

LCE0216
Introdução à Bioestatística Florestal
9. Estimação pontual e intervalar

Profa. Dra. Clarice Garcia Borges Demétrio

Monitor: Silvio Gomes

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

Piracicaba, 12 de maio de 2020

Avaliar características da população com base em informações da amostra



Estimar os parâmetros

Alguns parâmetros de interesse:

- ▶ Média (μ)
- ▶ Proporção (π)
- ▶ Variância (σ^2)

Exemplos:

- ▶ Produção média de determinada cultura;
- ▶ Proporção média de área foliar atacada por uma praga;
- ▶ Parâmetros estatísticos genéticos (variância genética, ambiental e fenotípica)...

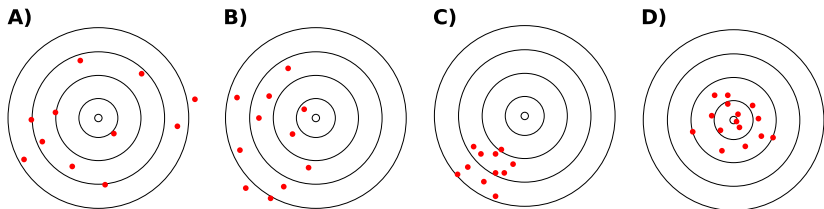
- ▶ Um **estimador** $\hat{\theta}$ do parâmetro θ é qualquer função dos elementos da amostra, ou seja, $\hat{\theta} = g(X_1, X_2, \dots, X_n)$.
- ▶ Uma **estimativa** $\hat{\theta}$ do parâmetro θ , é o valor assumido pelo estimador quando aplicado aos valores observados na amostra, ou seja, $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$.
- ▶ A notação para estimador e estimativa, geralmente, é a mesma $\hat{\theta}$ (para símbolo grego). No caso de símbolos latinos, usa-se letra maiúscula para estimador (\bar{X}) e letra minúscula para estimativa (\bar{x}).

Exemplos de estimadores:

- ▶ Para a média: $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$
- ▶ Para a proporção: $P = \frac{\text{número de sucessos}}{n}$
- ▶ Para a variância: $S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$

- ▶ **não viesado**: média da distribuição amostral igual ao parâmetro
- ▶ **preciso**: variância amostral pequena
- ▶ **acurado**: erro amostral pequeno

Propriedades dos estimadores



A): não viesado, pouca precisão e pouca acurácia;

B): viesado, pouca precisão e pouca acurácia;

C): viesado, boa precisão e baixa acurácia;

D): não viesado, boa precisão e boa acurácia.

Valor numérico assumido pelo estimador quando aplicado aos valores observados de uma amostra.

- ▶ Média: $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$;
- ▶ Proporção: $p = \frac{\text{número de sucessos}}{n}$;
- ▶ Variância: $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$.

Intervalo de confiança

Seja (X_1, X_2, \dots, X_n) uma amostra aleatória de tamanho n de uma população e θ o parâmetro de interesse. Sejam $\hat{\theta}_1$ e $\hat{\theta}_2$ estatísticas tais que:

$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha.$$

Então o intervalo $(\hat{\theta}_1; \hat{\theta}_2)$ é chamado intervalo de **100(1- α)% de confiança** para o parâmetro θ .

Usualmente, toma-se $1 - \alpha = 0,95$ ou $0,99$.

Interpretação: De todos os possíveis intervalos que possam ser construídos, espera-se que $100(1 - \alpha)\%$ deles contenham o verdadeiro valor do parâmetro θ .

Intervalos de confiança para média populacional

Casos

- ▶ População normal e variância da população conhecida;
- ▶ População normal e variância da população desconhecida;
- ▶ População não normal, grandes amostras ($n > 30$).

Intervalo de confiança para média

► População normal e variância populacional conhecida

$$X \sim N(\mu, \sigma^2) \Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1)$$

$$P\left(-z_T < \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} < z_T\right) = 1 - \alpha$$

...

$$P\left(\bar{X} - z_T \sqrt{\frac{\sigma^2}{n}} < \mu < \bar{X} + z_T \sqrt{\frac{\sigma^2}{n}}\right) = 1 - \alpha$$

$$IC(\mu)_{1-\alpha} = \left(\bar{X} - z_T \sqrt{\frac{\sigma^2}{n}}; \bar{X} + z_T \sqrt{\frac{\sigma^2}{n}}\right)$$

► População normal e variância populacional conhecida

Exemplo: A distribuição dos pesos de pacotes de determinadas sementes, enchidos automaticamente por uma certa máquina, é normal, com desvio padrão (σ) conhecido e igual a 0,20 kg. Uma amostra de 15 pacotes retirada ao acaso apresentou os seguintes pesos, em kg:

20,05	20,10	20,25	19,78	19,69	19,90	20,20	19,89
19,70	20,30	19,93	20,25	20,18	20,01	20,09	

Construir os intervalos de confiança de 95% e 99% para o peso médio dos pacotes de sementes.

