

LCE0216
Introdução à Bioestatística Florestal
7. Noções de Amostragem

Profa. Dra. Clarice Garcia Borges Demétrio

Monitor: Silvio Gomes

Escola Superior de Agricultura "Luiz de Queiroz"
Universidade de São Paulo

Piracicaba, 28 de abril de 2020

Conceitos iniciais

População

É o conjunto de indivíduos, ou objetos, tendo pelo menos uma variável observável em comum.

Unidade Amostral

Um item individual da população.

Amostra

É qualquer subconjunto de elementos da população. Em geral, no entanto, trabalhamos apenas com amostras aleatórias, em que todos os elementos da população têm a mesma probabilidade de serem escolhidos.

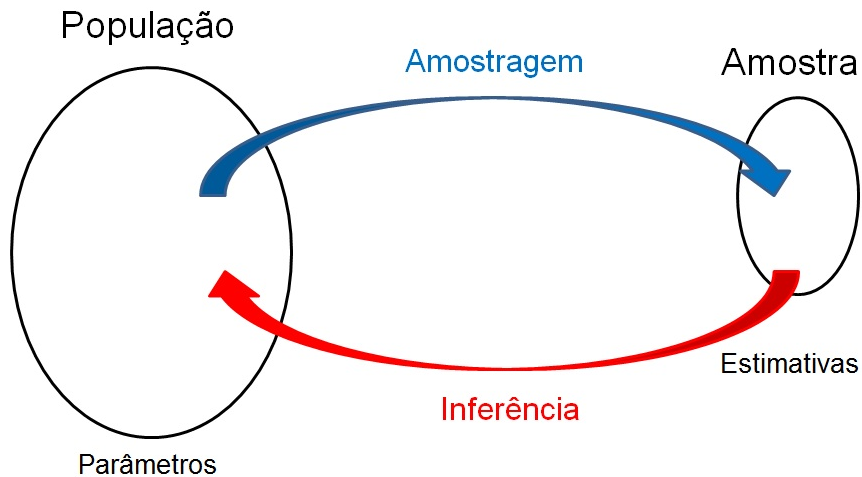
Variável

É a característica de interesse que é medida/observada em cada unidade amostral.

Censo: Exame de todos os elementos da população.

- ▶ O Censo do Brasil é um censo realizado a cada 10 anos pelo Instituto Brasileiro de Geografia e Estatística (IBGE).
- ▶ A população é contada em todo o território do Brasil e os resultados são usados pelo governo no desenvolvimento de políticas públicas e na destinação dos fundos governamentais para as unidades federativas.
- ▶ Em 1872 foi realizado o primeiro censo nacional no Brasil.
- ▶ O último censo realizado no Brasil ocorreu em 2010.

Desvantagens: Caro, demorado e nem sempre possível.



Exemplo: O objetivo de um levantamento florestal é determinar a taxa média de ocorrência de seca-do-ponteiro numa floresta de *Eucalyptus saligna*.

População:

Unidade Amostral:

Variável:

Amostra:

Parâmetro:

Estimador:

Exemplo: O objetivo de um levantamento florestal é determinar a taxa média de ocorrência de seca-do-ponteiro numa floresta de *Eucalyptus saligna*.

População: Floresta de *E. saligna*;

Unidade Amostral:

Variável:

Amostra:

Parâmetro:

Estimador:

Exemplo: O objetivo de um levantamento florestal é determinar a taxa média de ocorrência de seca-do-ponteiro numa floresta de *Eucalyptus saligna*.

População: Floresta de *E. saligna*;

Unidade Amostral: Árvore;

Variável:

Amostra:

Parâmetro:

Estimador:

Exemplo: O objetivo de um levantamento florestal é determinar a taxa média de ocorrência de seca-do-ponteiro numa floresta de *Eucalyptus saligna*.

População: Floresta de *E. saligna*;

Unidade Amostral: Árvore;

Variável: Ocorrência do seca-do-ponteiro;

Amostra:

Parâmetro:

Estimador:

Exemplo: O objetivo de um levantamento florestal é determinar a taxa média de ocorrência de seca-do-ponteiro numa floresta de *Eucalyptus saligna*.

População: Floresta de *E. saligna*;

Unidade Amostral: Árvore;

Variável: Ocorrência do seca-do-ponteiro;

Amostra: Levantamento florestal;

Parâmetro:

Estimador:

Exemplo: O objetivo de um levantamento florestal é determinar a taxa média de ocorrência de seca-do-ponteiro numa floresta de *Eucalyptus saligna*.

População: Floresta de *E. saligna*;

Unidade Amostral: Árvore;

Variável: Ocorrência do seca-do-ponteiro;

Amostra: Levantamento florestal;

Parâmetro: Média;

Estimador:

Exemplo: O objetivo de um levantamento florestal é determinar a taxa média de ocorrência de seca-do-ponteiro numa floresta de *Eucalyptus saligna*.

