

Presidente da República
Luíz Inácio Lula da Silva

Ministro do Planejamento, Orçamento e Gestão
Paulo Bernardo Silva

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE

Presidente
Eduardo Pereira Nunes

Diretor Executivo
Sérgio da Costa Côrtes

ÓRGÃOS ESPECÍFICOS SINGULARES

Diretoria de Pesquisas
Wasmália Socorro Barata Bivar

Diretoria de Geociências
Guido Gelli

Diretoria de Informática
Luiz Fernando Pinto Mariano

Centro de Documentação e Disseminação de Informações
David Wu Tai

Escola Nacional de Ciências Estatísticas
Pedro Luis do Nascimento Silva

Ministério do Planejamento, Orçamento e Gestão
Instituto Brasileiro de Geografia e Estatística - IBGE

REVISTA BRASILEIRA DE ESTATÍSTICA

volume 65 número 224 julho/dezembro 2004

ISSN 0034-7175

R. bras. Estat., Rio de Janeiro, v. 65, n. 224, p. 1-125, jul./dez. 2004

Instituto Brasileiro de Geografia e Estatística - IBGE

Av. Franklin Roosevelt, 166 - Centro - 20021-120 - Rio de Janeiro - RJ - Brasil

© IBGE. 2006

Revista Brasileira de Estatística, ISSN 0034-7175

Órgão oficial do IBGE e da Associação Brasileira de Estatística – ABE.

Publicação semestral que se destina a promover e ampliar o uso de métodos estatísticos (quantitativos) na área das ciências econômicas e sociais, através de divulgação de artigos inéditos.

Temas abordando aspectos do desenvolvimento metodológico serão aceitos, desde que relevantes para os órgãos produtores de estatísticas.

Os originais para publicação deverão ser submetidos em três vias (que não serão devolvidas) para:

Francisco Louzada-Neto
Editor responsável – RBEs – IBGE.
Av. República do Chile, 500 – Centro
20031-170 – Rio de Janeiro, RJ.

Os artigos submetidos às RBEs não devem ter sido publicados ou estar sendo considerados para publicação em outros periódicos.

A Revista não se responsabiliza pelos conceitos emitidos em matéria assinada.

Editor Responsável

Francisco Louzada-Neto (UFSCAR)

Editor de Estatísticas Oficiais

Denise Britz do Nascimento Silva (GAB/IBGE)

Editor de Metodologia

Enrico Antonio Colosimo (UFMG)

Editores Associados

Gilberto Alvarenga Paula (USP)

Dalton Francisco de Andrade (UFSC)
Ismenia Blavatsky de Magalhães (DPE/IBGE)
Helio dos Santos Migon (UFRJ)
Francisco Cribari-Neto (UFPE)

Editoração

Helem Ortega da Silva - Coordenação de Métodos e Qualidade - DPE/COMEQ/IBGE

Impressão

Gráfica Digital/Centro de Documentação e Disseminação de Informações - CDDI/IBGE, em 2004.

Capa

Renato J. Aguiar – Coordenação de Marketing/CDDI/IBGE

Ilustração da Capa

Marcos Balster – Coordenação de Marketing/CDDI/IBGE

Revista brasileira de estatística/IBGE, - v.1, n.1 (jan./mar.1940), - Rio de Janeiro:IBGE, 1940-

v.

Trimestral (1940-1986), semestral (1987-).

Continuação de: Revista de economia e estatística.

Índices acumulados de autor e assunto publicados no v.43 (1940-1979) e v. 50 (1980-1989).

Co-edição com a Associação Brasileira de Estatística a partir do v.58.

ISSN 0034-7175 = Revista brasileira de estatística.

I. Estatística – Periódicos. I. IBGE. II. Associação Brasileira de Estatística.

IBGE. CDDI. Div. de Biblioteca e Acervos Especiais

CDU 31 (05)

RJ-IBGE/88-05 (rev.98)

PERIÓDICO

Impresso no Brasil/Printed in Brazil

Sumário

Nota do Editor.....5

Artigos

Modelos de riscos proporcionais e aditivos para o tratamento de covariáveis dependentes do tempo em análise de sobrevivência.....7

*Tarciana Liberal Pereira
Enrico Antônio Colosimo
Maria Cristina F. Raposo
Pedro Israel C. de Lira
Marília de Carvalho Lima*

Dinâmica inflacionária brasileira: modelagem e previsão.....25

*Nilton José Neves Cordeiro
Francisco Cribari-Neto*

Avaliação de critérios de convergência univariados para o método de Monte Carlo via Cadeias de Markov.....59

*Denismar Alves Nogueira
Thelma Sáfyadi
Daniel Furtado Ferreira*

Uso de modelos de espaços de estados para a estimação do efeito de vício de grupos de rotação na PME/IBGE.....89

*Marcelo Trindade Pitta
Denise Britz do Nascimento Silva*

Política editorial.....123

Nota do Editor

É com prazer que apresento este volume da RBEs que encerra o ano de 2004.

