

Métodos de Previsão em Negócios
Soluções das Atividades

Lupércio França Bessegato- lupercio@est.ufmg.br
Instituto de Educação Continuada/PUCMinas
Curso de Especialização em Economia Empresarial

Resumo. *Solução das atividades extra-classe da disciplina.*

Keywords: *Econometria, modelos de regressão, métodos computacionais, ensino*

1. Distribuição Normal de Probabilidades

1.1 Atividade # 1

- Ativos:
Ativo $X \sim N(10, 1^2)$
Ativo $Y \sim N(10, 2^2)$
Ativo $V \sim N(10, 3^2)$
Ativo $W \sim N(10, 0,5^2)$
 - Carteiras:
Carteira 1: $P_1 = 1.000W$
Carteira 2: $P_2 = 500W + 500V$
Carteira 3: $P_3 = 250X + 250Y + 250V + 250W$
 - Retornos das Carteiras:
Carteira 1: $\mu_1 = E(P_1) = 1.000 \times 10 = \10.000
Carteira 2: $\mu_2 = E(P_2) = 500 \times 10 + 500 \times 10 = \10.000
Carteira 3: $\mu_3 = E(P_3) = 250 \times 10 + 250 \times 10 + 250 \times 10 + 250 \times 10 = \10.000
 - Risco (volatilidade) das Carteiras:
Supondo os ativos independentes
Carteira 1: $\sigma_1^2 = Var(P_1) = 1.000^2 \times 1^2 = 1.000 \Rightarrow \sigma_1 = \sqrt{1.000} = \1.000
Carteira 2: $\sigma_2^2 = Var(P_2) = 500^2 \times 0,5^2 + 500^2 \times 3^2 = 2.312.500 \Rightarrow \sigma_2 = \sqrt{2.312.500} = \$1.520,69$
Carteira 3: $\sigma_3^2 = Var(P_3) = 250^2 \times 1^2 + 250^2 \times 2^2 + 250^2 \times 3^2 + 250^2 \times 0,5^2 = 890.625 \Rightarrow \sigma_3 = \sqrt{890.625} = \$943,73$
1. $P\{P_i > 12.000\}$:
- carteira 1: $P\{Z_1 > \frac{12.000-10.000}{1.000}\} = P\{Z_1 > 2\} = 0,0228$
carteira 2: $P\{Z_2 > \frac{12.000-10.000}{1.520,69}\} = P\{Z_1 > 1,32\} = 0,0934$
carteira 3: $P\{Z_2 > \frac{12.000-10.000}{943,73}\} = P\{Z_1 > 2,12\} = 0,0170$
2. $P\{P_i > 5.000\}$:
- carteira 1: $P\{Z_1 > \frac{5.000-10.000}{1.000}\} = P\{Z_1 > -5\} \approx 1$
carteira 2: $P\{Z_2 > \frac{5.000-10.000}{1.520,69}\} = P\{Z_1 > -2,88\} = 0,9980$
carteira 3: $P\{Z_2 > \frac{5.000-10.000}{943,73}\} = P\{Z_1 > -3,08\} = 0,9990$

2. Teste de Hipóteses e Intervalo de Confiança

2.1 Atividade # 2

1. Estatística descritivas:

Variável	Média amostral	Desvio padrão amostral
Idade (x)	38,933	13,350
Desemprego (y)	16,800	10,345

2. Intervalo com 95% de confiança para Idade média:

$$\begin{aligned} \bar{x} \pm t_{n-1;0,025} \times \frac{s}{\sqrt{n}} \\ 38,933 \pm 2,14479 \times \frac{13,350}{\sqrt{15}} \\ [31,54; 46,33] \end{aligned}$$

3. Teste de hipóteses para duração média do desemprego:

$$H_0 : \mu \leq 14,6 \text{ vs. } H_1 : \mu > 14,6$$

Nível de significância $\alpha = 0,01$

Valor crítico do teste: $t_c = t_{15-1;0,01} = 2,62449$

Estatística de teste:

$$T = \frac{16,80 - 14,6}{10,345/\sqrt{15}} = 0,82364$$

Teste: $T = 0,824 < t_c = 2,624$

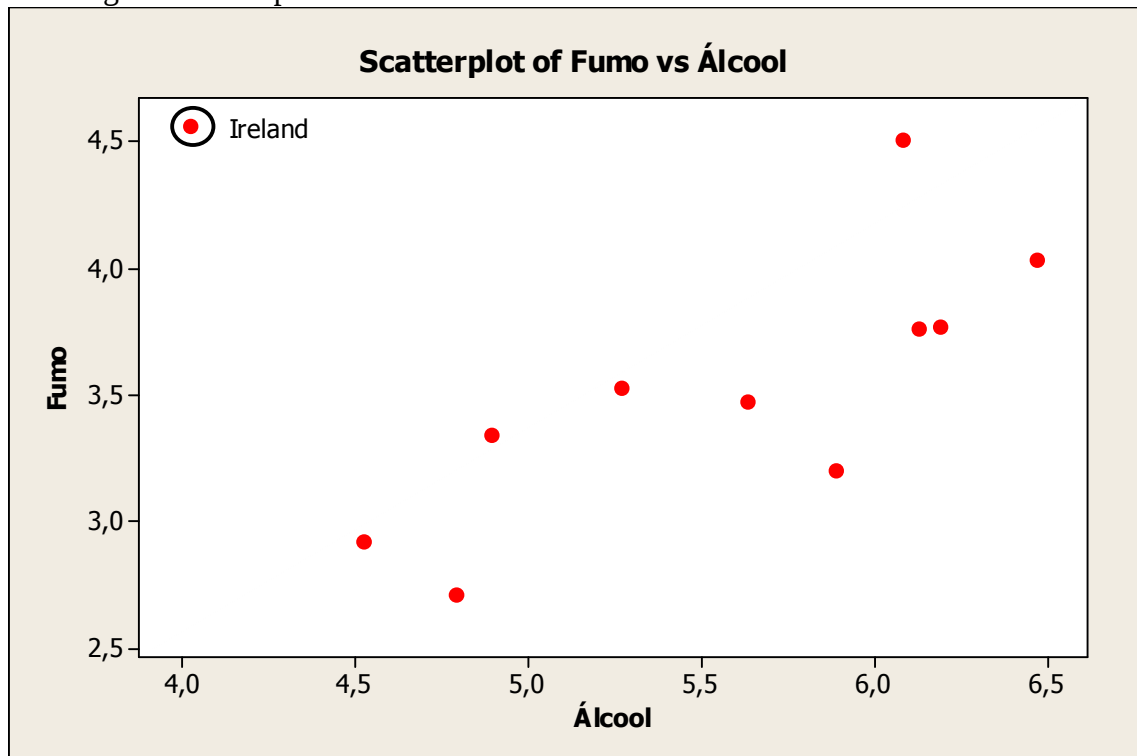
Decisão: não há evidências para se rejeitar H_0 , ou seja, a amostra não fornece evidências para concluir que o índice de desemprego da área de atuação do sindicato seja maior que a média nacional

