

# 1 Introdução

Definição: Estatística é um conjunto de conceitos e métodos científicos para coleta, organização, descrição, análise e interpretação de dados experimentais, que permitem conclusões válidas e tomadas de decisões razoáveis.

Classificação: Usualmente, a estatística se divide em:

- Estatística descritiva - é a parte que tem por objetivo organizar, apresentar e sintetizar dados observados de determinada população, sem pretenções de tirar conclusões de caráter extensivo.
- Teoria de probabilidade - objetiva descrever e prever as características de populações infinitas
- Inferência Estatística é a parte que, baseando-se em estudos realizados sobre os dados de uma amostra, procura inferir, induzir ou verificar leis de comportamento da população da qual a amostra foi retirada. A estatística inferencial tem sua estrutura fundamentada na teoria matemática das probabilidades. É, também definida como um conjunto de métodos para a tomada de decisões.

No estudo da estatística alguns conceitos são importantes:

- População (N) - Conjunto de todos os elementos relativos a um determinado fenômeno que possuem pelo menos uma característica em comum, a população é o conjunto Universo. Exemplos:
  - Todos os clientes de uma determinada empresa;
  - Todos os produtos fabricados em uma determinada empresa;
- Amostra ( $n$ ) - um subconjunto da população e deverá ser considerada finita, a amostra deve ser selecionada seguindo certas regras e deve ser representativa, de modo que ela represente todas as características da população como se fosse uma fotografia desta.
- Pesquisa Estatística: É qualquer informação retirada de uma população ou amostra, podendo ser através de Censo ou Amostragem
- Censo - atividade de inspecionar (observar) todos os elementos de uma população, objetivando conhecer, com certeza suas características;
- Amostragem - É o processo de retirada de informações dos "n" elementos amostrais, no qual deve seguir um método criterioso e adequado (tipos de amostragem).
- Dados estatísticos: é qualquer característica que possa ser observada ou medida de alguma maneira. As matérias-primas da estatística são os dados observáveis.
- Variável: É aquilo que se deseja observar para se tirar algum tipo de conclusão, geralmente as variáveis para estudo são selecionadas por processos de amostragem. Os símbolos utilizados para representar as variáveis são as letras maiúsculas do alfabeto, tais como X, Y, Z, ... que pode assumir qualquer valor de um conjunto de dados. As variáveis podem ser classificadas dos seguintes modos:

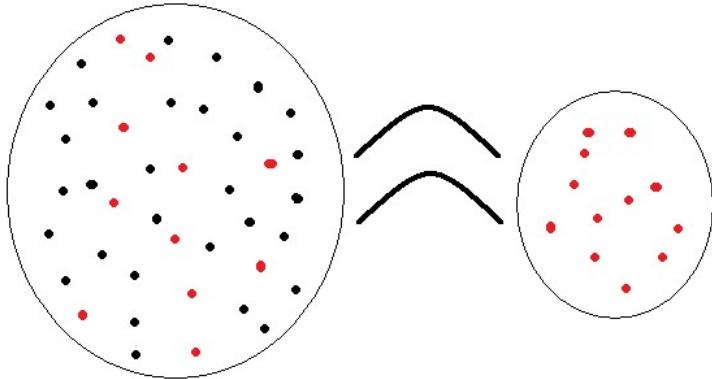


Figura 1: Representação de População×Amostra

## 1.1 Amostragem

Na realização de qualquer estudo quase nunca é possível examinar todos os elementos da população de interesse. Temos usualmente de trabalhar com uma amostra da população. A inferência estatística nos dá elementos para generalizar, de maneira segura, as conclusões obtidas da amostra para a população.

É errôneo pensar que, caso tivéssemos acesso a todos os elementos da população, seríamos mais preciosos. Os erros de coleta e manuseio de um grande número de dados são maiores do que as imprecisões a que estamos sujeitos quando generalizamos, via inferência, as conclusões de uma amostra bem selecionada.

Em se tratando de amostra, a preocupação central é que ela seja representativa. Assim que decidimos obter informações através de um levantamento amostral, temos imediatamente dois problemas:

- Definir cuidadosamente a população de interesse
- Selecionar a característica que iremos pesquisar.

Há duas grandes divisões no processo de amostragem: a probabilística e a não-probabilística.

A amostragem probabilística também é chamada de amostragem aleatória ou ao acaso. Este tipo de amostragem é submetida a tratamento estatístico que permite compensar erros amostrais. Hoje, dificilmente se aceita uma amostragem não-probabilística, exceto nos casos em que a amostragem probabilística não pode ser feita.

A amostragem não-probabilística, por não fazer uso de forma aleatória de seleção, não aceita diversas aplicações estatísticas e, por isto, é preterida.

Pontos importantes:

- É muito dispendioso entrevistar cada pessoa de toda uma população; recorremos, então, as amostras;
- Usa-se a proporção de pessoas em uma amostra, portadoras de determinada característica, para estimar a proporção, na população das que tem essa característica.
- O melhor método de escolha de uma amostra é a escolha aleatória, isto é, que toda amostra possível tenha a mesma chance de ser escolhida.

- Antes de se proceder a observação de uma determinada população surge a questão se a amostragem será com ou sem reposição. Se o tamanho da amostra é insignificante em relação à população o impacto da reposição será desprezível, porém, se a amostra for grande então a reposição ou não pode causar um impacto significativo no resultado da probabilidade.
- Como as características das populações estatísticas variam, às vezes, é necessário se adequar esta população estatística para submetê-la a um critério de seleção possível, sem, contudo, perder seu caráter aleatório.