Ainda continuamos a desprender esforços para colocar a revista em dia. Tarefa que, nos parece difícil, mas não impossível. No momento, acredito que em breve teremos este feito realizado.

Além disso, gostaria de comunicar que o conselho editorial está averiguando a possibilidade de incluirmos em volumes futuros da RBEs artigos convidados com discussão, o que daria uma dinâmica diferenciada à revista, proporcionando aos leitores diferentes visões sobre o mesmo tema.

Este volume é composto por quatro artigos. O primeiro artigo, relacionado à análise de sobrevivência, apresenta dois tipos de modelos de risco que podem ser utilizados para acomodar variáveis dependentes do tempo. No segundo artigo, os autores utilizam diferentes estratégias de modelagem para prever a dinâmica inflacionária brasileira e apontam a combinação de previsões como o procedimento mais eficaz. O terceiro artigo é direcionado à avaliação de critérios de convergência univariados para o método de Monte Carlo via Cadeias de Markov, direcionando para aplicação dos vários critérios no monitoramento da convergência de forma combinada, para aproveitar suas melhores características. O quarto artigo apresenta uma metodologia que permite a estimação do efeito de vício de grupos de rotação via modelos de espaço de estados, em pesquisa amostrais para dados provenientes de pesquisas repetidas no tempo, particularmente da pesquisa mensal de emprego do IBGE.

Aproveito a oportunidade para agradecer a equipe do IBGE, particularmente a nossa secretária Helem Ortega da Silva, bem como as editoras Ismenia Blavatsky de Magalhães e Denise Britz do Nascimento Silva.

Uma excelente leitura.

Francisco Louzada-Neto
Editor Responsável

Modelos de riscos proporcionais e aditivos para o tratamento de covariáveis dependentes do tempo em análise de sobrevivência

*Tarciana Liberal Pereira**
Enrico Antônio Colosimo***
Maria Cristina F. Raposo****
Pedro Israel C. de Lira*****
Marília de Carvalho Lima******

* Os autores agradecem à CAPES e ao CNPq pelo auxílio financeiro parcial.

** Endereços para correspondências: Dept^o de Estatística, Univ. Federal da Paraíba, Cidade Universitária - Campus I, João Pessoa/PB, 58059-900.

*** Dept^o de Estatística, Univ. Federal de Minas Gerais, Pampulha, Belo Horizonte/MG, 31270-901.

**** Dept^o de Estatística, Univ. Federal de Pernambuco, Cidade Universitária, Recife/PE, 50740-540.

***** Dept^o de Nutrição, Univ. Federal de Pernambuco, Cidade Universitária, Recife/PE, 50740-540.

***** Dept^o Materno-Infantil, Univ. Federal de Pernambuco, Cidade Universitária, Recife/PE, 50740-540.

Resumo

Em alguns estudos em análise de sobrevivência, os valores das covariáveis mudam no tempo e estes são registrados ao longo do estudo. Estas variáveis são denominadas covariáveis tempo-dependentes e quando consideradas podem fornecer resultados mais precisos e evitar sérios vícios de estimação. O modelo de riscos proporcionais de Cox pode ser estendido para incorporar o efeito destas covariáveis. Aalen (1989) propôs um modelo de risco aditivo que fornece uma alternativa útil ao modelo de Cox, especialmente quando o efeito das covariáveis variam no tempo. Ele permite através de técnicas gráficas a observação de possíveis mudanças no tempo da influência de cada uma das covariáveis. Neste trabalho estes dois modelos são apresentados e é mostrado o uso deles na presença de covariáveis dependentes do tempo. Um conjunto de dados relacionado à identificação de fatores de risco para o aleitamento materno ilustra o ajuste destes dois modelos.

Palavras-chave: modelo de Aalen, modelo de Cox, riscos aditivos, riscos proporcionais.