4. : Coeficiente de correlação: $r = 0,80414$

Há um forte grau de associação linear entre as variáveis, mas nada podemos afirmar sobre uma relação causal entre idade e tempo de desemprego

Atividade 3 – Correlação

1- Diagrama de Dispersão



2- Padrão do gráfico e desvios importantes:

Os pontos do gráfico mostram uma correlação positiva (mais despesa com fumo → mais despesa com álcool ou vice-versa). Ireland é um ponto discrepante do comportamento global

3- Correlação:

MTB > correlation álcool fumo

Correlations: Álcool; Fumo

Pearson correlation of Álcool and Fumo = **0,224 (correlação fraca)**
P-Value = 0,509

```
MTB > correlation álcool fumo;  
SUBC> exclude;  
SUBC> rows 11.
```

Correlations: Álcool; Fumo

Excluding specified rows: 11
1 rows excluded

Pearson correlation of Álcool and Fumo = **0,784 (correlação forte)**
P-Value = 0,007

A correlação sem Ireland é mais forte. De maneira diversa às outras regiões, as despesas domésticas com fumo são maiores do que as despesas com álcool (no mesmo nível das maiores)

Regressão Linear Simples

Atividade 4

Item a:

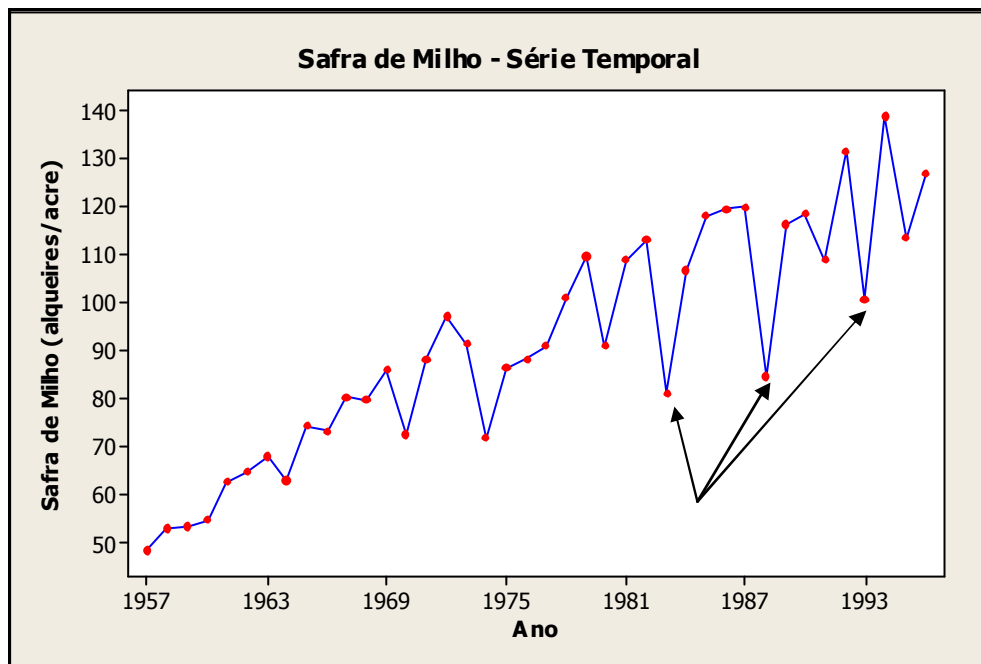


Figura 1- Safr de Milho vs. Ano

A tendência é crescente embora haja anos com safras muito baixas em relação àquelas dos anos próximos.

Há uma variabilidade crescente após a década de 70 (série ficou “mal-comportada”)

Item b:

Regression Analysis: milho versus ano

The regression equation is
 $\text{milho} = -3545 + 1,84 \text{ ano}$

Predictor	Coef	SE Coef	T	P
Constant	-3544,6	278,4	-12,73	0,000
ano	1,8396	0,1408	13,06	0,000

S = 10,2824 R-Sq = 81,8% R-Sq(adj) = 81,3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	18038	18038	170,60	0,000
Residual Error	38	4018	106		
Total	39	22055			

Unusual Observations

Obs	ano	milho	Fit	SE Fit	Residual	St Resid
27	1983	81,10	103,29	1,87	-22,19	-2,19R
32	1988	84,60	112,49	2,29	-27,89	-2,78R
37	1993	100,70	121,69	2,84	-20,99	-2,12R

R denotes an observation with a large standardized residual.

Regressão Significante

Equação de Regressão:

$milho = -3.545 + 1,8396 ano$	$R^2 = 81,8\%$
$(278,4) \quad (0,1408)$	$\hat{s}^2 = 106$

A produtividade da safra de milho aumenta 1,84 alqueires/acre por ano

Item c:

