

Lista 06b com respostas de Estatística II (aula 07)

1) (Barrow, Prob. 7.2 adaptado). Os dados apresentados abaixo fornecem informações sobre o consumo de margarina (em onças por pessoa por semana) e seu preço real no Reino Unido. Usando os dados do problema 7.2 do Barrow, obtemos os resultados abaixo.

Source	SS	df	MS	Number of obs =	19
Model	3.7308	1	3.7308	F(1, 53) =	45.9798
Residual	1.3794	17	0.0811	Prob > F =	0.0000
Total	5.1102	18	0.2839	R-squared =	0.7301
				Adj R-squared =	0.7142
				Root MSE =	0.2849

Cons	Coef.	Std. Error	t	P> t
Preco	-0.0221	0.0033	-6.7808	0.0000
_cons	5.9936	0.3657	18.3888	0.0000

Responda:

a) Escreva o modelo de regressão estimada (o modelo teórico, somente com as letras e o modelo empírico, com os valores dos coeficientes estimados).

$$\widehat{\text{consumo}} = \widehat{\beta}_0 - \widehat{\beta}_1 * \text{preço}$$

b) Qual sinal você espera para o coeficiente da variável preço? Por quê?

$\beta_1 < 0$, pois à medida que o preço aumenta, espera-se redução no consumo.

c) Interprete o coeficiente obtido e teste sua significância. Quantas observações há na amostra?

Há 19 observações na amostra. O coeficiente estimado do preço é -0,0221. Isto significa que se o preço aumentar em \$1, reduz-se o consumo de margarina em 0,0221 onça por pessoa por semana.

Para β_1 , temos que as hipóteses são: $\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$. Adotando um nível de significância de 5%, observa-se o p-valor calculado pelo Excel. Como o p-valor de $\hat{\beta}_1$ (0,000) é menor que 5%, rejeita-se H_0 , e, portanto, preço da margarina é uma variável importante para explicar a quantidade consumida.

d) Preveja qual será o consumo para 2011, se o índice de preço da margarina nesse ano for 122.

Como temos valores observados fora da amostra (pois a amostra vai até 1988), precisa-se prever o consumo para 2011. Da regressão estimada, sabemos que

$$\widehat{\text{consumo}} = \widehat{\beta}_0 + \widehat{\beta}_1 * \text{preço} \rightarrow \widehat{\text{consumo}} = 5,9936 - 0,0221 * (122).$$

Para um preço de 122, o consumo estimado para 2011 é de 3,30.

2. Um analista de seguradora está interessado em analisar o valor em dólares dos danos causados por acidentes automobilísticos. Ela coleta dados de 115 acidentes e registra a quantidade de danos, bem como a idade do motorista. Os resultados de sua regressão estão listados abaixo.

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.187
R Square	0.035
Adjusted R Square	0.026

Standard Error	5652.090
Observations	115.000

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1.000	130433116.219	130433116.219	4.083	0.046
Residual	113.000	3609911959.868	31946123.539		
Total	114.000	3740345076.087			

	<i>Coefficient</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	10725.802	1535.215	6.987	0.000
Age	69.964	34.625	2.021	0.046

A. Qual das seguintes afirmações é verdadeira?

i) A estatística F nos diz que este modelo é estatisticamente significativo para explicar a variação nos danos

ii) Não devemos colocar muita fé nesses resultados, pois o R² é inferior a 10%.

Resposta correta

iii) A idade de um motorista não é estatisticamente significativa para explicar os danos.

iv) Ao nível de significância de 0.05, não há evidência de uma relação linear entre idade e dano.

B. Em média, qual seria o valor em dólares de um acidente envolvendo um homem de 25 anos?

$$\hat{y} = 10725,80 + 69,964 * 25$$

$$\hat{y} = 12474$$

- a.) \$11,836
- b.) \$10,795
- c.) \$13,372
- d.) \$12,474 **correta**

C. Qual das seguintes afirmações é a melhor explicação do R²?

i) 3,5% dos danos causados por acidentes são explicados pela idade do motorista.

ii) 3,5% da variação nos danos do acidente podem ser explicados pela variação na idade do motorista.

Correta, pois o modelo de regressão simples explica quanto da variação em y pode ser explicado por variações em x .

iii) 3,5% das vezes, a quantidade de danos é explicada pela idade do motorista.

iv) 3,5% dos danos do acidente podem ser explicados pela variação da idade do motorista.

D. Como Podemos explicar o intercepto?

i) Para cada aumento adicional de 1 ano na idade do motorista, esperaríamos que o dano aumentasse em \$ 10.726.

ii) Para cada aumento adicional de 1 ano na idade do motorista, esperaríamos que o dano aumentasse em \$ 70.

iii) Não faz sentido, uma vez que não podemos ter um motorista com idade = 0.

iv) O valor médio dos danos foi de \$ 10.726. **correta**