1. Introdução

Em muitos estudos na área médica ou industrial, a variável resposta é o tempo transcorrido até a realização de algum evento de interesse. A análise de sobrevivência é um conjunto de modelos e técnicas estatísticas adequados a lidar com dados deste tipo.

O termo sobrevivência é usado porque o primeiro uso destas técnicas surgiu de uma empresa de seguro que estava desenvolvendo métodos de custo de prêmios de seguro de vida. Era necessário conhecer o risco, ou tempo de sobrevivência médio, associado a um tipo particular de cliente.

Em análise de sobrevivência é muito comum, tanto na engenharia quanto na medicina, que o interesse esteja nos fatores que influenciam o tempo de sobrevivência. O modelo de riscos proporcionais de Cox (Cox, 1972) é extensivamente usado em dados médicos com o objetivo de verificar o efeito de possíveis covariáveis. O efeito dessas covariáveis de interesse pode variar ao longo do tempo de duração do estudo. O modelo de Cox pode ser generalizado para incorporar o efeito destas covariáveis conhecidas como covariáveis dependentes do tempo.

O modelo proposto por Cox, incluindo sua idéia de verossimilhança parcial, é extremamente útil e representa um dos maiores avanços na análise de dados

censurados. Além disto, ele permite de forma simples a incorporação de covariáveis dependentes do tempo. Algumas alternativas para o modelo de Cox têm sido sugeridas na literatura ao longo dos anos. Em 1980, Aalen propôs um modelo de risco aditivo que tem mostrado freqüentemente vantagens práticas, especialmente quando as covariáveis têm efeitos variando no tempo. Aalen (1980, 1989, 1993) mostrou que seu modelo é útil também como uma ferramenta de diagnóstico para o modelo de Cox.

Considerando a grande utilidade dos modelos de regressão de dados censurados e também a necessidade de incluir covariáveis tempo-dependentes, pretende-se neste artigo comparar os resultados obtidos pelo modelo estendido de riscos proporcionais de Cox com os do modelo aditivo de Aalen, quando aplicados aos dados de um estudo de aleitamento materno.

Na Seção 2, apresentamos o modelo de riscos proporcionais de Cox estendido para incorporar covariáveis dependentes do tempo. Na Seção 3, apresentamos o modelo aditivo de Aalen e suas propriedades. A análise estatística realizada e apresentada neste artigo, na Seção 4, foi motivada por um projeto de pesquisa realizado pelos Departamentos de Nutrição, Materno-Infantil e de Fisiologia e Farmacologia da Universidade Federal de Pernambuco. Trata-se de uma coorte de 652 crianças selecionadas através de amostragem sistemática em seis maternidades nas áreas urbanas dos municípios pernambucanos de Palmares, Catende, Água Preta e Joaquim Nabuco e acompanhadas no período de setembro de 1997 a fevereiro de 2000. O objetivo do estudo, dentre outros, era identificar os fatores de risco associados ao desmame das crianças. O acompanhamento do aleitamento materno foi realizado através de visitas domiciliares realizadas duas vezes por semana nos primeiros 12 meses e uma vez por semana dos 12 aos 18 meses. Nestas visitas além da informação referente ao desfecho (desmame), registrou-se as medidas antropométricas de peso e comprimento. Estas covariáveis são dependentes do tempo. A aplicação envolvendo a identificação de fatores de risco para o tempo até o desmame é analisada usando ambos os modelos. O artigo termina, na Seção 5, com alguns comentários finais.

As análises foram realizadas utilizando o *software* R que se caracteriza pelo compromisso entre a grande flexibilidade e a conveniência de *softwares* estatísticos

tradicionais, incluindo uma ampla variedade de métodos estatísticos tradicionais e modernos. Para mais detalhes sobre esta linguagem de programação ver Cribari-Neto e Zarkos (1999). A versão utilizada foi a 1.7.0 e está disponível no endereço <http://www.r-project.org>. As funções necessárias para elaboração desse trabalho estão à disposição com o primeiro autor.

2. Modelo estendido de riscos proporcionais de Cox

O modelo de Cox estendido para incorporar covariáveis dependentes do tempo pode ser escrito, em termos da função de risco no tempo t , dado x , na seguinte forma:

$$h_i(t) = h_0(t) \exp(x_i'(t)\beta). \quad (1)$$

em que $h_0(t)$ é uma função arbitrária chamada de função de risco padrão ou de base, $x(t) = (x_1(t), x_2(t), \dots, x_p(t))'$ é o vetor de covariáveis dependentes do tempo, β é o vetor de parâmetros regressores associado com as covariáveis. É importante verificar que definindo desta forma, o modelo dado pela Equação (1) não é mais de riscos proporcionais. Os valores das covariáveis $x_i'(t)$ dependem do tempo t e a razão das funções de risco no tempo t para dois indivíduos i e j é também dependente do tempo. A interpretação dos coeficientes do modelo deve então considerar o tempo t .

