



## ESTIMANDO MUDANÇAS NA TENDÊNCIA EM SÉRIES TEMPORAIS COM INTERVENÇÕES A PARTIR DE UM MODELO DE REGRESSÃO COM ERROS AUTOCORRELACIONADOS

Débora Regina de Oliveira Moura Abreu<sup>1</sup>, Thais Aidar de Freitas Mathias<sup>2</sup>, Eniuce Menezes de Souza<sup>3</sup>

<sup>1</sup>Doutoranda no Programa de Pós-Graduação em Enfermagem da Universidade Estadual de Maringá e Docente colaborador do curso de Graduação em Enfermagem da Universidade Estadual de Maringá (UEM). E-mail: debora.drom@gmail.com

<sup>2</sup>Professora Titular, Departamento de Enfermagem, Centro de Ciências da Saúde, Universidade Estadual de Maringá, Maringá, PR, Brasil. E-mail: tafmathias@gmail.com

<sup>3</sup>Professora Adjunta Departamento de Estatística, Universidade Estadual de Maringá.(UEM). E-mail: emsouza@uem.br

### RESUMO

Este artigo aborda a estimação e mudanças de tendências em séries temporais quando ocorrem intervenções. São discutidas adequações de modelos de regressão na presença de autocorrelação dos erros a partir dos procedimentos iterativos de estimação de Cochrane-Orcutt e Prais-Winsten. É apresentada a análise de série epidemiológica das taxas de mortalidade por acidentes de transporte no estado do Paraná considerando como intervenção a implantação do novo Código Nacional de Transito e a Lei Seca. **Palavras chave:** Estudos de Séries Temporais, Acidentes de Trânsito; Mortalidade

### 1. INTRODUÇÃO

É muito frequente a necessidade de estimação de tendência em séries temporais epidemiológicas. Embora a classe de modelos SARIMA esteja bastante difundida na área de saúde e da epidemiologia em tais modelos a interpretação da tendência não ocorre de modo natural. Assim, modelos de regressão tem sido muito utilizados na literatura devido a facilidade de interpretação da tendência. Entretanto, por se tratar de séries temporais, a autocorrelação ou correlação serial é naturalmente presente em tais dados, salvo raras

exceções. A autocorrelação quebra um dos principais pressupostos destes modelos de regressão: a independência dos erros. Na presença de autocorrelação, as medidas de qualidade de ajuste são superestimadas, a estimativa de tendência tende a ser significativa, podendo induzir a erros de interpretação <sup>(1,2)</sup>.

Embora modelos mais elaborados possam ser construídos para resolver tal problema, uma possibilidade mais simples é estimar a correlação existente e readequar o modelo de regressão para que o mesmo seja mais realístico e atenda os pressupostos exigidos. Dois procedimentos abordam este princípio: Cochrane-Orcutt e Prais-Winsten <sup>(1,2)</sup>.

## 2. ESTIMAÇÃO POR COCHRANE-ORCUTT E PRAIS-WINSTEN

Para estimar a autocorrelação existente de ordem 1 ( $\rho$ ), por exemplo, pode-se construir um modelo autorregressivo AR (1) para a série residual do modelo de regressão, no qual os erros são autocorrelacionados [1]. Este modelo equivale a fazer um modelo de regressão dos resíduos em função dos mesmos defasados por uma observação:

$$e_t = \rho e_{t-1} + u_t \text{ para } t=2, \dots, n \quad (1)$$

Considerando um modelo usual de regressão simples

$$y_t = \beta_0 + \beta_1 x_t + e_t, \quad (2)$$

Então

$$y_{t-1} = \beta_0 + \beta_1 x_{t-1} + e_{t-1}. \quad (3)$$

Se multiplicarmos a equação 3 por  $\rho$  (estimada na equação 1) e subtraí-la da equação 2, obtém-se

$$y_t - \rho y_{t-1} = \beta_0(1 - \rho) + \beta_1(x_t - \rho x_{t-1}) + u_t \quad (4)$$

Pois  $u_t = e_t - \rho e_{t-1}$ .

Este resultado de uma "quase-diferença" resulta em um modelo sem correlação serial.

Este procedimento é conhecido como algoritmo ou modificação de Cochrane-Orcutt (CO) e é implementado iterativamente até estimação satisfatória de  $\rho$ , determinada quando em duas iterações sucessivas as estimativas de  $\rho$  não diferem mais do que um valor pré-estabelecido pequeno. Na implementação disponível no software R, pacote Orccut, esta convergência é dada na 8ª casa decimal. O primeiro passo do procedimento iterativo refere-se as equações 1 a 4. A partir do segundo passo, os  $\beta$ s estimados na equação 4 e divididos por  $(1-\rho)$ , são utilizados no modelo usual de regressão para obter novamente os resíduos da equação 1 e reestimar  $\rho$ .

