

Lista de Exercícios (Parte 1)

1. A tabela a seguir apresenta as pontuações relativas ao *ACT* (“vestibular”) e ao *GPA* (“coeficiente de rendimento” - CR) de 8 estudantes. A pontuação referente ao *GPA* é baseada em uma escala de 0 a 4 e os valores são arredondados para a primeira casa decimal.

Estudante	<i>GPA</i>	<i>ACT</i>
1	2,8	21
2	3,4	24
3	3,0	26
4	3,5	27
5	3,6	29
6	3,0	25
7	2,7	25
8	3,7	30

- a) Estime a relação entre *GPA* e *ACT* utilizando mínimos quadrados ordinários, ou seja, obtenha estimativas para o intercepto e a inclinação na seguinte equação:

$$\widehat{GPA} = \hat{\beta}_0 + \hat{\beta}_1 ACT.$$

Comente sobre a direção encontrada. O intercepto possui alguma interpretação útil? Explique. Quanto maior é o GPA previsto considerando um aumento de 5 pontos no ACT?

- b) Calcule os valores ajustados e os resíduos para cada observação. Verifique se a soma dos resíduos é (aproximadamente) zero.
- c) Qual o valor previsto para o GPA quando $ACT = 20$?
- d) Quanto da variação no GPA dos 8 estudantes é explicado pelo ACT? Explique.
2. A equação a seguir relaciona o valor do imóvel (*valor*) com sua distância para um incinerador de lixo (*dist*):

$$\ln(\widehat{valor}) = 9,40 + 0,312 \ln(dist) \\ n = 135 \text{ e } R^2 = 0,162.$$

- a) Interprete o coeficiente de $\ln(dist)$. O sinal referente à estimativa é o esperado?
- b) Interprete o coeficiente de determinação. Na sua interpretação, a distância, sozinha, explica bem o valor do imóvel?
- c) Quais outros fatores afetam o valor de um imóvel? Eles podem estar correlacionados com a distância do incinerador?
3. Utilizando dados sobre a renda anual (em reais) e o consumo (em reais) de 100 famílias, a seguinte equação é obtida:

$$\widehat{cons} = -124,84 + 0.853 \text{ renda}.$$

- a) Suponha que a renda seja medida em mil reais em vez de reais. Reescreva a regressão considerando essa mudança.
- b) Suponha agora que o consumo seja medido em mil reais em vez de reais. Reescreva a regressão considerando essa alteração.

4. Quais dos fatores a seguir podem causar viés nos estimadores de Mínimos Quadrados Ordinários?
- (a) Heterocedasticidade.
 - (b) Omissão de variável independente relevante.
 - (c) Correlação de 95% entre duas variáveis independentes especificadas no modelo.

5. Em um estudo relacionando o coeficiente de rendimento (GPA) na faculdade ao tempo gasto em diferentes atividades, os estudantes são perguntados sobre a quantidade de horas semanais despendidas em quatro atividades: estudos, sono, trabalho e lazer. A soma das horas nas quatro atividades deve ser 168.

(i) No modelo

$$GPA = \beta_0 + \beta_1 study + \beta_2 sleep + \beta_3 work + \beta_4 leisure + \varepsilon,$$

faz sentido manter fixos o sono, o trabalho e o lazer, enquanto se altera o estudo?

(ii) Explique por que esse modelo viola a Suposição A4 (ausência de colinearidade perfeita)?

(iii) Como você poderia reformular o modelo de modo que seus parâmetros tenham uma interpretação útil e satisfaçam a Suposição A4?

6. O salário inicial mediano para recém-graduados em Direito é determinado por

$$\ln(salary) = \beta_0 + \beta_1 LSAT + \beta_2 GPA + \beta_3 \ln(libvol) + \beta_4 \ln(cost) + \beta_5 rank + \varepsilon,$$

onde $LSAT$ é a pontuação mediana no exame LSAT da turma de formandos, GPA é o coeficiente de rendimento mediano da turma na faculdade, $libvol$ é o número de volumes na biblioteca da faculdade de Direito, $cost$ é o custo anual da faculdade de Direito, e $rank$ é a classificação da faculdade de Direito (sendo a melhor $rank = 1$).

(i) Explique por que esperamos que $\beta_5 \leq 0$.

(ii) Quais sinais são esperados para os demais coeficientes angulares? Justifique suas respostas.

(iii) Considerando um conjunto de dados, a seguinte equação foi estimada:

$$\begin{aligned} \widehat{\ln(salary)} &= 8,34 + 0,0047LSAT + 0,248GPA + 0,095\ln(libvol) \\ &+ 0,038\ln(cost) - 0,0033rank. \end{aligned}$$

Interprete o coeficiente da variável $\ln(libvol)$.

(iv) Você diria que é vantajoso frequentar uma faculdade de Direito melhor classificada? Quanto vale, em termos de salário inicial previsto, uma diferença de 20 posições no ranking?

7. A variável $rdintens$ representa o investimento em pesquisa e desenvolvimento (P&D) como porcentagem das vendas. As vendas são medidas em milhões de dólares. A variável $profmarg$ representa o lucro como porcentagem das vendas. Utilizando os dados de 32 empresas da indústria química, foi estimada a seguinte equação:

$$\widehat{rdintens} = \underset{(1,369)}{0,472} + \underset{(0,216)}{0,321} \ln(sales) + \underset{(0,046)}{0,050} profmarg.$$

(i) Interprete o coeficiente de $\ln(sales)$. Em particular, se as vendas aumentam em 10%, qual é a variação estimada em $rdintens$?

(ii) Teste a hipótese de que não há mudança em $rdintens$ com as vendas, contra a alternativa de que ela aumenta com as vendas. Faça o teste nos níveis de 5% e 10%.

(iii) A variável $profmarg$ tem efeito estatisticamente significativo sobre $rdintens$?

8. Considere uma equação para explicar os salários de CEOs em função das vendas anuais da empresa, do retorno sobre o patrimônio líquido (*roe*, em percentual) e do retorno das ações da empresa (*ros*, em percentual):

$$\ln(\textit{salary}) = \beta_0 + \beta_1 \ln(\textit{sales}) + \beta_2 \textit{roe} + \beta_3 \textit{ros} + \varepsilon,$$

(i) Em termos dos parâmetros do modelo, formule a hipótese nula de que, após controlar pelas vendas e pelo *roe*, o *ros* não tem efeito sobre o salário do CEO. Formule a alternativa de que um melhor desempenho no mercado de ações aumenta o salário do CEO.

(ii) Considerando um conjunto de dados, a seguinte equação foi estimada:

$$\ln(\textit{salary}) = 4,32 + 0,280 \ln(\textit{sales}) + 0,0174 \textit{roe} + 0,00024 \textit{ros} + \varepsilon,$$

(0,32)(0,035)(0,0041)(0,00054)

Qual é o aumento percentual previsto para o salário em caso de aumento de 50 pontos no *ros*?

(iii) Teste a hipótese nula de que o *ros* não tem efeito sobre o salário, contra a hipótese alternativa de que o *ros* tem efeito positivo. Realize o teste ao nível de significância de 10%.