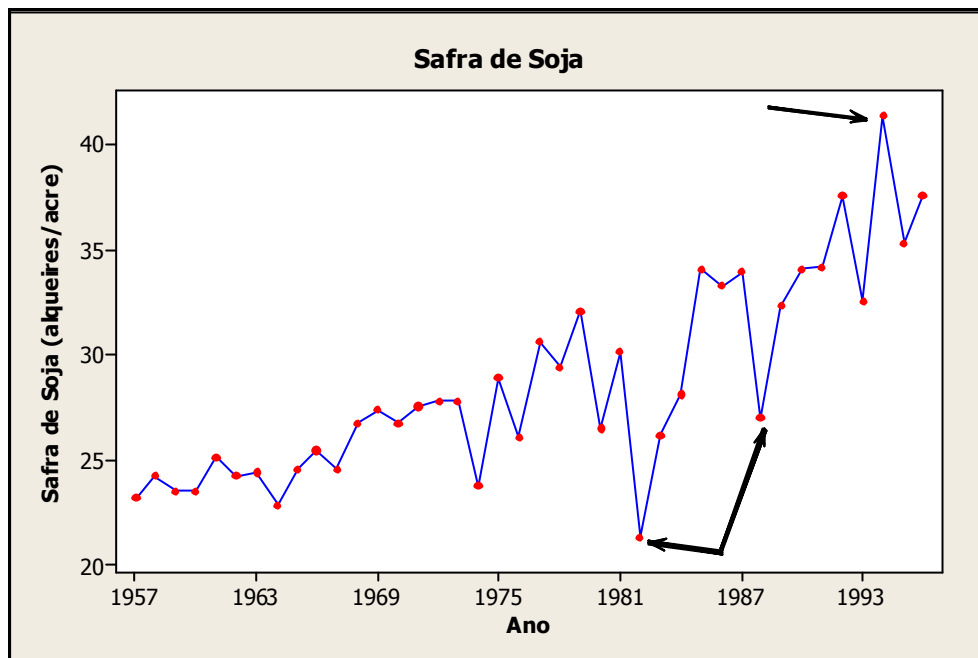


Figura 2 - Safra de Soja vs. Ano

Regression Analysis: soja versus ano					
The regression equation is					
soja = - 639 + 0,338 ano					
Predictor	Coef	SE Coef	T	P	
Constant	- 638,92	74,61	-8,56	0,000	
ano	0,33775	0,03775	8,95	0,000	
S = 2,75576 R-Sq = 67,8% R-Sq(adj) = 67,0%					
Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	608,02	608,02	80,06	0,000
Residual Error	38	288,58	7,59		
Total	39	896,60			
Unusual Observations					
Obs	ano	soja	Fit	SE Fit	Residual
26	1982	21,300	30,498	0,483	-9,198
32	1988	27,000	32,524	0,615	-5,524
38	1994	41,400	34,551	0,791	6,849
St Resid					
					-3,39R
					-2,06R
					2,59R
R denotes an observation with a large standardized residual.					

Regressão Significante

Equação de Regressão:

$soja = -638,92 + 0,33775 ano$	$R^2 = 67,8\%$
$(74,61) \quad (0,03775)$	$\hat{s}^2 = 7,59$

Testes de Significância para o coeficiente de inclinação:

Nível de significância: 5%

Estatística T = 8,95

Limite região crítica: $t_c = t_{0,025; 38} = 2,02439$

Teste: $T > t_c \rightarrow$ rejeita-se H_0 , ou seja, a regressão é significativa

A produtividade da safra de soja aumenta 0,338 alqueires/acre por ano

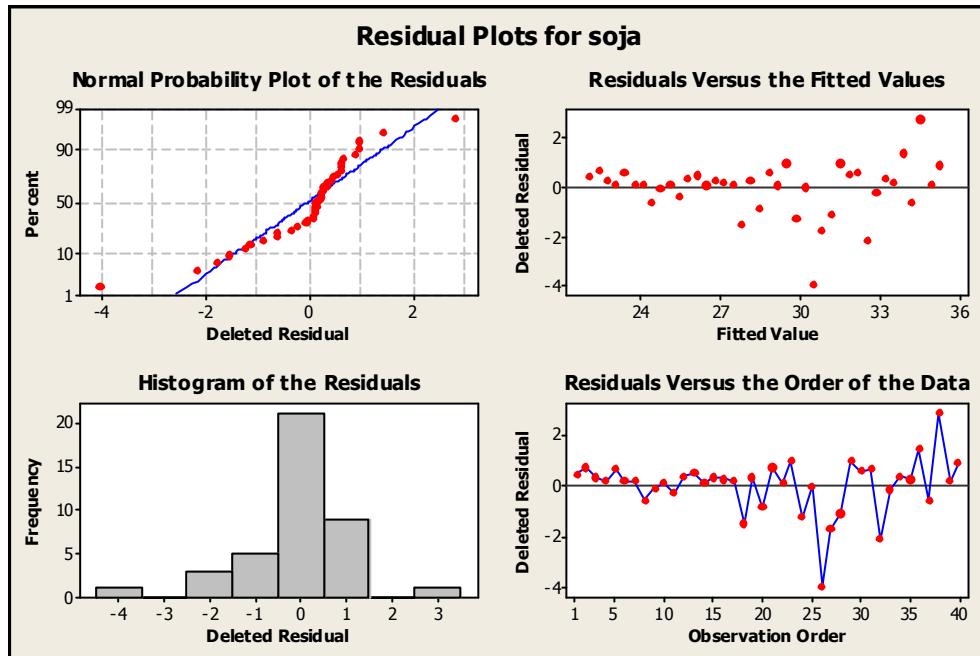


Figura 3 - Gráfico de Resíduos Safra de Soja

Resíduos:

Hipóteses violadas:

- Normalidade
- Homocedasticidade
- Correlação

Os resíduos devem estar correlacionados por se tratar de série temporal

Conclusões:

- A partir da década de 70, há um aumento significativo na variabilidade;
- O modelo de regressão linear parece não ser adequado para explicar a produtividade da safra de soja

Item d

Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	145,60	4,46	(136,57; 154,64)	(122,91; 168,29)X

X denotes a point that is an outlier in the predictors.

Values of Predictors for New Observations

New Obs	ano
1	2006

Obs.: Não é apropriada, pois, trata-se de uma extrapolação

Item e:

Gráfico simultâneo:

