

1.3 POPULAÇÃO E AMOSTRA

Na linguagem comum, população significa o conjunto de habitantes de um território, de um país, de uma região, de uma cidade etc. Em Estatística, no entanto, o termo *população* tem significado bem mais geral. Entende-se por *população* o conjunto de elementos que têm, em comum, determinada característica.

As populações podem ser finitas como, por exemplo, os alunos matriculados em determinada escola, ou infinitas como, por exemplo, os resultados obtidos quando se joga um dado sucessivamente. Existem populações que, embora finitas, são consideradas infinitas para qualquer finalidade prática. Como exemplo, imagine o número de peixes que há no mar. Esse número é, em dado instante, matematicamente finito, mas tão grande que pode ser considerado infinito para qualquer finalidade prática.

Entende-se por *amostra* qualquer conjunto de elementos, retirado da população, desde que esse conjunto seja não-vazio e tenha menor número de elementos do que a população. Imagine uma população constituída por quatro elementos, isto é, por:

Ana, Pedro, Luís, Clara.

Dessa população podem ser retiradas:

a) quatro amostras de três elementos cada uma:

Ana; Pedro; Luís	Ana; Luís; Clara
Ana; Pedro; Clara	Pedro; Luís; Clara

b) seis amostras de dois elementos cada uma:

Ana; Pedro	Ana; Clara	Pedro; Clara
Ana; Luís	Pedro; Luís	Luís; Clara

c) quatro amostras de um elemento cada uma:

Ana	Luís
Pedro	Clara

Quando são coletados dados sobre todos os elementos da população, diz-se que foi feito um recenseamento. No Brasil, os recenseamentos gerais são feitos por um órgão oficial, a Fundação Instituto Brasileiro de Geografia e Estatística — IBGE —, que tem sede na cidade do Rio de Janeiro. Denomina-se *censo* o conjunto de dados obtidos através de recenseamento. Existem censos demográficos (que descrevem as caracterís-

tivas da população), censos agropecuários, censos industriais, censos comerciais e de serviços etc. No Brasil foram feitos nove recenseamentos gerais, isto é, em 1872, 1890, 1900, 1920, 1940, 1950, 1960, 1970 e 1980.

O Recenseamento Geral de 1980, a exemplo dos anteriores, atendeu às determinações da ONU (Organização das Nações Unidas) para se integrar ao Censo das Américas. Essas determinações visam garantir a uniformidade de conceitos utilizados nos censos das nações americanas.

É preciso deixar claro, porém, que nem sempre é possível coletar dados de toda a população. Um censo pode revelar-se impossível por razões de economia. Assim, por exemplo, para determinar o índice de audiência de um programa de rádio, que atinge todo o Brasil, seria absurdo realizar um censo, devido ao custo muito elevado. Então, os índices de audiência de programas de rádio e de televisão são calculados com base em amostras.

Também não se usam censos nos casos de estudos demorados, que exigem atenção cuidadosa do pesquisador. Considere, como exemplo, um estudo que tenha a finalidade de obter uma descrição pormenorizada de como uma família gasta sua renda, isto é, um estudo sobre orçamento familiar. Nesse tipo de estudo é preciso coletar grande número de dados de cada família. É impraticável coletar dados de todas as famílias brasileiras ou, até mesmo, de todas as famílias de uma cidade. Então, os estudos de orçamento familiar são, necessariamente, feitos com base em amostras.

Estes exemplos servem para lembrar que boa parte do conhecimento se baseia no estudo de amostras. No entanto, a validade de uma amostra depende da técnica utilizada para sua obtenção.

1.4 TÉCNICAS DE AMOSTRAGEM

Uma amostra só é representativa da população quando é constituída por elementos selecionados de acordo com uma técnica conhecida. Existem diversas técnicas de amostragem, mas as mais conhecidas são:

- amostragem casual simples;
- amostragem sistemática;
- amostragem estratificada.

1.4.1 Amostragem casual simples

No processo de *amostragem casual simples*, todos os elementos da população têm igual probabilidade de serem selecionados para constituir

a amostra. Como exemplo, imagine que o gerente de uma indústria pretenda obter uma amostra dos empregados. Para isso, o gerente pode conferir um número a cada empregado e sortear aqueles que irão constituir a amostra.

