

Amostragem

Catarina Marques

Conceitos: População, Unidade Estatística e Amostra

- **População** (ou Universo) – dimensão **N**

- Conjunto de unidades com uma ou mais características comuns

- **População alvo** - Conjunto de unidades que se pretendem estudar

- **População inquirida** - Conjunto de unidades para as quais é operacionalmente possível recolher uma amostra

- **Unidade Estatística**

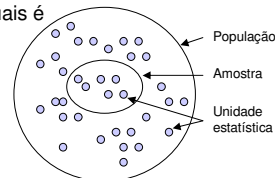
- Um elemento da população

- **Amostra**

- Subconjunto da População (n° de elementos é **n**).

- Estudam-se apenas algumas características específicas, sendo cada uma delas representada por uma variável, **X**.

- O processo de recolha de informação de toda a população é denominado de **censo** ou **recenseamento**. A consulta a um subconjunto da população é denominada de **amostragem** ou **sondagem**.



Inferência Estatística

- **Inferência Estatística**

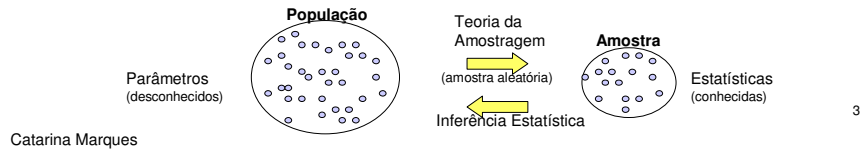
- Permite generalizar à população certas conclusões, tendo em conta o resultados amostrais, com a indicação da precisão dessas inferências. O seu uso permite estimar as características (**parâmetros**) desconhecidas da população e testar se as hipóteses formuladas sobre as características desconhecidas são plausíveis.

- **Parâmetros** – Resumem as características da população e assumem valores numéricos fixos (constantes).

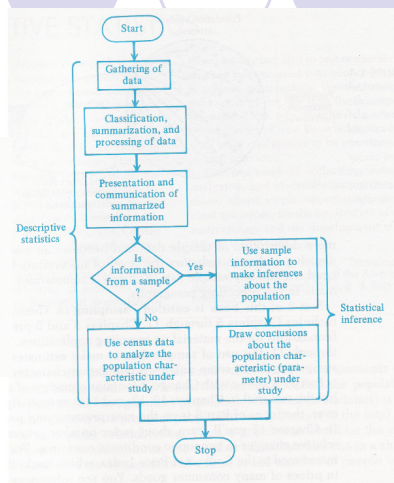
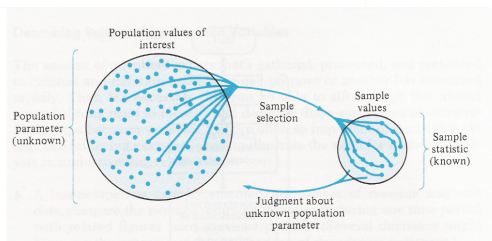
- Ex: média da população (μ), variância populacional (σ^2), desvio-padrão populacional (σ)

- **Estatísticas** – É toda a função que opera sobre a amostra. Dado que de uma população se podem retirar diversas amostras estamos perante valores que variam de amostra para amostra.

- Ex: média amostral ($\bar{x}$), a variância amostral (S^2), desvio-padrão amostral (S)



3



4

Catarina Marques

Fases do Processo de Amostragem

- Definição clara dos objectivos de estudo
- Definição dos dados a recolher e escolha do instrumento a utilizar (ex: inquérito)
- Definição do processo adequado ao tipo de dados e instrumento de recolha
 - Estabelecer um plano de amostragem
 - Construção da amostra
 1. População alvo/inquirida
 2. Método de selecção da amostra
 3. Dimensão da amostra

Métodos de Amostragem

Métodos de amostragem aleatória

- Amostragem aleatória simples
- Amostragem sistemática
- Amostragem estratificada
- Amostragem por clusters
- Amostragem multi-etápica
- Amostragem multi-fásica

Métodos de amostragem não aleatória

- Amostragem por conveniência
- Método intencional
- Amostragem Snowball
- Amostragem por quotas

Amostragem aleatória simples

- Cada elemento da população tem a mesma probabilidade de ser seleccionado que qualquer um dos restantes
- Cada amostra de dimensão n tem a mesma probabilidade de ser seleccionada que qualquer outra da mesma dimensão
- Implica a existência de uma lista exaustiva da população (base de sondagem)
- Duas formas de proceder:
 - Lotaria
 - Números aleatórios

Distribuição de uma amostra aleatória

Amostra aleatória é uma variável aleatória:

- Pretendemos estudar uma certa característica de uma população, que denotamos por X
- Uma amostra de dimensão n dessa população X será $(X_1, X_2, \dots, X_n)$
 - x_i é o valor observado de X para o i -ésimo elemento da amostra
 - É uma concretização de X para o i -ésimo elemento ou uma concretização de X_i
- De cada vez que recolhemos uma amostra da mesma dimensão, obtemos valores diferentes em cada um dos elementos
- Logo uma amostra aleatória é uma variável aleatória multidimensional.

Distribuição de uma amostra aleatória

Função de probabilidade (se X discreta) ou função densidade de probabilidade (fdp) (se X contínua) da amostra aleatória:

- As v.a. $X_1, X_2, \dots, X_n$ assumem os mesmos valores de X, pois são elementos de uma amostra, todos retirados da mesma população, pelo que têm a mesma função (densidade) de probabilidade da população
 $f(x_1)=f(x_2)=\dots=f(x_n)=f(x)$
- Como $(X_1, X_2, \dots, X_n)$ é uma amostra recolhida segundo um processo aleatório (vamos considerar a.a.s.), os seus elementos são v.a. independentes entre si.

$X_1, X_2, \dots, X_n$ são v.a. independentes e identicamente distribuídas (iid)

- Então a função (densidade) de probabilidade da amostra é dada pela função (densidade) de probabilidade conjunta:

$$f(x_1, x_2, \dots, x_n) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n) = \prod_{i=1}^n f(x_i)$$

Distribuições amostrais

- Quando não é possível estudar a População, é necessário recolher uma Amostra representativa, que será usada para tirar conclusões sobre a População.
- Por exemplo, poderemos calcular a média amostral, a **estatística** $\bar{X}$, e utilizá-la como um **estimador** para o **parâmetro** média populacional, μ .
- O erro de amostragem é dado pela diferença entre o parâmetro populacional e a estatística amostral $\bar{X} - \mu$.
- Se, para todas as amostras possíveis de uma certa dimensão n , obtermos todos os valores possíveis para a estatística e a probabilidade associada a cada um destes valores, estamos a construir a **distribuição amostral da estatística**.