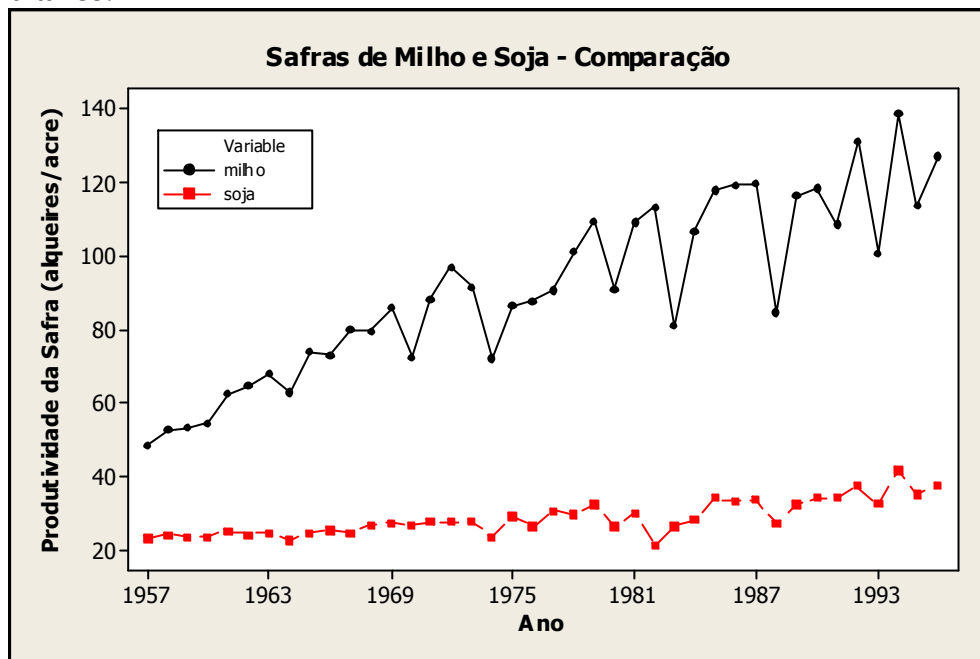


Figura 4 - Comparação anual das safras de milho e soja

Associação entre as duas safras:

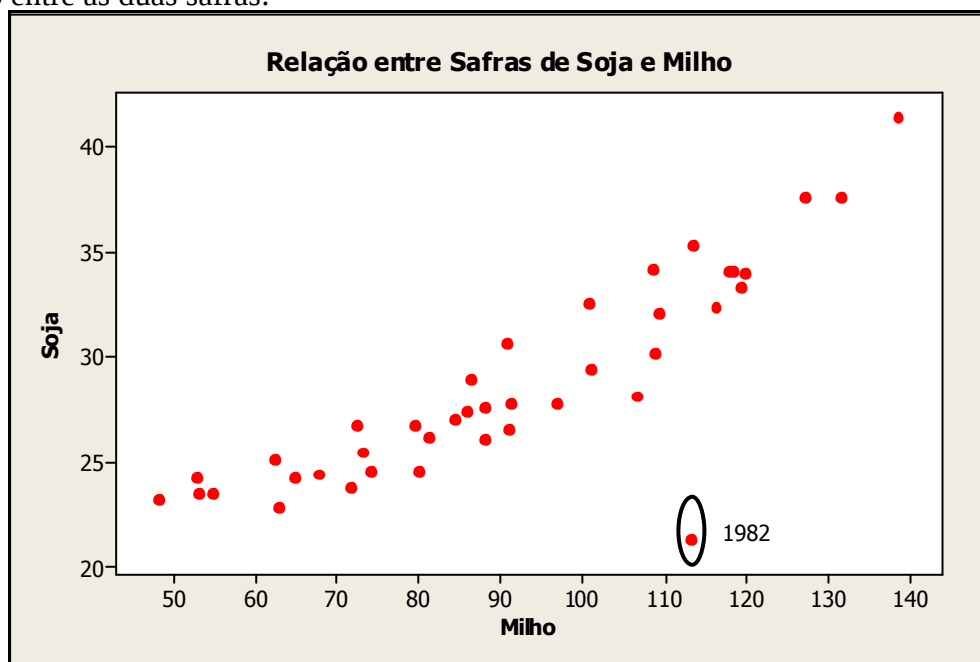


Figura 5 - Associação entre as safras de soja e milho

Correlação entre as safras:

```
MTB > correlation milho soja
```

Correlations: milho; soja

Pearson correlation of milho and soja = 0,858

P-Value = 0,000

Há indícios de uma forte associação linear entre as safras, mas nada podemos afirmar sobre causalidade.

Safra de soja explicada pela safra de milho:

Regression Analysis: soja versus milho

The regression equation is
soja = 12,8 + 0,173 milho

Predictor	Coef	SE Coef	T	P
Constant	12,844	1,586	8,10	0,000
milho	0,17295	0,01681	10,29	0,000

S = 2,49694 R-Sq = 73,6% R-Sq(adj) = 72,9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	659,68	659,68	105,81	0,000
Residual Error	38	236,92	6,23		
Total	39	896,60			

Unusual Observations

Obs	milho	soja	Fit	SE Fit	Residual	St Resid
26	113	21,300	32,421	0,539	-11,121	-4,56R

R denotes an observation with a large standardized residual.

Equação da regressão:
$$\text{soja} = 12,884 + 0,17295\text{milho}$$
$$(1,586) \quad (0,01681) \quad R^2 = 73,6\%$$
$$\hat{s}^2 = 6,23$$

Testes de Significância para o coeficiente de inclinação:

Nível de significância: 5%

Estatística T = 10,29

Limite região crítica: $t_c = t_{0,025; 38} = 2,02439$

Teste: $T > t_c \rightarrow$ rejeita-se H_0 , ou seja, a regressão é significativa e há correlação entre as safras

A safra de milho aumentou 0,338 alqueires/acre por ano

Resíduos da Regressão (soja = f(milho)) vs. tempo:

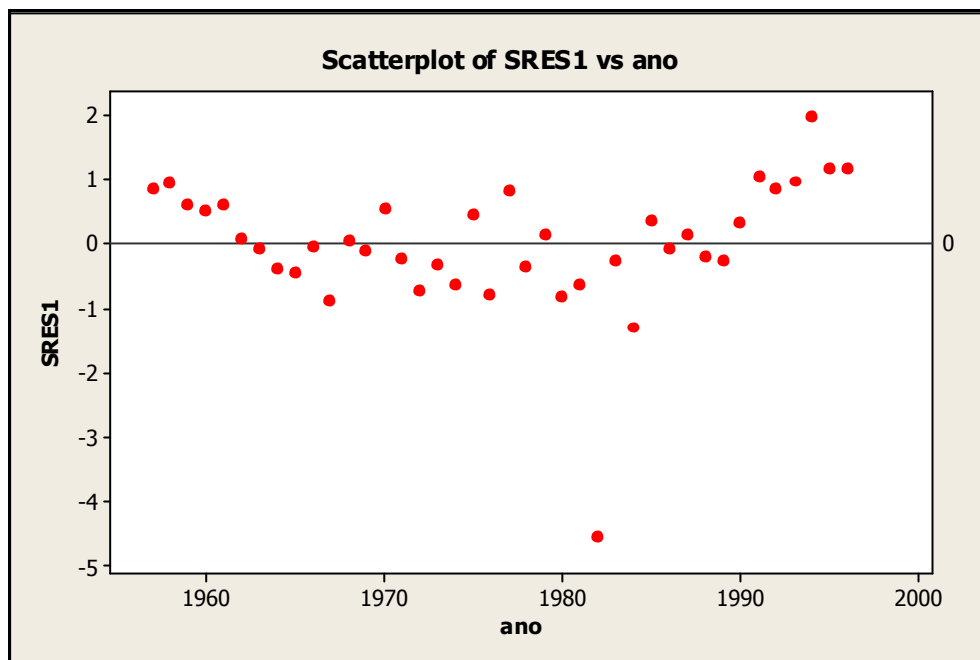


Figura 6 - Resíduos da regressão vs. ano

Há um padrão identificável nos resíduos (positivos no tempos anteriores, negativos no período médio e positivos novamente nos períodos posteriores). É um indício de que o tempo está omitido do modelo.