Solução:

$$\sigma_X = 0,20 \quad n = 15$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{300,32}{15} = 20,02;$$

$$IC(\mu)_{1-\alpha} = \left(\bar{X} - z_T \sqrt{\frac{\sigma^2}{n}}; \bar{X} + z_T \sqrt{\frac{\sigma^2}{n}} \right)$$

$$1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_T = z_{0,975} = 1,96$$

$$IC(\mu)_{0,95} = \left(20,02 - 1,96 \frac{0,20}{\sqrt{15}}; 20,02 + 1,96 \frac{0,20}{\sqrt{15}} \right) = (19,92; 20,12)$$

$$1 - \alpha = 0,99 \Rightarrow \frac{\alpha}{2} = 0,005 \Rightarrow z_T = z_{0,995} = 2,58$$

$$IC(\mu)_{0,99} = \left(20,02 - 2,58 \frac{0,20}{\sqrt{15}}; 20,02 + 2,58 \frac{0,20}{\sqrt{15}} \right) = (19,89; 20,15)$$

► População normal e variância populacional desconhecida

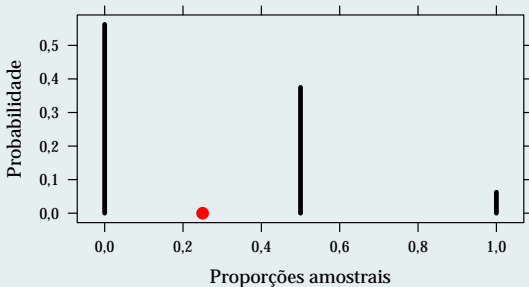
Nova estatística:

$$T = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}} \sim t_{(n-1)}$$

$$IC(\mu)_{1-\alpha} = \left(\bar{X} - t_T \sqrt{\frac{S^2}{n}}; \bar{X} + t_T \sqrt{\frac{S^2}{n}} \right)$$

Distribuição t de Student

- ▶ Simétrica em relação ao zero;
- ▶ Semelhante à distribuição normal padrão, porém com “caudas mais grossas”;
- ▶ Para $n \rightarrow \infty$ ($n \geq 30$) a distribuição t tende para a normal padrão



Intervalo de confiança para média

► População normal e variância populacional desconhecida

Exemplo:

Para estudar a maturidade de certo reflorestamento, um comprador tomou uma amostra aleatória simples de 16 árvores, obtendo os dados de D.A.P., a seguir:

16,6	23,2	17,0	21,3	19,2	20,3	20,4	21,5
17,1	19,3	20,4	22,0	19,6	18,2	19,9	18,7

Supondo que a distribuição dos dados de D.A.P. é aproximadamente normal,

- Determinar estimativas por ponto para a média e para a variância dos D.A.P.s desse reflorestamento;
- Construir um intervalo de 95% de confiança para μ ;
- Calcule o tamanho de n da amostra necessária para que se obtenha um intervalo de confiança com semi-amplitude $d=0,5$ cm, ao nível de significância $\alpha = 0,05$.

Solução:

(a) $n = 16$

$$\text{Média: } \bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{314,7}{16} = 19,67;$$

$$\text{Variância: } s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right] = \\ \frac{1}{16-1} \left[6241,79 - \frac{(314,7)^2}{16} \right] = 3,47$$

(b) $IC(\mu)_{1-\alpha} = \left(\bar{X} - t_T \sqrt{\frac{S^2}{n}}; \bar{X} + t_T \sqrt{\frac{S^2}{n}} \right)$

$$1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow t_T = t_{0,975;15} = 2,1315$$

$$IC(\mu)_{0,95} : \left(19,67 \pm 2,1315 \sqrt{\frac{3,47}{15}} \right) = (18,68; 20,66)$$

(c) $1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_T = z_{0,975} = 1,96$

$$d = z_T \sqrt{\frac{s^2}{n}} \Rightarrow n = \frac{z_T^2 s^2}{d^2} = \frac{1,96^2 \times 3,47}{0,5^2} \simeq 53$$