### 1.1.1 Amostragem Simples

Objetivo: Obter uma amostra representativa quando os elementos da população são todos homogêneos. Neste processo de amostragem todos os elementos da população têm a mesma probabilidade de serem amostrados. A característica principal é que todos os elementos da população têm igual probabilidade de pertencer à amostra.

Procedimento: Na prática a amostragem aleatória simples pode ser realizada numerando-se a população de 1 a N e sorteando-se, a seguir, por meio de um dispositivo aleatório qualquer,  $k$  números dessa seqüência, os quais corresponderão aos elementos pertencentes à amostra.

Exemplo: Vamos obter uma amostra representativa, de 10% dos valores, para obtermos a estatura média de noventa alunos de uma escola:

- Numeramos os alunos de 01 a 90
- Sorteamos os números, de 01 a 90, um a um, nove números que formarão a amostra.

### 1.1.2 Amostragem Sistemática

Objetivo: Aumentar a representatividade da amostra dando maior cobertura à população. É usada quando todos os elementos são homogêneos.

Procedimento: Quando os elementos da população já estão ordenados, não há necessidade de construirmos um sistema de referência, para selecionarmos a amostra. São exemplos os prontuários médicos de um hospital, os prédios de uma rua, uma linha de produção, os nomes em uma lista telefônica, etc. Nestes casos a seleção dos elementos que constituirão a amostra pode ser feita por um sistema imposto pelo pesquisador. A esse tipo de amostragem denominamos de sistemática.

Exemplo: Consideremos uma população, com elementos ordenados, de tamanho  $N$  e dela tiramos uma amostra de tamanho  $n$ , através de uma amostragem sistemática, da seguinte maneira:

- Definimos  $FS$  como fator de sistematização, dado por  $FS = N/n$ .
- Sorteamos um número entre 1 e  $FS$ . Esse número é simbolizado por  $m$ , que será o primeiro elemento da amostra.
- O segundo elemento da amostra é o de número  $FS + m$ .
- O terceiro elemento da amostra é o de número  $2FS + m$ .

- O  $k$ -ésimo elemento da amostra é o número  $(k - 1)FS + m$

Exemplo: Uma rua contém 1000 prédios, dos quais desejamos obter uma amostra sistemática formada por 100 deles.

$$FS = 1000/100 = 10$$

$m$  será um número entre 1 e 10. Vamos supor que  $m = 7$ . Então temos:

- 1º elemento da amostra =  $(1 - 1)10 + 7 = 7 \gg 7$ º elemento da população.
- $\vdots$
- 100º elemento da amostra =  $(100 - 1)10 + 7 = 997 \gg 997$ º elemento da população.

### 1.1.3 Amostragem Estratificada

Objetivo: Melhorar a representatividade da amostra quando os elementos da população são heterogêneos, porém, podem ser agrupados em subpopulações (ESTRATOS) contendo elementos homogêneos.

Procedimento: A população é dividida em grupos ou estratos contendo elementos homogêneos e as amostras são retiradas separadamente de cada um desses grupos.

Exemplo: Dada a população de 50.000 operários da indústria, selecionar uma amostra proporcional estratificada de 5% de operários para estimar seu salário médio. Usando a variável critério "cargo" para estratificar essa população, e considerando amostras de 5% de cada estrato obtido, chegamos ao seguinte quadro:

CARGO	POPULAÇÃO	5%	AMOSTRA
Chefes de seção	5000	$5(5000)/100 = 250$	250
Operários especializados	15000	$5(15000)/100 = 750$	750
Operários não especializados	30000	$5(30000)/100 = 1500$	1500
<b>TOTAL</b>	<b>50000</b>	<b><math>5(50000)/100 = 2500</math></b>	<b>2500</b>

## 1.2 Método Estatístico

O Método Estatístico pode ser descrito pelas etapas a seguir:

- Definição do problema - Consiste na:
  - formulação correta do problema;
  - examinar outros levantamentos realizados no mesmo campo (revisão da literatura);
  - saber exatamente o que se pretende pesquisar definindo o problema corretamente (variáveis, população, hipóteses, etc.)
- Planejamento -Determinar o procedimento necessário para resolver o problema:
  - Como levantar informações;
  - Tipos de levantamentos: Por Censo (completo); Por Amostragem (parcial).
  - Cronograma, Custos, etc.

- Coleta da dados - Consiste na obtenção dos dados referentes ao trabalho que desejamos fazer.;
  - A coleta pode ser: Direta - diretamente da fonte ou Indireta - feita através de outras fontes.
  - Os dados podem ser obtidos pela própria pessoa (primários) ou se baseia no registro de terceiros (secundários).
- Apuração dos dados - Consiste em resumir os dados, através de uma contagem e agrupamento. É um trabalho de coordenação e de tabulação.
- Apresentação dos dados -É a fase em que vamos mostrar os resultados obtidos na coleta e na organização. Esta apresentação pode ser:
  - Tabular (apresentação numérica)
  - Gráfica (apresentação geométrica)
- Análise e interpretação dos dados - É a fase mais importante e também a mais delicada. Tira conclusões que auxiliam o pesquisador a resolver seu problema.