Questão f:

Transformação dos dados da safra de milho:

Soja vs. log(milho)

Regression Analysis: soja versus log(milho)

The regression equation is
soja = - 35,3 + 14,3 log(milho)

Predictor	Coef	SE Coef	T	P
Constant	-35,251	7,250	-4,86	0,000
log(milho)	14,264	1,616	8,83	0,000

S = 2,78086 R-Sq = 67,2% R-Sq(adj) = 66,4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	602,73	602,73	77,94	0,000
Residual Error	38	293,86	7,73		
Total	39	896,60			

Unusual Observations

Obs	log(milho)	soja	Fit	SE Fit	Residual	St Resid
26	4,73	21,300	32,207	0,597	-10,907	-4,02R
38	4,93	41,400	35,095	0,853	6,305	2,38R

R denotes an observation with a large standardized residual.

Equação de regressão:

Transformada:

$$\begin{array}{lcl} \text{soja} = -35,251 + 14,264 \log(\text{milho}) & R^2 = 67,2\% \\ (7,250) & (1,616) & \hat{s}^2 = 7,73 \end{array}$$

Original:

$$\begin{array}{lcl} \text{soja} = 12,884 + 0,17295 \text{milho} & R^2 = 73,6\% \\ (1,586) & (0,01681) & \hat{s}^2 = 6,23 \end{array}$$

O modelo transformado, embora significativo, mostra-se menos adequado que o original

Log(milho) vs. tempo:

Regression Analysis: log(milho) versus ano

The regression equation is
 $\log(\text{milho}) = -37,6 + 0,0213 \text{ ano}$

Predictor	Coef	SE Coef	T	P
Constant	-37,633	3,235	-11,63	0,000
ano	0,021306	0,001637	13,02	0,000

S = 0,119497 R-Sq = 81,7% R-Sq(adj) = 81,2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	2,4196	2,4196	169,45	0,000
Residual Error	38	0,5426	0,0143		
Total	39	2,9622			

Unusual Observations

Obs	ano	log(milho)	Fit	SE Fit	Residual	St Resid
32	1988	4,4379	4,7241	0,0267	-0,2862	-2,46R

R denotes an observation with a large standardized residual.

Equação de regressão:

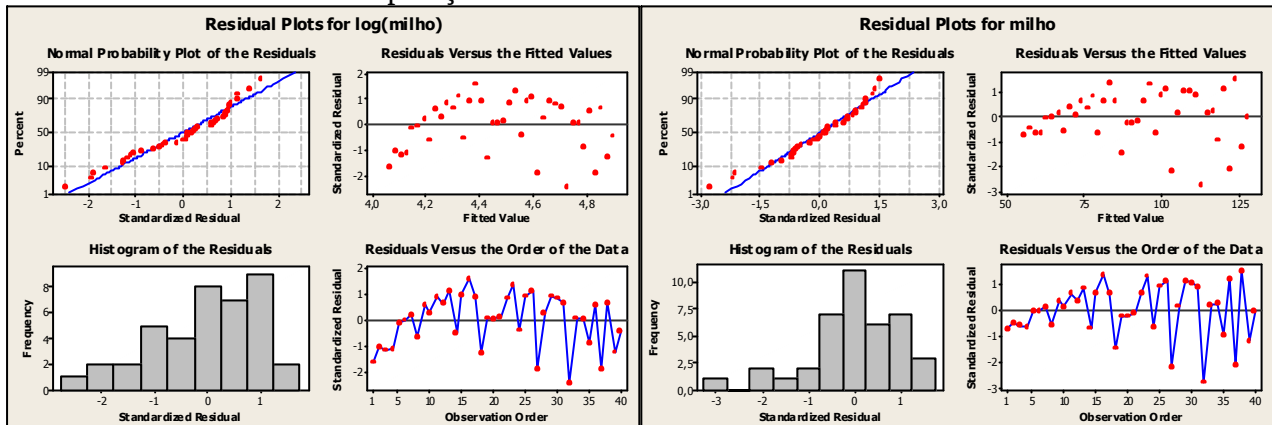
Transformada:

$$\begin{array}{lcl} \log(\text{milho}) = -37,6 + 0,0213 \text{ano} & R^2 = 81,7\% \\ (3,235) & (0,001637) & \hat{s}^2 = 0,0143 \end{array}$$

Original:

$$\begin{array}{lcl} \text{milho} = -3,545 + 1,8396 \text{ano} & R^2 = 81,8\% \\ (278,4) & (0,1408) & \hat{s}^2 = 106 \end{array}$$

Gráficos dos Resíduos – Comparação



A heterogeneidade diminuiu de intensidade com a transformação

A variação explicada continua a mesma, embora a variância do modelo seja grandemente diminuída com a transformação

Regressão Linear Múltipla – Preço Carros Seminovos

Atividade 5

Item 1.a: Modelo 1 – Modelo de regressão simples

Regression Analysis: preço de rev. versus preço novo

The regression equation is
 preço de rev. = 2654 + 0,476 preço novo

Predictor	Coef	SE Coef	T	P
Constant	2654,1	431,2	6,15	0,000
preço novo	0,47568	0,02069	22,99	0,000

S = 1778,48 R-Sq = 79,1% R-Sq(adj) = 78,9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	1671895220	1671895220	528,58	0,000
Residual Error	140	442820588	3163004		
Total	141	2114715808			

Resultados importantes:

Resultados	Modelo 1 – R.simples
R^2	0,791
R^2 ajustado	0,789
$ep(\beta^{\wedge}_0)$	431,2
$ep(\beta^{\wedge}_1)$	0,02069
$\sigma^{\wedge 2}$	3.163.004

Comentários:

- Cerca de 79% da da variação do preço de revenda pode ser explicada por uma relação linear com o preço do automóvel, os demais 21% podem ser considerados como a variação provocada por outros fatores não considerados no modelo de regressão;
- Análise de Variância do modelo (Testes de Significância Conjunta do Modelo)
 H_0 : o coeficiente de inclinação é igual a zero (não significativo)
 H_1 : ele é diferente de zero (coeficiente significativo):
 Nível de significância: 5%
 Estatística F = 528,58
 Graus de liberdade da estatística:
 Numerador: 1
 Denominador: $142 - (1+1) = 140$
 Limite região crítica: $f_c = F_{0,05; (1, 140)} = 3,90874$
 Teste: $F = 528,58 >> f_c = 3,90874 \rightarrow$ rejeita-se H_0 , ou seja, a regressão é significativa
 p-valor ~ 0 .
- Teste de significância dos coeficientes:
 Nível de significância: 5%
 Limite região crítica: $t_c = t_{0,025; 140} = 1,97705$
 Teste: Se $T > t_c \rightarrow$ rejeita-se H_0 , ou seja, o coeficiente é significativo.

<i>Coeficientes</i>	<i>Estatística de Teste</i>	<i>p-valor</i>	<i>Decisão</i>
β_0	$T = 6,15 > t_c = 1,97705$	~ 0	Rejeita-se H_0 : coeficiente significativo
β_1	$T = 22,99 > t_c = 1,97705$	~ 0	Rejeita-se H_0 : coeficiente significativo

- Interpretação:
Cada real de diferença no carro novo, espera-se uma diferença de R\$ 0,476 na revenda;
- Intervalos de 95% de Confiança para os coeficientes:

<i>Coeficientes</i>	<i>Limite inferior</i>	<i>Limite Superior</i>
β_0	1801,56	3506,67
β_1	0,43	0,52

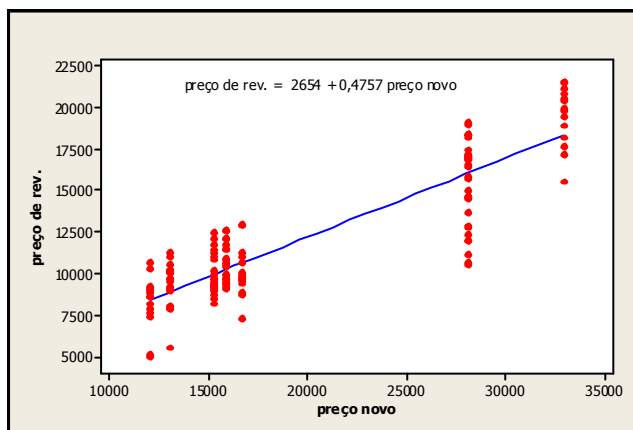


Fig. 1 – Reta de regressão (modelo simples)

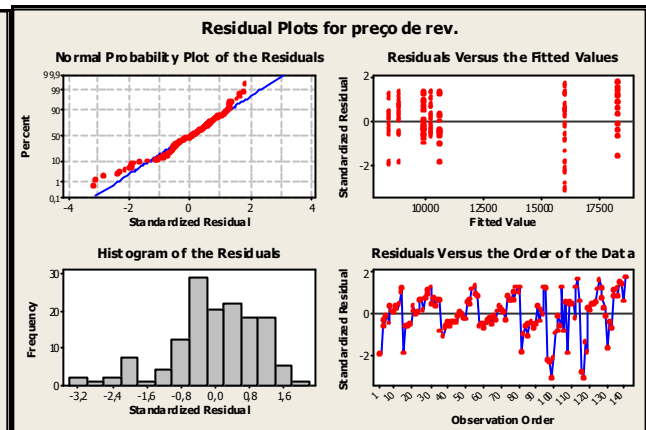


Fig. 2 – Gráfico de resíduos (modelo simples)

- Pelo diagrama de dispersão percebe-se que X só assume alguns determinados valores, isto porque os automóveis em estudo são de 7 modelos e, para cada modelo, o preço 0 km é único.
- O gráfico dos resíduos indica que parece não haver fortes violações nas suposições do modelo de regressão, a não ser a ocorrência maior de valores pequenos com respeito às duas variáveis, o que sugere tentar-se uma transformação logarítmica em X e em Y .

Item 2 – Modelo Simples considerados os logs dos preços:

Regression Analysis: log(revenda) versus log(novo)

The regression equation is
 $\log(\text{revenda}) = 1,80 + 0,768 \log(\text{novo})$

Predictor	Coef	SE Coef	T	P
Constant	1,8025	0,3716	4,85	0,000
log(novo)	0,76775	0,03783	20,30	0,000

$S = 0,155398$ $R\text{-Sq} = 74,6\%$ $R\text{-Sq}(\text{adj}) = 74,5\%$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	9,9479	9,9479	411,94	0,000
Residual Error	140	3,3808	0,0241		
Total	141	13,3287			

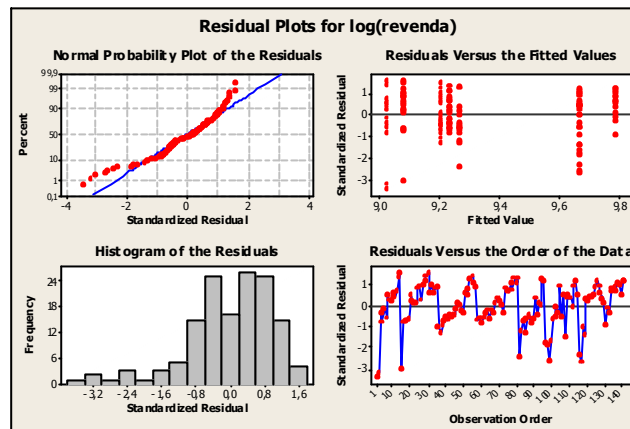


Fig. 3 - Gráfico dos resíduos (modelo simples - log)

- A transformação logarítmica levou uma redução do R^2 .
- O histograma dos resíduos levou a apontar uma distribuição assimétrica, com cauda mais longa à esquerda.
- Em função destes resultados, prefere-se manter o modelo original (regressão simples)
- Na realidade, o preço de um carro seminovo depende de vários outros fatores, levando a um modelo de regressão múltipla.