População: Floresta de *E. saligna*;

Unidade Amostral: Árvore;

Variável: Ocorrência do seca-do-ponteiro;

Amostra: Levantamento florestal;

Parâmetro: Média;

Estimador: $\bar{X} = \sum_{i=1}^n X_i / n$ (média aritmética).

Amostragem

É o processo da escolha da amostra. Consiste na escolha criteriosa dos elementos a serem submetidos ao estudo.

Quando uma amostra é boa?

- ▶ Viés...
- ▶ Precisão...



Viés

Tendenciosidade presente no processo de amostragem que impede que a amostra seja perfeitamente representativa da população, independentemente do tamanho da amostra.

Vários tipos de vieses podem ocorrer no processo de amostragem, como **viés de seleção**, **viés de não resposta** e **viés de resposta**, como será visto a seguir.

Amostragem: conceitos iniciais

- ▶ **Viés de seleção:** Tendência sistemática de subrepresentar ou excluir da amostra um certo tipo de unidade;
Ex.: Pesquisa de opinião por telefone.
- ▶ **Viés de não resposta:** Numa pesquisa de opinião, o resultado de pessoas não responderem às questões do entrevistador;
Ex.: Pesquisa de opinião por correio.
- ▶ **Viés de resposta:** Numa pesquisa de opinião, a forma de apresentação das questões e a atitude do entrevistador podem influenciar a resposta.
Ex.: Ordem do nome dos candidatos numa pesquisa de intenção de votos.

Exemplo: Estudando a influência do sistema de amostragem na estrutura de tamanho das árvores de uma floresta tropical, um Ecologista tomou 10 Amostras compostas de 5 árvores, 10 amostras compostas de 20 árvores e 10 amostras compostas por 40 árvores.

- ▶ Qual o efeito do tamanho da amostra (número de árvores) nos resultados obtidos?
- ▶ Qual tipo de amostra se mostra menos variável?

Tabela: Média do diâmetro de árvores amostradas numa floresta nativa

Amostras	Tamanho das amostras (número de árvores)		
	5	20	40
1	26,6	34,1	30,5
2	32,9	31,2	29,8
3	39,2	40,1	37,3
4	30,3	30,4	29,4
5	31,8	29,7	32,7
6	48,0	38,8	31,5
7	37,7	37,3	29,3
8	20,0	35,6	30,6
9	31,8	37,7	31,2
10	33,2	26,7	29,3
Desvio padrão	7,48	4,48	2,42
CV(%)	22,55	13,13	7,75

Precisão

Capacidade dos sistemas de amostragem de gerar Estatísticas com pequena variabilidade entre amostras (de mesmo tamanho) tomadas numa mesma população.



Sem Viés
Sem Precisão



Com Viés
Com Precisão



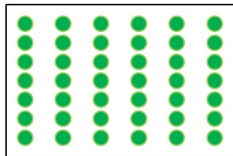
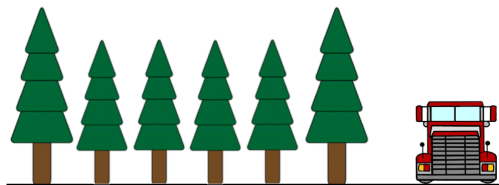
Sem Viés
Com Precisão

Conceitos iniciais

Exemplo: Uma equipe precisa determinar a altura média de um talhão para a determinação do volume de madeira produzido.

Como você realizaria a coleta dessa informação? Comente sobre:

- 1 Equipamentos;
- 2 Equipe de coleta;
- 3 Plano de amostragem.



Alguns métodos mais comuns de extração da amostra:

- ▶ Amostragem aleatória simples
- ▶ Amostragem estratificada
- ▶ Amostragem sistemática
- ▶ Amostragem por conglomerados

É um método de selecionar, sem reposição, n elementos de uma população de tamanho N , conhecido e finito, em que cada elemento tem a mesma chance de ser selecionado.

Procedimento para a obtenção de uma amostra aleatória simples de tamanho n :

- 1 Enumerar os N elementos da população de 1 a N .
- 2 Sortear, sem reposição, n números compreendidos de 1 a N .
- 3 Os elementos correspondentes aos números escolhidos formarão a amostra.

Na amostragem estratificada, a população (por exemplo de uma escola) é dividida em estratos (alunos do 3º, 4º e 7º ano) e os elementos que formarão a amostra são retirados de dentro de cada um desses estratos.

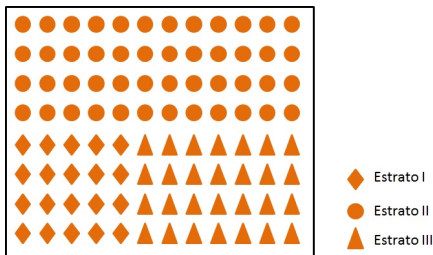
Uma amostragem estratificada é indicada quando:

- ▶ É fácil o acesso à uma lista contendo todos os nomes dos elementos da população;
- ▶ As informações sobre a população estão disponíveis e embora ela seja heterogênea posso identificar grupos homogêneos dentro dessa mesma população e assim dividi-la em diferentes estratos para depois obter uma amostra.

