1. a) $H_0: \beta_1 = 0$

$H_A: \beta_1 > 0$

Sob H_0

$$t_{obs} = \frac{\hat{\beta}_1 - \beta_1^{\text{sob } H_0}}{ep(\hat{\beta}_1)} \sim t_{(30)} \qquad t_{obs} = \frac{2}{1,2} = 1,67$$

$t_{crítico} = 1,697$

$RC = [1,697; +\infty)$ $t_{obs} \notin RC \Rightarrow$ não há evidências p/ rejeitar H_0.

Não há evidências que o efeito seja positivo!

b) $\underset{\sim}{\hat{\beta}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1} \cdot \underset{\sim}{X}'\underset{\sim}{y}$ tanto X e y possam a ser multiplicados por uma constante θ.

$$\underset{\sim}{\hat{\beta}}^* = [(\theta \underset{\sim}{X})'\theta \underset{\sim}{X}]^{-1} \cdot (\theta \underset{\sim}{X})' \cdot (\theta \underset{\sim}{y})$$

$$\underset{\sim}{\hat{\beta}}^* = (\theta^2 \underset{\sim}{X}'\underset{\sim}{X})^{-1} \cdot \theta^2 \cdot \underset{\sim}{X}'\underset{\sim}{y}$$

$$\underset{\sim}{\hat{\beta}}^* = \frac{1}{\theta^2} \cdot \theta^2 \cdot (\underset{\sim}{X}'\underset{\sim}{X})^{-1} \cdot \underset{\sim}{X}'\underset{\sim}{y} = \underset{\sim}{\hat{\beta}}$$

Portanto, não haveria mudança no resultado.

c) $H_0: \beta_1 = 3$

$H_A: \beta_1 \neq 3$

$$t_{obs} = \frac{2-3}{1,2} \sim t_{(30)}$$

$t_{obs} = -0,833$ $t_{crítico} = 2,042$

$RC = (-\infty; -2,042] \cup [2,042; +\infty)$

$t_{obs} \notin RC \Rightarrow$ não há evidências p/ rejeitar H_0.

2. a) 1 ano adicional de experiência acarreta um aumento esperado de em média 0,9 % no salário, mantendo-se os demais variáveis constantes.

b) Para o homem: 9,3%

Para a mulher: 8,9%

c) H_0: Retorno esperado de 1 ano adicional de educação é o mesmo, em média, para homens e mulheres, tudo o mais constante.

H_A: Efeito de 1 ano a mais de educação no salário não é, em média, o mesmo para homens e mulheres, tudo o mais constante.

$t_{obs} = \dfrac{-0{,}004}{0{,}014} \sim t_{(55)}$ $t_{obs} = -0{,}2857$

$t_{crítico} = 2$ $RC = (-\infty; -2] \cup [+2; +\infty)$

$t_{obs} \notin RC \rightarrow$ não há evidências p/ rejeitar H_0.

3. a) Variação de 1 unidade em educ acarreta um aumento esperado de $(100 \cdot \beta_1)\%$ no salário, tudo o mais constante.

b) Variação percentual no QI acarreta uma variação esperada de $\beta_4\%$ no salário, tudo o mais constante.

c) $\beta_2 > 0 \quad \beta_3 < 0$

É esperado que o aumento de anos de experiência impacte positivamente o salário, mas de uma forma cada vez menor. À medida que os anos de experiência crescem, o impacto de um ano adicional no salário é cada vez menor.

d) $$\frac{\partial \log(\text{salário}_i)}{\partial \text{exper}_i} = \beta_2 + 2 \cdot \beta_3 \cdot \text{exper}_i$$

1 ano adicional de experiência acarreta uma variação esperada de $100(\beta_2 + 2\beta_3 \text{exper}_i)\%$ no salário.

e) Ponto de máximo (mínimo). Indica a experiência ótima.

f) $$\frac{\partial^2 \log(\text{salário}_i)}{\partial \text{exper}_i^2} = 2\beta_3$$

Fornece a concavidade da função.
Indica se os retornos marginais são crescentes ou decrescentes.

g) Ocorre apenas quando $\beta_3 = 0$, o que implica que o efeito marginal da experiência no log do salário será constante (β_2)

4. a) Custo médio esperado será de 27,2 milhares de reais

b) Homens com até 60 anos

$\hat{y}_i = 44,75 - 0,45 X_2$

Mulheres com até 60 anos

$\hat{y}_i = 33,95 - 0,27 X_2$

Homens e mulheres com mais do que 60 anos

$\hat{y}_i = 4,55 + 0,22 X_2$

Pode-se notar que o decréscimo esperado no custo médio é maior para os homens até 60 anos do que para mulheres até 60 anos quando-se aumenta 1 ano na idade. (Reta c/ maior inclinação) Para pessoas com mais que 60 anos, um ano a mais de idade acarreta um aumento esperado no custo da seguradora c/ danos materiais (Reta c/ inclinação positiva).

