



UNIVERSIDADE FEDERAL DA BAHIA  
INSTITUTO DE MATEMÁTICA E  
ESTATÍSTICA  
DEPARTAMENTO DE ESTATÍSTICA  
PROVA II — MATD44 — 22/07/2025



**Professor:** Raydonal Ospina

**Regras:** Leia com atenção as perguntas. **Todas as questões devem ser detalhadas.** A prova deve ser claramente resolvida. Seja claro e organizado.

**1** ▶ **(Estimação por conglomerados)** Uma amostra de 5 fazendas na Bahia (conglomerados) foi selecionada por amostragem aleatória simples sem reposição de uma população de 100 fazendas para estimar a produção média de soja por hectare. Os dados da amostra são:

- ▷ Fazenda 1: 80 hectares, produção total 400 toneladas
- ▷ Fazenda 2: 120 hectares, produção total 540 toneladas
- ▷ Fazenda 3: 100 hectares, produção total 480 toneladas
- ▷ Fazenda 4: 90 hectares, produção total 450 toneladas
- ▷ Fazenda 5: 110 hectares, produção total 495 toneladas

Estime a produção total de soja por hectare ( $\hat{t}$ ). Forneça uma expressão para a estimativa da variância.

**2** ▶ **(Eficiência na estimação por conglomerados)** No contexto da amostragem por conglomerados, como o coeficiente de correlação intraclasse impacta a eficiência do estimador de Horvitz-Thompson para a média populacional?

**3** ▶ **(Amostragem com probabilidade proporcional ao tamanho (PPT))** Explique por

que a amostragem com probabilidade proporcional ao tamanho (PPT) é frequentemente mais eficiente do que a amostragem aleatória simples (AAS) sem reposição para estimar totais populacionais quando existe uma correlação positiva entre a variável de interesse ( $y$ ) e uma variável auxiliar de tamanho  $x$  (como área, densidade demográfica, etc)?

**4** ▶ **(Estimação de parâmetros lineares de totais em populações finitas)** Uma pesquisa foi feita para estimar a diferença no número de horas semanais gastas em redes sociais por jovens de dois cursos diferentes na UFBA (chamemos A e B). Os resultados das amostragens independentes são:

- ▷ Curso A:  $\hat{t}_A = 50.000$  horas;  $\hat{V}(\hat{t}_A) = 1.200.000$
- ▷ Curso B:  $\hat{t}_B = 42.000$  horas;  $\hat{V}(\hat{t}_B) = 900.000$ .

Estime o contraste entre os totais ( $D = 3t_A - 2t_B$ ) e a variância dessa estimativa.

**5** ▶ **(Estimação de parâmetros não-lineares de totais em populações finitas)** Estamos interessados em estimar a variância de um estimador do parâmetro populacional não-linear (produto de totais)  $\theta = f(t_y, t_z) = t_y \cdot t_z$ , através do produto de dois estimadores de Horvitz-Thompson de totais totais  $\hat{\theta} = \hat{t}_{\pi y} \cdot \hat{t}_{\pi z}$  das características populacionais  $y$  e  $z$ . Descreva passo a passo como poderia aproximar a variância de  $\hat{\theta}$ .

BOA PROVA