

**Lista 5**

**Parte 1. Pressupostos do Modelo de Mínimos Quadrados Ordinários**

Exercício 1. Responda se as afirmações estão corretas ou erradas. Justifique sua resposta.

a. Mesmo no modelo teórico, os fenômenos sociais estão sujeitos a choques aleatórios que, em última análise, não podem ser modelados.

R: Certo, mesmo o modelo teórico está a sujeito a choques aleatórios não controlado. Pensar o contrário nos levaria ao modelo determinista.

b. Em teoria, se soubermos o quanto que um conjunto de variáveis explicativas estão correlacionadas com determinado fenômeno (como o crescimento do PIB ou o a votação presidencial), podemos prever perfeitamente o resultado desse fenômeno caso soubermos os valores das variáveis explicativas.

R: Errado, mesmo dominando o modelo causal verdadeiro e o valor das variáveis explicativas, ainda haverá fatores estocásticos não determináveis.

c. O "resíduo" representa os choques aleatórios que os fenômenos observados podem sofrer, enquanto o "erro" representa apenas erros de mensuração.

R: Errado, o erro representa os choques aleatório, enquanto o resíduo apresenta tanto o erro quanto outros problemas, como erros de mensuração.

**Parte 2. Teste de Hipóteses do Modelo Clássico de Regressão Linear Normal**

Considere a estatística teste individual de cada parâmetro  $\beta_i$ :

$$\frac{(\hat{\beta}_i)}{ep(\hat{\beta}_i)} \sim t_{n-k-1}$$

Sendo  $ep$  o erro-padrão individual e  $t$  a distribuição  $t$ -Student com  $n - k - 1$  graus de liberdade (com  $n$  observações e  $k$  quantidade de parâmetros desconhecidos, incluindo o intercepto  $\beta_0$ ).

Exercício 2. Com base no banco de dados *fair.dta*

a) [0,5] Obtenha as estimativas para as regressões abaixo:

$$VOTE = \beta_0 + \beta_1 GROWTH + \varepsilon \quad (1)$$

$$VOTE = \beta_0 + \beta_2 INFLATION + \varepsilon \quad (2)$$

No Stata: `regress VOTE GROWTH`

```
. regress VOTE GROWTH
```

Source	SS	df	MS			
Model	406.285326	1	406.285326	Number of obs =	32	
Residual	736.446506	30	24.5482169	F( 1, 30) =	16.55	
Total	1142.73183	31	36.8623172	Prob > F =	0.0003	
				R-squared =	0.3555	
				Adj R-squared =	0.3341	
				Root MSE =	4.9546	

  

VOTE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
GROWTH	.6535869	.1606563	4.07	0.000	.325483	.9816909
_cons	51.85977	.8816524	58.82	0.000	50.05919	53.66034

No Stata: regress VOTE INFLATION

```
. regress VOTE INFLATION
```

Source	SS	df	MS			
Model	27.4162336	1	27.4162336	Number of obs =	32	
Residual	1115.3156	30	37.1771866	F( 1, 30) =	0.74	
Total	1142.73183	31	36.8623172	Prob > F =	0.3973	
				R-squared =	0.0240	
				Adj R-squared =	-0.0085	
				Root MSE =	6.0973	

  

VOTE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
INFLATION	-.4385161	.5106459	-0.86	0.397	-1.481394	.604362
_cons	53.43527	1.732742	30.84	0.000	49.89654	56.97401

Exercício 3. Com base nas estimativas das regressões (1) e (2) do Exercício 2, responda as seguintes perguntas.

a) [0,5] Qual é o efeito estimado de crescimento e inflação obtido no item a (Exercício 2) em relação à variável dependente. Como varia  $y$  em função de um aumento em uma unidade de  $x$ ? Neste momento, preocupe-se apenas com a relação teórica. Deixe a discussão sobre a significância estatística para o item c.

Regressão	$\hat{\beta}$	Coefficiente	Relação com $y$
(1)	$\beta_0$	51.85977	VOTO médio inicial sem efeito de Crescimento
	$\beta_1$	+0.6535869	VOTO médio de acordo com alterações no Crescimento (proporções na tabela abaixo)
(2)	$\beta_0$	53.43527	VOTO médio inicial sem efeito de Inflação
	$\beta_2$	-0.4385161	VOTO médio de acordo com alterações na Inflação (proporções na tabela abaixo)

Regressão	$\hat{\beta}$	Var. Independente	VOTO
(1)	$\beta_1$	↑ 1 Unidade Crescimento	↑ +0.6535869
(2)	$\beta_2$	↑ 1 Unidade Inflação	↓ -0.4385161

De acordo com esse modelo de regressão, a variável independente na relação (1) é diretamente proporcional à variável dependente e a (2) é inversamente proporcional. Isto é, um aumento do crescimento corresponde a um aumento de votos. Por outro lado, um aumento da inflação corresponde a uma redução de votos. As proporções estão na segunda tabela.

b) [0,5] Somente analisando o sinal dos estimadores obtidos, você percebe alguma eventual incoerência com a relação esperada pela teoria do voto econômico? Isto é, existem sinais que são contrários ao que normalmente se esperaria?