) $H_0: \beta_6 = 0$ decréscimo esperado é o mesmo p/ ambos

$H_A: \beta_6 > 0$ decréscimo maior p/ homens

Sob H_0

$$t_{obs} = \frac{0,18}{ep(\hat{\beta}_6)} \sim t_{(84)}$$

$t_{crítico} = 2,37$

$RC = [2,37; +\infty)$

Rejeitar H_0 se $t_{obs} \geq 2,37$

Questão 5

a) Maximizar o R^2 é um critério errado. A conclusão do aluno seria que todas as variáveis deveriam ser colocadas no modelo.

b) Colocar uma variável a mais no modelo nunca pode piorar o ajuste que já existia sem ela. No pior dos casos, se a nova variável não tiver nenhuma capacidade de explicação, o R^2 permanecerá o mesmo de antes.

c) $$\bar{R}^2 = 1 - \frac{n-1}{n-(k+1)}(1-R^2) \qquad (1)$$

Sabemos que $R^2 = 1 - \frac{SSR}{SST} \Leftrightarrow \frac{SSR}{SST} = 1 - R^2$, substituindo em (1)

$$\bar{R}^2 = 1 - \frac{n-1}{n-(k+1)} \cdot \frac{SSR}{SST}$$

$$\bar{R}^2 = 1 - \frac{SSR/(n-(k+1))}{SST/(n-1)}$$

$$\bar{R}^2 = 1 - \frac{\hat{\sigma}^2}{S_y^2}$$

Note que para maximizar $\bar{R}^2$ temos que minimizar $\hat{\sigma}^2$. Vale lembrar que S_y^2 não varia com diferentes modelos. Minimizar a variância dos resíduos é igual a minimizar o ~~desv~~ desvio padrão destes, que por sua vez é o erro padrão da regressão.

d) Por esse critério, uma variável é colocada no modelo se:

$$\bar{R}^2 > \bar{R}^2_R$$

↳ R^2-ajustado do modelo Restrito (sem a nova variável)

↳ R^2-ajustado do modelo Irrestrito (com a nova variável)

$$1 - \frac{n-1}{n-(k+1+1)}(1-R^2) > 1 - \frac{n-1}{n-(k+1)}(1-R^2_R)$$

$$\frac{n-1}{n-(k+1+1)}(1-R^2) < \frac{n-1}{n-(k+1)}(1-R^2_R)$$

$$(1-R^2) < \frac{n-(k+1)-1}{n-(k+1)}(1-R^2_R)$$

$$(1-R^2) < (1-R^2_R) - \frac{(1-R^2_R)}{n-(k+1)}$$

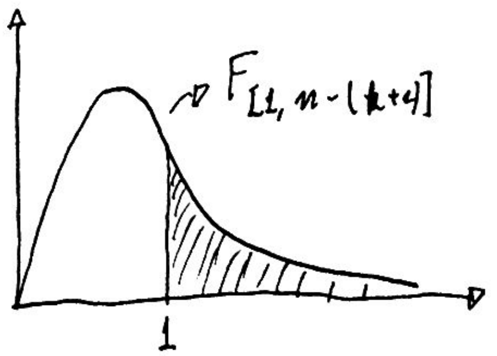
$$R^2 - R^2_R > \frac{(1-R^2_R)}{n-(k+1)}$$

$$\frac{(R^2-R^2_R)}{(1-R^2)/(n-(k+1))} > 1$$

Portanto o R^2-ajustado só aumenta com uma nova variável x_j se a estatística-F do teste $H_0: \beta_j = 0$ for maior do que 1

$$F_{obs} > 1$$

O nível de significância do teste é dado pela área:



Com a ajuda do excel, temos que essa área é igual a 0,3173, que é um nível de significância muito maior que os usualmente usados.

