

Avaliação Parcial I

Disciplina: Microeconometria - IBM0288

Professor(a): Raphael Gouvea

Data: 29/09/2025

Tempo de Prova: 1h40

Nome do(a) Aluno(a):

Orientações:

1. Não é permitido a utilização de material de consulta.
2. Não é permitido o uso do celular ou computadores. Caso necessite ir ao banheiro, deixe seu celular em cima da mesa desligado ou em modo avião.

Questão 1 (1 ponto) Responda Falso ou Verdadeiro para as afirmações abaixo. Justifique brevemente sua escolha.

- (a) A hipótese de que $\text{Var}(u | X) = \sigma^2$ não é necessária para que os estimadores de mínimos quadrados ordinários sejam não-viesados e consistentes.
- (b) Se Z_i for uma variável omitida no modelo de regressão $y_i = \beta_0 + \beta_1 X_i + u_i$, sua omissão gerará viés de variável omitida **se e somente** $\text{corr}(X_i, Z_i) \neq 0$.
- (c) O IBGE pesquisa a cor ou raça da população brasileira com base na autodeclaração. Ou seja, quando questionada, a pessoa pode se declarar como preta, parda, branca, amarela ou indígena. Suponha que você queira utilizar essas informações como variáveis de controle em uma regressão. Seja R_1, \dots, R_5 variáveis indicadoras (*dummies*) de cada uma das raças. A especificação $y_i = \beta_1 R_{1i} + \beta_2 R_{2i} + \beta_3 R_{3i} + \beta_4 R_{4i} + \beta_5 R_{5i} + \beta_6 X_i + u_i$ é equivalente à especificação $y_i = \beta_0 + \beta_2 R_{2i} + \beta_3 R_{3i} + \beta_4 R_{4i} + \beta_5 R_{5i} + \beta_6 X_i + u_i$.

- (d) Suponha o seguinte modelo de dados em painel: $Y_{it} = \alpha_i + \lambda_t + \beta_1 X_{it} + u_{it}$. Devemos utilizar erros-padrão robustos ao fazermos inferência estatística pois o modelo é heterocedástico.

Questão 2 (3 pontos) Um pesquisador deseja avaliar o efeito de um programa que incentiva o uso de redes mosquiteiras ($D_i = 1$ se o indivíduo utiliza rede, $D_i = 0$ caso contrário) sobre o risco de ser infectado por malária (Y_i) (resultado contínuo, quanto maior, pior). Suponha que para cada indivíduo (i), ele observa a variável (Z_i) que indica se o indivíduo é jovem (até 60 anos) ou idoso (acima de 60 anos). O superscrito 1 indica o resultado potencial com tratamento e o superscrito 0 o resultado potencial sem tratamento.

Com base nesse contexto, responda:

- (a) Preencha os campos faltantes na tabela abaixo.

ID	Z_i	D_i	Y_i^1	Y_i^0	$Y_i^1 - Y_i^0$	Y_i
1	idoso		80	68		80
2	idoso	1		70	9	79
3	idoso	1	88		8	
4	jovem	0	76	70	6	70
5	jovem	0	87	90		
6	idoso	1	70	64	6	70
7	jovem	1	75	70	5	75
8	jovem	0	98	80		
9	jovem		90		10	80
10	jovem	0	89	80	9	80

- (b) Utilizando a tabela completa (ou seja, com os dados inseridos no item anterior), calcule o efeito médio do tratamento (ATE), o efeito médio do tratamento sobre os tratados (ATT) e o efeito médio do tratamento sobre os não tratados (ATU). Caso o resultado não seja inteiro, pode deixar o valor indicado na forma de fração.
- (c) Suponha que você não tenha nenhuma informação sobre qual o mecanismo de alocação dos indivíduos nos dois grupos, ou seja, você desconhece o processo gerador da variável D_i . Com essas informações você consegue identificar o efeito causal das redes mosquiteiras sobre o risco de infecção? Explique.
- (d) Agora assumamos que a variável D_i tenha sido aleatorizada. Um pesquisador, então, decide estimar o seguinte modelo de regressão: $Y_i = \beta_0 + \beta_1 D_i + \beta_2 Z_i + u_i$. O estimador de MQO $\hat{\beta}_1$ é um estimador não viesado do efeito populacional β_1 ? Justifique.

Tabela 1: Retornos da educação (erros-padrão robustos)

	ln salário
Constante	0.448 (0.025)
Educ (anos)	0.142 (0.003)
n	3000
R ²	0.406
SER (Root MSE)	0.43

- (e) Qual o motivo da inclusão da variável Z_i na regressão? Justifique. Caso julgue a inclusão de Z_i desnecessária, defenda esse ponto de vista.

Questão 3 (3 pontos) Suponha que um pesquisador queira estimar o efeito causal de anos de educação (Educ, em anos) sobre o salário em logaritmo ($\ln W$). **Habilidade é uma variável não observada**, que afeta salários e tende a ser positivamente correlacionada com educação (pessoas mais hábeis estudam mais). O pesquisador possui uma amostra de dados de corte transversal. Diante desse cenário o pesquisador estima a seguinte regressão $\ln W_i = \beta_0 + \beta_1 \text{Educ}_i + u_i$, cujos resultados se encontram na tabela 1.

- (a) Interprete o valor do coeficiente da variável Educ. O valor do coeficiente é estatisticamente diferente de zero?
- (b) Explique por que o modelo sofre de viés de variável omitida.
- (c) Qual o sinal esperado do viés de variável omitida? Explique sua resposta.
- (d) Qual problema o pesquisador pretende evitar ao usar erros-padrão robustos conforme descrito no título da tabela?
- (e) Algum tempo depois de estimar a regressão acima, o pesquisador encontra uma base de dados em painel em que cada indivíduo é observado por um período de 20 anos. Utilizando essa nova base de dados, seria possível eliminar o viés de variável omitida causado pelo fato de que habilidade é não observada? Seria necessário fazer algum ajuste para os erros-padrão do modelo a ser estimado? Justifique.

Questão 4 (3 pontos) Elabore uma pergunta de pesquisa em economia. Use sua criatividade. **A única regra é que não valem exemplos discutidos em sala de aula.**

- (a) Declare de forma clara e objetiva a pergunta que pretende responder.
- (b) Descreva como você responderia a essa pergunta utilizando uma das técnicas discutidas em sala: experimento aleatório, regressão ou dados em painel. Descreva também a base de dados que você teria (pode ser um exemplo fictício), ou seja, explique quais variáveis você observa, o tipo de dados que dispõe, período observado, etc.
- (c) Indique de forma explícita qual hipótese deve ser satisfeita para que seja possível interpretar o resultado como efeito causal.
- (d) Especifique que tipo de erro-padrão deve ser utilizado para a realização da inferência estatística dado o desenho de pesquisa descrito nos itens anteriores.
- (e) Descreva e justifique uma possível ameaça à validade interna do seu desenho de pesquisa e um possível problema para validade externa.