Dependendo de como se trata a primeira observação o algoritmo é chamado de CO ou Prais-Winsten. Na estimação por Prais-Winsten [2], faz-se uma transformação em  $t = 1$ :

$$\sqrt{1-\rho^2}y_1 = \beta_0\sqrt{1-\rho^2} + \beta_1\sqrt{1-\rho^2}x_1 + \sqrt{1-\rho^2}e_1 \quad (5)$$

A partir desta transformação o procedimento usual de estimação é realizado. Embora bastante parecidos, a estimação por Prais-Winsten é preferida quando se tem séries curtas, por exemplo uma observação ao ano.

### 3. APLICAÇÃO

Foi construído um modelo de regressão para a taxa de mortalidade por acidentes de transporte no estado do Paraná (em 100000 habitantes), no período de 1980 a 2013, para avaliar a mudança na tendência ocorrida com a implantação no Brasil do Novo Código de Trânsito em 1998 e a Lei Seca em 2008:

$$y_t = \beta_0 + \beta_1 t_{1980-1997} + \beta_2 t_{1998} + \beta_3 t_{1998-2007} + \beta_4 t_{2008-2013} + e_t. \quad (6)$$

Em que  $y_t$  é a taxa de mortes por ano ( $t$ ). Neste modelo, a taxa de mortalidade por acidentes de transporte ajustada no início da análise em 1980 é estimada por  $\beta_0$ , e os demais  $\beta$ s representam as tendências e suas modificações de acordo com cada intervenção, abrupta ( $\beta_2$ ) ou gradual ( $\beta_3$  e  $\beta_4$ ).

Na Tabela 1 pode-se verificar as estimativas dos parâmetros, bem como seus respectivos erros padrão (EP) e valor p.

Tabela 1: Estimativas dos parâmetros do modelo dado pela Equação 6, erros padrão (EP) e valor p.

	$\beta_0$	EP	p	$\beta_1$	EP	p	$\beta_2$	EP	p	$\beta_3$	EP	p	$\beta_4$	EP	p
Total de Acidentes	22,5	1,2	<0,0001	0,7	0,12	<0,0001	-2,22	0,65	0,002	1,69	0,76	0,03	0,25	0,48	0,60
Sexo Masculino	35,6	1,9	<0,0001	1,0	0,19	<0,0001	-3,01	1,03	0,0068	2,33	1,20	0,06	0,24	0,77	0,75
Sexo Feminino	9,1	0,6	<0,0001	0,3	0,05	<0,0001	-1,43	0,30	<0,0001	1,05	0,35	0,006	0,27	0,22	0,2344
0 a 9 anos	5,26	1,0	<0,0001	0,4	0,10	0,0003	-0,03	0,54	0,94	-0,99	0,63	0,12	0,45	0,40	0,271
10 a 19 anos	12,9	1,0	<0,0001	0,6	0,11	<0,0001	-1,89	0,57	0,01	1,45	0,66	0,03	0,37	0,42	0,3914
20 a 29 anos	29,4	1,9	<0,0001	0,8	0,19	0,0002	-3,20	1,06	0,01	3,11	1,24	0,01	-0,22	0,79	0,775
30 a 39 anos	31,7	1,6	<0,0001	0,7	0,16	0,0005	-2,71	0,90	0,01	1,80	1,05	0,09	0,66	0,67	0,33
40 a 49 anos	34,6	1,9	<0,0001	0,7	0,19	0,0015	-3,16	1,03	0,01	2,33	1,20	0,06	0,07	0,77	0,92
50 a 59 anos	40,3	2,3	<0,0001	0,5	0,22	0,0315	-2,95	1,26	0,02	2,20	1,44	0,13	-0,30	0,92	0,74
60 anos ou mais	51,0	2,3	<0,0001	0,1	0,23	0,5443	-2,42	1,27	0,06	2,36	1,48	0,12	-1,12	0,95	0,24

Na Figura 1 consta, para ilustração, as séries da taxa do total de acidentes e para faixa etária de 0 a 9 anos. Estas séries foram escolhidas para ilustrar os dois comportamentos distintos ocorridos, o de aumento na década de 2000 e o de declínio.