Questão 6

a) A forma funcional sugerida não é linear nos parâmetros, ou seja, os consultores não podem aplicar o estimador de Mínimos Quadrados ordinários. Para contornar o ~~pr~~ problema, eles podem aplicar a transformação log, que resulta em:

$$\log(Q_i) = \beta_0 + \beta_1 \log(K_i) + \beta_2 \log(L_i) + \varepsilon_i \qquad (1)$$

Este último é linear nos parâmetros e pode ser estimado por MQO.

b) $H_0 : \beta_0 = 0$ $\quad t_{obs} = \dfrac{0,701}{0,415} = 1,689$ $\quad$ não rejeito H_0

$H_1 : \beta_0 \neq 0$

$H_0 : \beta_1 = 0$ $\quad t_{obs} = \dfrac{0,242}{0,110} = 2,2$ $\quad$ rejeito H_0

$H_1 : \beta_1 \neq 0$

$H_0 : \beta_2 = 0$ $\quad t_{obs} = \dfrac{0,756}{0,091} = 8,307$ $\quad$ rejeito H_0

$H_1 : \beta_2 \neq 0$

$t_{crit,\ 23,\ 5\%} = 2,068$ $\quad RC =]-\infty; -2,068] \cup [2,068; +\infty[$

Interpretação: Um aumento de 1% em K gera um aumento médio esperado de 24,2% na quantidade produzida.

c) $H_0: \beta_1 + \beta_2 = 1$

Sob H_0 temos que $\beta_1 = 1 - \beta_2$, então a forma funcional da função de produção fica:

$$Q_i = e^{\beta_0} K_i^{1-\beta_2} L_i^{\beta_2} e^{\varepsilon_i}$$

podemos dividir os dois lados da equação por K_i

$$\frac{Q_i}{K_i} = e^{\beta_0} \left(\frac{L_i}{K_i}\right)^{\beta_2} e^{\varepsilon_i}$$

aplicando o log

$$\log\left(\frac{Q_i}{K_i}\right) = \beta_0 + \beta_2 \log\left(\frac{L_i}{K_i}\right) + \varepsilon_i \qquad (2)$$

Então temos (1) como modelo Irrestrito e (2) como o restrito.

$$F_{SSR} = \frac{(1,825652 - 1,825544)/1}{1,825544/23} = 0,001$$

$$F_{R^2} = \frac{(0,95688 - 0,751397)/1}{(1-0,95688)/23} = 109,628$$

Note que a estatística-F é diferente para cada maneira de escrevê-la, mas sabemos que isso não deveria acontecer.

Neste caso, isso ocorre porque a variável dependente é diferente nos dois modelos. Por isso, o SST é diferente em cada modelo, o que torna os R^2 não comparáveis entre os modelos. Por isso o teste-F deve ser conduzido com a estatística baseada em SSR. Como $F_{SSR} = 0{,}001$ não rejeitamos a hipótese de retornos constantes de ~~[illegible]~~ escala.

Questão 7

a) Sabemos que

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$$

Como só existem 2 observações em que $x_i = 1$, dentre as 3141, temos que $\bar{x} \approx 0$ e $\hat{y}_i \approx \bar{y}$. Isso faz com que o numerador seja muito pequeno.

b) $$t_{\beta_0} = \frac{50,3}{16,4} = 3,067$$

$$t_{\beta_1} = \frac{14,1}{5,2} = 2,711$$

$\hat{\beta}_0$: Al gore teve, em média, 50,3% dos votos nos municípios que não tinham problema na contagem dos votos

$\hat{\beta}_1$: Al gore teve, em média, 14,1 p.p. a mais de porcentagem de votos ~~[illegible]~~ nos municípios em que houve problema de contagem, em relação à média dos demais municípios.

c) O fato da média nos municípios que tiveram problema de contagem não anula a possibilidade que Al gore tenha recebido menos votos do que deveria.

Provavelmente, o partido Republicano teria gostado dessa análise, pois ela diz contra a alegação do Al gore.

d) $\hat{\beta}_0$: Nos municípios em que não houve problema de contagem e que Bill Clinton não recebeu votos em 1996, Al gore teve, em média, 40,1% dos votos.

$\hat{\beta}_1$: Nos municípios com problema de contagem, Al Gore teve, em média, 2,9 p.p. a menos de votos.

$\hat{\beta}_2$: Para cada p.p. a mais de votos que Clinton teve em 1996, em média, aumentaria 0,93 p.p. de voto para Al gore.

A estimativa de β_1 corrobora a alegação do Partido Democrata.