Item 3 – Modelo de regressão múltipla:

Regression Analysis: preço de rev. versus preço novo; tempo de uso; km

The regression equation is

preço de rev. = 6240 + 0,477 preço novo - 433 tempo de uso - 45,1 km

Predictor	Coef	SE Coef	T	P
Constant	6240,1	352,1	17,72	0,000
preço novo	0,47728	0,01275	37,45	0,000
tempo de uso	-432,9	136,6	-3,17	0,002
km	-45,114	8,997	-5,01	0,000

S = 1087,28 R-Sq = 92,3% R-Sq(adj) = 92,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	1951574151	650524717	550,27	0,000
Residual Error	138	163141657	1182186		
Total	141	2114715808			

Resultados importantes:

Resultados	Modelo 3 – R.Múltipla
R^2	0,923
R^2 ajustado	0,921
$ep(\beta^{\wedge}_0)$	352,1
$ep(\beta^{\wedge}_1)$	0,01275
$ep(\beta^{\wedge}_2)$	136,6
$ep(\beta^{\wedge}_3)$	8,997
$\sigma^{\wedge 2}$	1.182.186

Comentários:

- Cerca de 92% da variação do preço de revenda pode ser explicada por uma relação linear que envolve o preço do automóvel do automóvel 0 km (X_1), o tempo de uso (X_2) e a quilometragem (X_3). O resultado é razoavelmente maior do que o obtido quando foi usada apenas X_1 como variável independente;
- Análise de Variância do modelo (Testes de Significância Conjunta do Modelo)
 H_0 : todos os coeficientes de inclinação são iguais a zero (não significativos)
 H_1 : pelo menos um deles é diferente de zero:
Nível de significância: 5%
Estatística F = 528,58
Graus de liberdade da estatística:
Numerador: 3
Denominador: $142 - (3+1) = 138$
Limite região crítica: $f_c = F_{0,05; (3, 138)} = 2,67020$
Teste: $F = 550,27 >> f_c = 2,67020 \rightarrow$ rejeita-se H_0 , ou seja, as variáveis independentes escolhidas são significativas para explicar a variação de Y ;
p-valor ~ 0 .
- Teste de significância dos coeficientes:
Nível de significância: 5%
Limite região crítica: $t_c = t_{0,025; 138} = 1,97705$
Teste: Se $T > t_c \rightarrow$ rejeita-se H_0 , ou seja, o coeficiente é significativo.

<i>Coeficientes</i>	<i>Estatística de Teste</i>	<i>p-valor</i>	<i>Decisão</i>
β_0	$T = 17,72 > t_c = 1,97730$	~ 0	Rejeita-se H_0 : coeficiente significativo
β_1	$T = 37,45 > t_c = 1,97730$	~ 0	Rejeita-se H_0 : coeficiente significativo
β_2	$T = -3,17 < t_c = 1,97730$	$= 0,002$	Rejeita-se H_0 : coeficiente significativo
β_3	$T = -5,01 < t_c = 1,97730$	~ 0	Rejeita-se H_0 : coeficiente significativo

Rejeitam-se as 4 hipóteses nulas, concluindo que nenhuma das variáveis independentes pode ser excluída do modelo.

- Interpretação:
 - o A cada real de diferença no carro novo, espera-se uma diferença de R\$ 0,48 na revenda (mantendo-se constantes o tempo de uso e a quilometragem);
 - o A cada ano de envelhecimento do automóvel, espera-se R\$ 433,00 a menos na revenda (mantendo-se constantes o valor do carro novo e a quilometragem);
 - o A cada 1.000 km rodados, espera-se R\$ 45,11 a menos na revenda (mantendo-se constantes o valor de novo e o tempo de uso).
- Intervalos de 95% de Confiança para os coeficientes:

<i>Coeficientes</i>	<i>Limite inferior</i>	<i>Limite Superior</i>
β_0	5543,89	6936,36
β_1	0,45	0,50
β_2	-703,10	-162,75
β_3	-62,90	-27,32

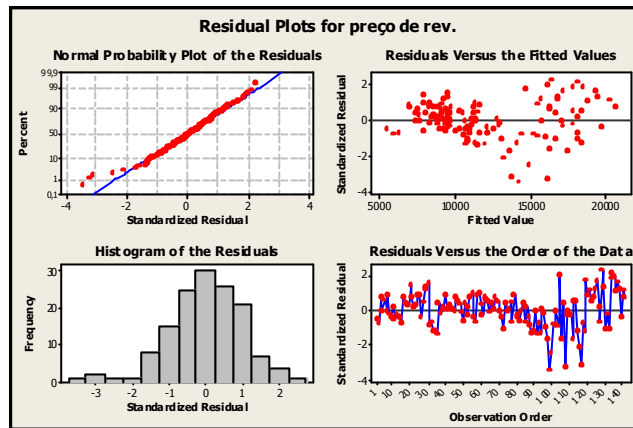


Fig. 4 - Gráfico de resíduos - regressão múltipla

- O gráfico dos resíduos mostra um certo padrão. Para valores preditos pequenos, os resíduos tendem a ser positivos, depois eles tendem a ser negativos e, para valores preditos grandes, eles tendem a ser positivos de novo;
- Além disso, observa-se que dispersão aumenta para os valores preditos maiores;
- Estas características sugerem a aplicação de uma transformação logarítmica na variável dependente.
- Pensando em termos de tempo de uso (X_2) e o valor do automóvel (Y) é mais natural considerar que a cada ano de uso, o automóvel tenha uma redução percentual do seu valor, reforçando a transformação sugerida pelo gráfico de resíduos.

Item 4 – Modelo de regressão múltipla com $\log(Y)$:

Item 4 – Predição de Preço de revenda:

X_1 = R\$ 16.000

X_2 = 2 anos

X_3 = 50.000 km

Resultados	Modelo 1 – R.simples	Modelo 2 – R. simples bilog	Modelo 3 – R.múltipla	Modelo 3 – R.múltipla $\log(Y)$
Preço de revenda predito				
IC (média dos preços de revenda de todos os carros com as características)				
IC (preço de revenda de um carro com as características)				

Item 6 – Escolha do Modelo:

Resultados	Modelo 1 – R.simples	Modelo 2 – R. simples bilog	Modelo 3 – R.múltipla	Modelo 3 – R.múltipla log(Y)
R^2	0,791	0,746	0,923	
R^2 ajustado	0,789	0,745	0,921	
$\beta^{\wedge}_0$	2,654,1	$\log(\beta^{\wedge}_0) = 1,8025$	6,240,1	
$ep(\beta^{\wedge}_0)$	431,2	0,3716	352,1	
$IC(\beta^{\wedge}_0)$	[1.802; 3.507]		[5.544; 6.936]	
$\beta^{\wedge}_1$	0,47568	$\log(\beta^{\wedge}_1) = 0,76775$	0,47728	
$ep(\beta^{\wedge}_1)$	0,02069	0,03783	0,01275	
$IC(\beta^{\wedge}_1)$	[0,45; 0,52]		[0,45; 0,50]	
$\beta^{\wedge}_2$			-432,9	
$ep(\beta^{\wedge}_2)$			136,6	
$IC(\beta^{\wedge}_2)$			[-703,10; -162,75]	
$\beta^{\wedge}_3$			-45,114	
$ep(\beta^{\wedge}_3)$			8,997	
$IC(\beta^{\wedge}_3)$			[-62,90; -27,32]	
$\sigma^{\wedge^2}$	3.163.004	0,0241	1.182.186	

Regressão Linear

Avaliação

Item 1:

Testes de Significância Conjunta do Modelo

H_0 : todos os coeficientes de inclinação são iguais a zero (não significativos)

H_1 : pelo menos um deles é diferente de zero:

Nível de significância: 5%

Estatística $F = 18,02$

Graus de liberdade da estatística:

Numerador: 3

Denominador: $56 - (3+1) = 52$

Limite região crítica: $f_c = F_{0,05; (3, 52)} = 2,78260$

Teste: $F = 18,02 \gg f_c = 2,78260 \rightarrow$ rejeita-se H_0 , ou seja, a regressão é significativa
p-valor ~ 0 .

Testes de Significância para os coeficientes de inclinação:

Nível de significância: 5%

Limite região crítica: $t_c = t_{0,025; 52} = 2,00665$

Teste: Se $T > t_c \rightarrow$ rejeita-se H_0 , ou seja, a regressão é significativa

Significativas (Levam à rejeição de H_0)	Não significativas (Não levam à rejeição de H_0)
Intercepto $10,59 > 2,01$	qualidade $0,57 < 2,01$
qte $ -5,89 > 2,01$	trend $ -1,70 > 2,01$

Sinais:

Quanto maior a quantidade menor o preço;

Quanto maior a qualidade maior o preço

O preço decai com o tempo.

Entretanto não há evidências para se considerar os coeficientes de qualidade e tempo. Neste caso, o sinal não faz sentido.

Item 2:

51,0% da variação d do preço é explicada pela variação na quantidade, qualidade e tempo.

Item 3:

Testes da relação entre preço e quantidade:

$H_0: \beta_1 = 0$ (preço maior para quantidades maiores)

$H_1: \beta_1 < 0$ (preço menor para quantidades maiores):

Nível de significância: 5%

Estatística $T = -5,89$

Graus de liberdade da estatística: 52

Limite região crítica: $t_c = t_{0,05; 52} = -1,67469$

Teste: $T = -5,89 < t_c = -1,67469 \rightarrow$ rejeita-se H_0 , ou seja, há evidências amostrais para concluir que os vendedores aceitam preços mais baixos se puderem vender quantidades maiores.

Item 4:

Teste de significância de β_2 :

$H_0: \beta_2 = 0$ (qualidade não influencia o preço)

$H_1: \beta_2 \neq 0$ (qualidade influencia o preço):

p-valor ~ 0

Estatística $T = -0,57$

Graus de liberdade da estatística: 52

Limite região crítica: $t_c = t_{0,025; 52} = -2,00665$

Teste: $T = -0,57 > t_c = -2,00665 \rightarrow$ não se rejeita H_0 , ou seja, não há evidências amostrais para concluir que a qualidade influencia o preço.

Item 5:

Teste da relação entre preço e tempo:

$H_0: \beta_3 = 0$ (preço não decai com o tempo)

$H_1: \beta_3 < 0$ (preço decai com o tempo):

Nível de significância: 5%

Estatística $T = -1,70$

Graus de liberdade da estatística: 52

Limite região crítica: $t_c = t_{0,05; 52} = -1,67469$

Teste: $T = -1,70 < t_c = -1,67469 \rightarrow$ rejeita-se H_0 , ou seja, há evidências amostrais para concluir que o preço decai com o tempo.

De qualquer maneira, do problema, sugere-se duas situações conflitantes, sendo que deveríamos ter mais informações para decidir-se:

- Por se tratar de um bem exaurível, poderia se esperar que o preço crescesse com o tempo
- Por haver um aumento na quantidade oferecida, o preço apresentaria queda com o tempo (que é a conclusão que chegamos a partir das evidências amostrais)

Item 6:

Os coeficientes da qualidade e trend não são significativo de acordo ao teste de significância individual, assim é razoável estudar o modelo linear simples considerando a quantidade como variável explicativa, já que foi a única variável significante individualmente, no ajuste efetuado no item 1.

Verifica-se que a regressão simples é significante neste caso.

Item 7:

Resultados	Modelo 1 – R.múltipla	Modelo 2 – R.simples
R^2	51,0%	48,1%
R^2 ajustado	48,1%	47,2%
$\beta^{\wedge}_0$	90,847	87,101
$ep(\beta^{\wedge}_0)$	8,580	3,169
$\beta^{\wedge}_1$	-0,05997	-0,057130
$Ep(\beta^{\wedge}_1)$	0,01018	0,008069

$\beta^{\wedge}_2$	0,1162	
$Ep(\beta^{\wedge}_2)$	0,2033	
$\beta^{\wedge}_3$	-2,355	
$Ep(\beta^{\wedge}_3)$	1,386	
$\sigma^{\wedge 2}$	402,3	410

O modelo 2 (regressão simples) têm um erro padrão menor para o coeficiente da quantidade, e levaria a estimativa mais precisa do valor deste coeficiente.

A variância residual do modelo 1(regressão múltipla) é ligeiramente menor, podendo levar a uma precisão ligeiramente melhor para predições de preço.

Item 8:

IC 95% - Modelo 1- Regressão Múltipla: $[-0,05997 \pm 2,00665 \times 0,01018]$

IC 95% - Modelo 2- Regressão Simples: $[-0,057130 \pm 2,00665 \times 0,008069]$

Modelo	Limite Inferior	Limite Superior
(1) R. Múltipla	-0,08040	-0,03954
(2) R. Simples	-0,07332	-0,04094

O modelo 2 fornece uma estimativa intervalar mais precisa do coeficiente.