Dica: verifique cada sinal obtido e compare com o que seria esperado.

Considerando-se apenas o sinal, a relação fica:

Variável Dependente	Variável Independente	Sinal Esperado	Sinal Obtido	Conclusão
↑ VOTE	↑ GROWTH	(+)	(+)	Coerente
↑ VOTE	↓ INFLATION	(-)	(-)	Coerente

Para este item não foi pedida uma análise de significância de cada estimador.

c) [0,5] Explique os resultados do *item b* em termos de significância estatística.

Resgatando-se a tabela do *item b*:

Variável Dependente	Variável Independente	Sinal Esperado	Sinal Obtido	Conclusão	Signif. Estat.
↑ VOTE	↑ GROWTH	(+)	(+)	Coerente	Signif.
↑ VOTE	↓ INFLATION	(-)	(-)	Coerente	Não-Signif.

Em ambos os casos os sinais são coerentes com o esperado, mas apenas o crescimento possui significância estatística a 5%. A inflação, segundo essa especificação de modelo, não é estatisticamente significativa (não é estatisticamente diferente de nula).

d) [1,0] Considere a seguinte passagem do artigo de Beck (2010):

“Focusing on the one simple test of the null [hypothesis] that  $\beta = 0$  also misleads students into not thinking about the hypothesis of interest. Sometimes we are interested in a series of coefficients, sometimes we are interested in the equality of coefficients, sometimes we care if they are near one, and so on. Current regression output makes it appear that the thing we naturally care about is one specific null. So whatever one thinks about hypothesis testing logic, current regression output is highly misleading”

Seguindo o conselho de Beck, estime os modelos (1) e (2) com hipóteses nulas que sejam mais interessantes do que  $\beta = 0$ . Justifique a hipótese nula escolhida e compare os resultados com os resultados originais dos modelos (1) e (2).

Dica: use o comando **test** no Stata para testar hipóteses nulas diferentes de zero. Por exemplo, para testar a hipótese nula de que o efeito da inflação é 1, use o comando **test \_b[INFLATION] = 1**

O aluno é livre para escolher o beta que será testado, desde que seja capaz de defendê-lo. A discussão dos resultados deve incluir a comparação com os resultados da regressão, na qual  $\beta = 0$ .

A escolha de um beta deve ser justificada de acordo com a teoria. Por exemplo: “um pesquisador alega que o coeficiente para o beta na relação entre voto e crescimento é de 2. Isso é, para cada aumento em uma unidade de crescimento, os votos aumentam em 2 unidades.”.

Existem duas maneiras de testar essa hipótese: através de um teste-t para o coeficiente ou através de um teste-F (teste de Wald para verificar a “veracidade” do parâmetro).

Para o teste-t, não há maneira direta de testar o coeficiente de uma regressão para valores diferentes de zero. Então, deve-se utilizar o teste “imediato” do Stata com o comando `ttesti`.

```
ttesti 32 .6535869 .90880927 2
```

Os argumentos para esse teste são:

```
ttesti <n> <coeficiente obtido na regressão> <desvio-padrão do coeficiente> <valor do teste>
```

Conforme a dica do exercício, desejamos testar o coeficiente de uma maneira menos trabalhosa do que o teste-t imediato, que seria o mais convencional, mas demandaria dos alunos o conhecimento dos argumentos do `ttesti` do Stata. Assim, escolheu-se realizar o teste proposto, que também serve aos propósitos do exercício, mas possui sintaxe menos trabalhosa, conforme demonstrado no próprio enunciado.

Exemplo para  $GROWTH = 2$

```
. test _b[GROWTH] = 2
( 1)  GROWTH = 2
      F( 1, 30) = 70.24
      Prob > F = 0.0000
```

De acordo com a resposta de cada aluno, seja com o teste-t, ou com o teste sugerido pelo enunciado, deve-se demonstrar capacidade de apresentar uma hipótese teórica, apresentar e concluir o teste com base na resposta do Stata.

e) [1,0]. Um colega observa os resultados dos modelos 1 e 2 e fala “Então podemos concluir que o crescimento tem um impacto sobre o voto do partido do incumbente em quanto a inflação não tem nenhum efeito.” É possível chegar a esta conclusão com os resultados obtidos nos dois modelos? Justifique sua resposta.

Não é possível comparar os efeitos de cada variável com base apenas em seus níveis de significâncias ou seus *p-valores*. Isso é um erro comum e a justificativa se encontra na leitura indicada à aula GELMAN & STERN (2006). Os *p-valores* são probabilidades calculadas a partir dos dados específicos de cada variável e não são diretamente comparáveis. GELMAN & STERN (2006) alertam para o erro comum de comparar variáveis com base nos níveis de significância associados aos estimadores.