueles julgados como típicos da população que se deseja estudar;
- c) **Amostragem por quotas**: Assemelha-se com a amostragem estratificada proporcional, onde a população é dividida em diversos subgrupos.

### 2.4 Inferência Estatística

Realizado o processo de amostragem, os resultados obtidos podem ser generalizados para a população. Estes estudos são chamados *indutivos ou inferenciais*. A estimativa da amostra pode ser um indicativo do parâmetro populacional, considerando a existência de um erro amostral.

- **Quanto maior a dispersão** maior o erro padrão
- **Quanto maior n, menor o erro padrão**
- **Nível de confiança** da estimativa  $(1 - \alpha)$ , é o grau de certeza dos resultados. (mais utilizados: 90%, 95%, 99%).
- **Nível de significância**  $(\alpha)$ , é a probabilidade de erro.
- **VALOR CRÍTICO**  $\left( Z_{\alpha/2}, \text{ou}, t_{\alpha/2(n-1)} \right)$ : É o número nas fronteiras que separa os valores estatísticos amostrais prováveis de ocorrerem. (tabelados)

**MARGEM DE ERRO OU ERRO MÁXIMO DA ESTIMATIVA:** A margem de erro  $E$  é chamado erro máximo da estimativa, e pode ser obtida multiplicando o valor crítico pelo erro padrão das médias.  $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$  ou  $E = t_{\alpha/2, (n-1)} \cdot \frac{S}{\sqrt{n}}$

O tamanho da amostra é um passo importante no processo da inferência estatística. O tamanho da amostra depende do grau de confiança desejado, da dispersão entre os valores individuais e do erro tolerável do processo, e também depende de o desvio padrão populacional ser conhecido ou não.

Após a análise do questionário, escolha a(s) variável que represente o objetivo da pesquisa. Segundo Fonseca & Martins (2012), o procedimento para o cálculo do tamanho da amostra deve seguir as etapas:

**Verificar a natureza da variável** (quantitativa ou qualitativa), e se a é população (finita ou infinita).

Para a variável intervalar e população finita:  $n = \left( Z \frac{\sigma}{E} \right)^2$ , onde:

$Z$  = valor crítico da curva normal padrão.

$\sigma$  = desvio padrão da população (especificações técnicas, valor de estudos similares, ou estimativas )

$E$  = erro amostral. É máxima diferença que o pesquisador admite entre  $\mu$  (média da população) e  $\bar{x}$  (estimativa amostral).

**Se a variável for quantitativa e a população for finita (conhecemos o total da população):**

$$n = \frac{Z^2 \cdot \sigma^2 \cdot N}{Z^2 \cdot \sigma^2 + E^2 \cdot (N-1)}$$

**Se a variável for qualitativa:**

População infinita  $\rightarrow n = \frac{Z^2 \times p \times q}{E^2}$

Onde  $p$  = a estimativa da verdadeira proporção da variável escolhida;

$q$  = o complemento de  $p \rightarrow (1-p) \rightarrow p+q=1$  ou 100%.

População finita  $\rightarrow n = \frac{Z^2 \cdot p \cdot q \cdot N}{Z^2 \cdot p \cdot q + E^2 \cdot (N-1)}$

Exemplos:

- 1- Suponha que a variável escolhida seja o peso de certa peça produzida e que não conheço o tamanho da população. Pelas especificações do produto o desvio padrão é de 10 kg, e admitindo um nível de confiança  $(1-\alpha)$  de 95% e um erro amostral de 1,5 kg. Qual o tamanho mínimo da amostra a ser levantada?

$$n = \left( Z \frac{\sigma}{E} \right)^2 \quad \rightarrow \quad n = \left( \frac{1,96 \times 10}{1,5} \right)^2 = 170,74 \approx 171$$

- 2- Considere o exemplo anterior e que a população seja finita de 500 peças.

Logo:

$$n = \frac{Z^2 \cdot p \cdot q \cdot N}{Z^2 \cdot p \cdot q + E^2 \cdot (N-1)} \quad \rightarrow \quad n = \frac{1,96^2 \cdot 10^2 \cdot 500}{1,96^2 \cdot 10^2 + 1,5^2 \cdot (500-1)} = 127,47 \approx 128$$

- 3- Um pesquisador precisa determinar o tamanho da amostra para estimar a verdadeira percentagem populacional com um erro máximo igual a 5% e utilizando um nível de confiança de 95%. Suponha-se um valor de  $p$  igual ou menor que 0,30.

$$n = \frac{Z^2 \cdot p \cdot q}{E^2} \quad \rightarrow \quad n = \frac{1,96^2 \times 0,30 \times 0,70}{0,05^2} = 322,6944 \approx 323$$

### Considerações Finais

No término do artigo é esperado que as dúvidas existentes em relação a utilização das técnicas de amostragem e determinação do tamanho mínimo da amostra sejam elucidadas, e que facilite ao leitor, o processo de pesquisa e uso da ferramenta estatística.

## REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSON, D. R.; SWEENEY D. J., WILLIAMS, T. A. **Estatística Aplicada a Administração e Economia**. São Paulo: Thomson Learning, 2007, 2. ed.

CLARK, J.; DOUGLAS, D. **Estatística Aplicada**. São Paulo: Saraiva, 1998.

FONSECA, J.S. ,MARTINS,G.A. **Curso de estatística**. São Paulo: Atlas, 2012. 6. ed.

FREUND, J. E. SIMON, G. A.. **Estatística Aplicada**. Porto Alegre: Bookman, 2000. 9 ed.

MORETTIN, L. G **Estatística Básica, Probabilidade**. São Paulo: Mc Graw Hill. s. a. 6. Ed.

TRIOLA,M.F. **Introdução à Estatística**. Rio de janeiro: LTC, 1999. 7. ed.