Amostragem Estratificada

Exemplo: Seja uma área florestal, onde, por meio de técnicas de interpretação de imagens fotográficas, foram detectados três estratos: Estrato I (floresta alta sem babaçu), Estrato II (floresta alta com babaçu), Estrato III (floresta baixa cipoálica).

- ▶ Deseja-se estimar o volume de madeira dessa área (m^3) observando-se apenas 10 unidades amostrais.
- ▶ Como você realizaria essa amostragem?
- ▶ Você recomendaria o uso da amostragem aleatória simples?



Amostragem Estratificada

Na **amostragem estratificada** a população de tamanho N é dividida em L subpopulações disjuntas, denominadas estratos de tamanho N_1, N_2, \dots, N_L . Em seguida, é feita uma amostragem aleatória simples dentro de cada estrato que pode ser proporcional ou não ao tamanho de cada estrato.

Amostragem estratificada proporcional

É uma amostragem estratificada em que o número n_i de elementos amostrados em cada estrato é proporcional ao tamanho N_i de cada um dos estratos ($i = 1, 2, \dots, L$).

Exemplo: Uma comunidade universitária com 8000 indivíduos está estratificada da seguinte forma

Estrato	População	Amostra
Professores	800	80
Funcionários	1200	120
Estudantes	6000	600

Amostragem Estratificada

Procedimento para a obtenção de uma amostra estratificada de tamanho n :

Considerando:

- ▶ N o número de elementos da população
- ▶ L o número de estratos
- ▶ N_i o número de elementos do i -ésimo estrato tal que $N = N_1 + N_2 + \dots + N_L$
- ▶ n o tamanho da amostra a ser selecionada,

- 1 Calcular a fração da amostragem dada por $f = \frac{n}{N}$
- 2 Calcular o número de elementos a serem sorteados em cada estrato:

$$n_1 \approx N_1 f \quad n_2 \approx N_2 f \quad \dots \quad n_L \approx N_L f,$$

com $n = n_1 + n_2 + \dots + n_L$.

Cálculo da média

$$\bar{x}_{AE} = \frac{N_1 \bar{x}_1 + N_2 \bar{x}_2 + \dots + N_L \bar{x}_L}{N},$$

em que \bar{x}_i corresponde à média aritmética simples dos n_i elementos amostrados do estrato i .

Exercício: Considerando-se o esquema apresentado anteriormente, obter uma amostra estratificada de tamanho $n = 10$ e indicar as plantas selecionadas.

Amostragem por Conglomerado

A área da população é dividida em seções (ou conglomerados, ex.: bairros, quarteirões, famílias, organizações). Os conglomerados são selecionados aleatoriamente. Dentro de um conglomerado, todos os elementos são amostrados.

- ▶ O número de elementos de um conglomerado deve ser pequeno e o número de conglomerados razoavelmente grande.



Amostragem por Conglomerado

Uma amostragem por conglomerado é indicada quando:

- ▶ Não se possui uma lista contendo todos os nomes dos elementos da população;
- ▶ Existe grande heterogeneidade entre os elementos da população;
- ▶ É preciso fazer entrevistas ou observações em grandes áreas geográficas;
- ▶ O custo de obtenção dos dados cresce com o aumento da distância entre os elementos;

Amostragem por Conglomerado

Procedimento para a obtenção de uma amostra por conglomerado

É uma amostra aleatória simples na qual cada unidade de amostragem é um grupo de elementos, chamado conglomerado.

- 1 Definir conglomerado
- 2 Selecionar amostras aleatórias simples de grupos (conglomerados)
- 3 Todos os elementos dentro do grupo selecionado constituirão a amostra.

Procedimento para a obtenção de uma amostra sistemática (1 em cada k elementos)

1. Definir a porcentagem $P\%$ de elementos da população que farão parte da amostra
2. Obter o valor de k , inteiro, dado por:

$$k \approx \frac{1}{P} \times 100$$

3. Sortear um número r inteiro entre 1 e k
4. A amostra será composta pelos elementos de ordem

$$r, r + k, r + 2k, r + 3k, \dots$$

Exemplo: Um restaurante deseja conhecer melhor o perfil de seus clientes e tem condições para entrevistar aproximadamente 15% dos mesmos.

Os clientes que visitaram o restaurante num certo dia, por ordem de chegada, foram:

JJ	MR	SSZ	RA	RMA
CRG	CGB	DB	JEC	AAS
TZC	CTD	SMP	EM	DF
LB	SS	ROS	GS	SMC
MCN	JB	AFI	VFC	RSM
CF	MHT	PHA	MAP	AO
...				

Como você faria essa amostragem?

Fatores que determinam o tamanho da amostra

- ▶ Nível de confiança (quanto maior o nível de confiança, maior o tamanho da amostra);
- ▶ Erro máximo permitido (quanto menor o erro permitido, maior o tamanho da amostra);
- ▶ Variabilidade do fenômeno que está sendo pesquisado (quanto maior a variabilidade, maior o tamanho da amostra).