Na Figura 2, está a função autocorrelação (FAC) e função de autocorrelação parcial (FACP) dos resíduos do modelo (6) estimados pelo procedimento tradicional e de CO.

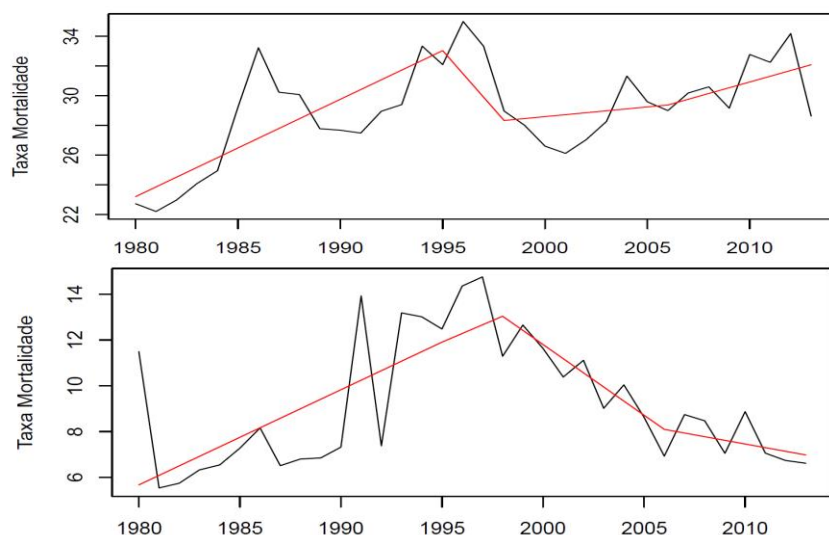
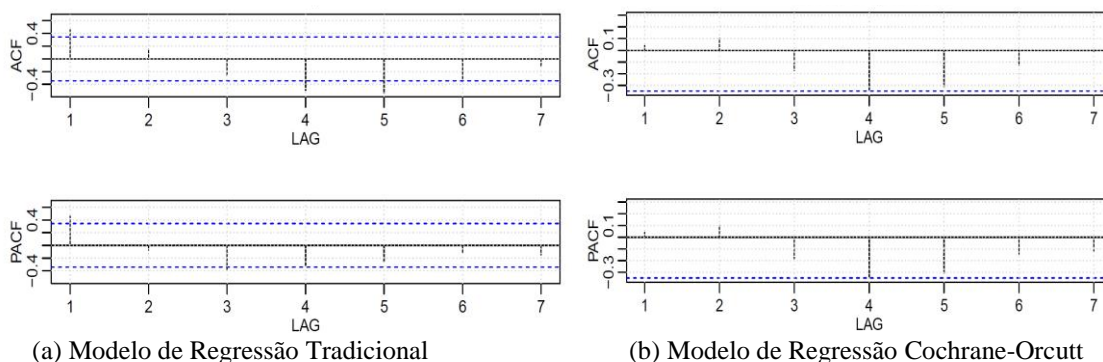


Figura 1 - Modelos ajustados em vermelho para as taxas de mortalidade por acidentes de transporte terrestre (todas as vítimas) (topo) e para a faixa etária de 0 a 9 anos - em preto, Paraná, Brasil, 1980 a 2013.



(a) Modelo de Regressão Tradicional (b) Modelo de Regressão Cochrane-Orcutt  
Figura 2 - Função de Autocorrelação (FAC) e Função de autocorrelação Parcial (FACP) para o modelo de Regressão Tradicional (a) Modelo de Regressão CO (b) para a série do total de mortes

A partir da Figura 2 observa-se que a autocorrelação existente nos resíduos do modelo convencional desaparece após a correção de CO.

## 4. CONSIDERAÇÕES FINAIS

Os resultados permitiram avaliar o impacto de tais leis na mortalidade por acidentes e discutir que, embora tenha-se resultado em uma desaceleração na tendência, a tendência atual ainda é forte, estatisticamente significativa e preocupante. Recomenda-se que sejam considerados os procedimentos de CO ou Prais-Winsten para adequação de modelo de regressão simples para modelagem de tendência em séries temporais. Após o procedimento os resíduos tornam-se independentes (ruído branco) e a inferência estatística é válida.

## REFERÊNCIAS

- [1] COCHRANE, D.; ORCUTT, G. H. "Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms". Journal of the American Statistical Association 44 (245): 32-61, 1949.
- [2] PRAIS S.J., WINSTEN C.B. Trend estimators and serial correlation. Chicago: Cowles Commission, CCDP statistics; n.383, 1954.