Para fazer o sorteio, o gerente pode usar uma urna que contenha fichas numeradas, de zero a nove, inclusive. Então o gerente retira uma ficha da urna, ao acaso, e anota o número, que dará o primeiro dígito do número do empregado que será selecionado. Depois o gerente recoloca a ficha na urna, mistura e retira outra ficha, que dará o segundo dígito do número do empregado que será selecionado. O gerente deve repetir esse processo até que sejam obtidos todos os dígitos da numeração utilizada.

Se a indústria tiver, por exemplo, 900 empregados, devem ser sorteados três dígitos para definir cada empregado que será selecionado para constituir a amostra. Então, no primeiro sorteio pode ocorrer, por exemplo:

Primeira ficha retirada:	4
Segunda ficha retirada:	7
Terceira ficha retirada:	4

O resultado do primeiro sorteio indica que foi selecionado, para a amostra, o empregado de número 474. O gerente deve repetir esse processo até completar a amostra. Evidentemente, serão desprezados números iguais a zero (0,0,0), números já sorteados (já estão na amostra) e números maiores do que 900 (não existem nessa população).

1.4.2 Amostragem sistemática

No processo de *amostragem sistemática*, os elementos são selecionados para a amostra por um sistema preestabelecido. Assim, por exemplo, imagine que um gerente de vendas quer obter uma amostra dos clientes cadastrados de sua firma. Se as fichas dos clientes estiverem organizadas por ordem alfabética, o gerente obterá uma amostra sistemática se escolher a quinta de cada cinco fichas, ou a décima, de cada dez etc.

O processo de obtenção de uma amostra sistemática é bem mais simples do que o processo de obtenção de uma amostra casual simples nos casos em que a população está organizada. Assim, para obter uma amostra dos assinantes de um jornal é mais fácil localizar os nomes dos assinantes no arquivo e, de cada 20, tirar, por exemplo, o vigésimo do que dar um número a cada assinante e fazer sorteio até completar uma amostra que corresponderia a $1/20$ da população.

Nas pesquisas domiciliares também se usa amostragem sistemática. Nesses casos é comum estabelecer um ponto de partida ao acaso em cada bairro e, a partir daí, percorrer as ruas usando determinado sistema como, por exemplo, de cada dez domicílios selecionar, para a amostra, o décimo.

1.4.3 Amostragem estratificada

Usa-se o processo de *amostragem estratificada* quando a população se apresenta dividida em estratos, isto é, quando a população está dividida em grupos distintos. Como exemplo, imagine que o diretor de uma faculdade quer obter uma amostra da comunidade acadêmica. O diretor precisa considerar que a faculdade é constituída por três grupos distintos de pessoas, isto é, professores, funcionários e alunos.

Então, para obter uma amostra mais representativa da comunidade acadêmica, o diretor deve selecionar uma amostra dentro de cada estrato, isto é, uma amostra dos professores, uma amostra dos funcionários e uma amostra dos alunos e depois reunir essas três amostras em uma só, constituindo então uma amostra estratificada.

1.5 TENDENCIOSIDADE DA AMOSTRA

Muitas vezes as amostras são constituídas por pessoas que devem executar algum tipo de tarefa, como responder perguntas, preencher um questionário ou até mesmo testar um produto. Algumas pessoas se recusam a cooperar. Nesses casos é preciso ter muito senso crítico para avaliar se a amostra efetivamente utilizada é representativa da população. Sempre é possível que a amostra obtida seja *tendenciosa* ou *viciada*, isto é, não representativa da população.

O senso crítico ainda é mais importante quando as amostras são constituídas por voluntários. Muitas vezes, o procedimento usado para solicitar voluntários conduz à formação de amostras tendenciosas. Como exemplo, imagine que um professor de educação física peça que três alunos da turma se apresentem como voluntários para apostar uma corrida. Ora, é bastante razoável imaginar que, neste caso, se apresentarão como voluntários apenas os alunos que sabem ser bons corredores. Então, os três alunos não constituirão uma amostra representativa da turma, mas uma amostra tendenciosa ou viciada.