► **População não normal, grandes amostras ($n > 30$)**

Pelo Teorema Central do Limite, se n for razoavelmente grande ($n > 30$), então

$$\frac{\bar{X} - \mu}{\sqrt{\frac{s^2}{n}}} \sim N(0, 1)$$

e o intervalo de $100(1 - \alpha)\%$ de confiança para a média μ da população é dada por:

$$IC(\mu)_{1-\alpha} = \left(\bar{X} - z_T \sqrt{\frac{s^2}{n}}; \bar{X} + z_T \sqrt{\frac{s^2}{n}} \right)$$

► **População não normal, grandes amostras ($n > 30$)**

Exemplo: Para se avaliar o número médio de árvores de uma certa espécie por ha, numa determinada área, foram observadas 32 unidades amostrais de 1 ha, obtendo-se uma média de 3,3 árvores por ha e variância 3,2 (árvores por ha)².

- (a) Construir os intervalos de 95% e 99% de confiança para o número médio lagartas na área total.
- (b) Calcular o tamanho n da amostra necessária para que se obtenha um intervalo de 95% de confiança com precisão $d = 0,4$ árvores por ha.
- (c) Calcular o tamanho n da amostra necessária para que se obtenha um intervalo de 99% de confiança com precisão $d = 0,4$ árvores por ha.

Solução:

$$n = 32, \quad \bar{x} = 3,3 \text{ árvores/ha}, \quad s^2 = 3,2 \text{ (árvores/ha)}^2$$

$$(a) \quad IC(\mu)_{1-\alpha} = \left(\bar{X} - z_T \sqrt{\frac{s^2}{n}}; \bar{X} + z_T \sqrt{\frac{s^2}{n}} \right)$$

$$1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_T = z_{0,975} = 1,96$$

$$IC(\mu)_{0,95} : \left(3,3 \pm 1,96 \sqrt{\frac{3,2}{32}} \right) = (2,7; 3,9)$$

$$1 - \alpha = 0,99 \Rightarrow \frac{\alpha}{2} = 0,005 \Rightarrow z_T = z_{0,995} = 2,58$$

$$IC(\mu)_{0,99} : \left(3,3 \pm 2,58 \sqrt{\frac{3,2}{32}} \right) = (2,5; 4,1)$$

$$(b) \quad 1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_T = z_{0,975} = 1,96$$

$$d = z_T \sqrt{\frac{s^2}{n}} \Rightarrow n = \frac{z_T^2 s^2}{d^2} = \frac{1,96^2 \times 3,2}{0,4^2} \simeq 76$$

$$(c) \quad 1 - \alpha = 0,99 \Rightarrow \frac{\alpha}{2} = 0,005 \Rightarrow z_T = z_{0,995} = 2,58$$

$$d = z_T \sqrt{\frac{s^2}{n}} \Rightarrow n = \frac{z_T^2 s^2}{d^2} = \frac{2,58^2 \times 3,2}{0,4^2} \simeq 133$$

Intervalo de confiança para proporção

$$IC(\pi)_{1-\alpha} = \left(\hat{\pi} - z_T \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{n}}; \hat{\pi} + z_T \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{n}} \right)$$

Exemplo: Coletou-se uma amostra de 35 peixes da espécie *Xenomelaniris brasiliensis*, na localidade da praia da Barra da Lagoa, Florianópolis, SC, a qual apresentou 45,7% de peixes com comprimento total acima de 50 mm.

- ▶ Encontre um intervalo com 95% de confiança, dentro do qual deve estar a verdadeira proporção de peixes dessa espécie com comprimento acima de 50 mm.
- ▶ Qual o tamanho da amostra necessário para que tenhamos 95% de confiança de que o erro de nossa estimativa não seja superior a cinco pontos percentuais (0,05)?

Solução:

$$n = 35, \quad \hat{\pi} = 0,457$$

$$(a) \quad 1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_T = z_{0,975} = 1,96$$

$$IC(\pi)_{1-\alpha} = \left(\hat{\pi} - z_T \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{n}}; \hat{\pi} + z_T \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{n}} \right)$$

$$IC(\mu)_{0,95} : \left(0,457 \pm 1,96 \sqrt{\frac{0,457(1 - 0,457)}{35}} \right) = (0,31; 0,64)$$

$$(b) \quad 1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_T = z_{0,975} = 1,96$$

$$d = z_T \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{n}}$$

$$n = \frac{z_T^2 \hat{\pi}(1 - \hat{\pi})}{d^2} = \frac{1,96^2 \times 0,457(1 - 0,457)}{0,05^2} \simeq 383$$