

Econometria I Lista 5

Profa. Lorena Hakak

Entrega: 31/10/2022

Exercício 3

(i)

Neste caso, $\beta_1 < 0$, pois um aumento de 1 minuto de trabalho total por semana representará uma redução de β_1 minutos de sono por semana.

(ii)

Indivíduos com mais anos de educação terão mais responsabilidades para administrar durante boa parte de sua vida no mercado de trabalho, reduzindo o tempo disponível para dormir. $\beta_2 < 0$. Analogamente para idade, $\beta_3 < 0$.

No entanto, cabe pontuar que indivíduos com muitos anos de educação e idade mais elevada, por terem trabalhado por mais tempo e, provavelmente, adquirido estabilidade financeira, podem dispor de mais tempo para dormir. Para captar esse fenômeno, bastaria incluir no modelo econométrico as variáveis $educ^2$ e age^2 . O sinal de seus respectivos β ditaria a concavidade da curva de cada variável.

(iii)

Se alguém trabalha 5 horas a mais por semana, trabalha 300 minutos a mais. Logo, dorme $-0,148 * 300 = 44,4$ minutos a menos por semana.

(iv)

O sinal do parâmetro dos anos de educação é negativo, o que representa que indivíduos com mais anos de educação possuem menos minutos por semana para dormir. Comparando a magnitude de todas as variáveis independentes, o parâmetro da educação é o de maior impacto.

(v)

Como $R^2 = 0,113$, apenas 11,3% da variação amostral do tempo de sono por semana é explicada pelas variáveis independentes do modelo. Exposição a telas, trabalhar e estudar ao mesmo tempo, se é pai/mãe são variáveis que possivelmente explicam o tempo de sono semanal, e não necessariamente estão correlacionadas com *totwork*.

Exercício 6

Obs: Há um erro no enunciado da 6ª ed do livro texto, como pode ser verificado na versão inglesa. O correto é $\theta_1 = \beta_1 + \beta_2$.

(i)

$E[\theta_1] = E[\beta_1 + \beta_2] = E[\beta_1] + E[\beta_2]$; Cp,p as hipóteses RLM.1 RLM.4 são válidas, o estimador de MQO do parâmetro é o melhor. Logo, vem que: $E[\beta_1] + E[\beta_2] = \hat{\beta}_1 + \hat{\beta}_2 = \hat{\theta}_1$. Portanto, o estimador não é viesado.

(ii)

$$Var[\theta_1] = Var[\beta_1 + \beta_2]$$

$$Var[\theta_1] = Var[\beta_1] + Var[\beta_2] + 2Cov[\beta_1, \beta_2]$$

Sendo que:

$$Correl[\beta_1, \beta_2] = \frac{Cov[\beta_1, \beta_2]}{dp[\beta_1]dp[\beta_2]}$$

e

$$dp[\beta] = \sqrt{Var[\beta]}$$

então:

$$Var[\theta_1] = Var[\beta_1] + Var[\beta_2] + 2\sqrt{Var[\beta_1]Var[\beta_2]}Correl[\beta_1, \beta_2]$$

Exercício 7

Das opções de hipóteses apresentadas, apenas a omissão de uma variável importante poderia fazer com que o estimador de MQO seja viesado. A heteroscedastidade afeta a variância do estimador dos parâmetros, afetando os testes de hipótese destes; e o coeficiente de correlação de 0,95 entre duas variáveis independentes não faz com que os estimadores de MQO tenham problemas.

Exercício 8

$$avg\hat{prod} = \hat{\beta}_0 + \hat{\beta}_1 avgtrain + \hat{\beta}_2 avgabil$$

$$avg\tilde{prod} = \tilde{\beta}_0 + \tilde{\beta}_1 avgtrain$$

$$avgtrain = \tilde{\delta}_1 avgabil + \mu_1$$

Como $\tilde{\beta}_1 = \hat{\beta}_1 + \delta\tilde{ta}_1\hat{\beta}_2$, o viés de $\tilde{\beta}_1$ é $\delta\tilde{ta}_1\hat{\beta}_2$. Como $\tilde{\delta}_1$ é a covariância amostral de $avgtrain$ e $avgabil$, seu sinal acompanhará o da correlação das duas variáveis independentes – neste caso, será negativo.

Exercício 14

$$\hat{\sigma} = \frac{SQR}{(n - k - 1)}$$

Com n sendo o número de variáveis do modelo e k o número de parâmetros. Adicionar uma variável ao modelo diminui o SQR, mas também diminui o denominador da razão (graus de liberdade), não sendo possível garantir se a inclusão/remoção de uma variável aumenta ou não o erro padrão da regressão.

Observando primeiro o caso em que x_2 não tem significância estatística, ou seja, $\beta_2 = 0$. Se isso for verdade, então a inclusão não irá afetar o poder de explicação da regressão, sendo assim, R^2 não irá alterar e por consequência FIV também será mantido. Entretanto, inserir essa variável no modelo irá fazer com que a variância, fazendo com que:

$$ep(\hat{\beta}_1) > ep(\tilde{\beta}_1)$$

Agora, caso x_2 seja uma variável estatisticamente relevante ($\beta_2 \neq 0$) então excluir no modelo irá causar um problema do viés de variável omitida, causando uma variância maior no modelo. Logo:

$$ep(\hat{\beta}_1) < ep(\tilde{\beta}_1)$$

Exercício 15

(i)

Na primeira regressão possui 351 graus de liberdade e, na segunda, 350. O erro padrão da regressão do segundo modelo é menor do que o primeiro pois a inclusão de uma variável adicional, apesar de reduzir o grau de liberdade, aumentou a capacidade preditiva do modelo, reduzindo o SQR e, consequentemente, o erro padrão da regressão.

(i)

Faz sentido haver uma correlação amostral razoável entre as variáveis, pois jogadores que duram mais na liga principal vão apresentar um desempenho em rebatidas melhor do que os que possuem menos tempo.

$$FIV = \frac{1}{1 - R^2} = \frac{1}{1 - 0,487} = \frac{1}{0,513} \approx 1,95$$

Pela fórmula, é possível observar que, se $R^2 \rightarrow 1$, $FIV \rightarrow \infty$. Caso contrário, se $R^2 \rightarrow 0$ isto é, se a correlação amostral for baixa, $FIV \rightarrow \infty$. Pelo valor calculado, concluímos que há baixa colinearidade entre as variáveis independentes.

(iii)

Como *rbisyr* é variável relevante para o modelo, sua inclusão torna a estimação do impacto de years mais eficiente, logo com menor erro padrão.