O modelo de regressão de Cox é caracterizado pelos coeficientes β que medem o efeito das covariáveis sobre a função de risco. Para fazer inferência nos coeficientes β , Cox propôs em 1975 (Cox, 1975) o método de máxima verossimilhança parcial que condiciona a verossimilhança à história dos tempos de sobrevivência e censuras anteriores. Considere uma amostra de n indivíduos, onde se tem $k(\leq n)$ falhas distintas nos tempos $t_1 \leq t_2 \leq \dots \leq t_k$. A função de verossimilhança parcial $L(\beta)$ é dada por:

$$L(\beta) = \prod_{i=1}^n \left[\frac{\exp(x_i'(t_i)\beta)}{\sum_{j \in R(t_i)} \exp(x_j'(t_i)\beta)} \right]^{\delta_i}; \quad (2)$$

em que

$$\delta_i = \begin{cases} 0, & \text{se o } i\text{-ésimo tempo de vida é censurado,} \\ 1, & \text{caso contrário,} \end{cases}$$

e $R(t_i)$ é o conjunto dos índices dos indivíduos sob risco no tempo t_i .

Estimadores para o vetor de parâmetros β podem ser obtidos maximizando o logaritmo da função de verossimilhança parcial (2), ou seja, resolvendo o sistema de equações definido pela função escore $U(\beta) = 0$. Isto é o equivalente a

$$U(\beta) = \sum_{i=1}^n \delta_i \left[x_i'(t_i)\beta - \log \sum_{j \in R(t_i)} \exp(x_j'(t_i)\beta) \right] = 0. \quad (3)$$

O procedimento de estimação requer um método iterativo que é geralmente o método de Newton-Raphson, pois as equações encontradas em (3) não apresentam forma fechada. Para construir intervalos de confiança e testar hipóteses sobre os coeficientes do modelo são necessárias propriedades assintóticas dos estimadores de máxima verossimilhança parcial. As provas mais gerais das propriedades para covariáveis dependentes do tempo foram apresentadas por Andersen e Gill (1982).

3. Modelo aditivo de Aalen

Um modelo alternativo ao de Cox, baseado no modelo de risco multiplicativo para processo de contagem, foi sugerido originalmente por Aalen (1980). Este modelo apresentado de forma mais simples em Aalen (1989) é um modelo de risco aditivo para análise de regressão de dados censurados. Este modelo aditivo de Aalen fornece uma alternativa útil ao modelo de riscos proporcionais de Cox, pois permite que tanto os parâmetros quanto os vetores de covariáveis variem com o tempo. Já que os efeitos temporais não são assumidos de serem proporcionais para cada covariável, o modelo de Aalen é capaz de fornecer informações detalhadas a respeito da influência temporal de cada covariável. Os modelos de Cox e Aalen diferem fundamentalmente. O de Cox tem uma função básica não-paramétrica, mas o efeito das covariáveis é modelado parametricamente. Por outro lado, o modelo de Aalen é completamente não-paramétrico no sentido de que funções são ajustadas e não os parâmetros. Ou seja, na estimação

dos parâmetros o modelo de Aalen usa apenas informação local o que faz este modelo bastante flexível.

A forma mais geral do modelo para $h_i(t)$ que é acessível a análises estatísticas é

$$h_i(t) = \alpha(t, x_i(t)), \quad (4)$$

em que α é uma função do tempo geral e desconhecida. Apesar de esse modelo ser atrativo do ponto de vista teórico, a exigência de tamanhos de amostras grandes torna-o difícil de ser utilizado na prática.

Assumindo que $\alpha(t,0) = 0$ e ignorando todos os termos de ordens maiores da expansão de Taylor de primeira ordem de $\alpha(t, X)$ sobre $X = 0$, ou seja, o primeiro termo da expansão da série de Taylor de uma função de risco geral sobre o vetor de covariáveis igual a zero, então o modelo (4) se reduz ao modelo de risco aditivo de Aalen dado por

$$h_i(t) = \alpha_0(t) + \sum_{j=1}^p \alpha_j(t) x_{ij}(t). \quad (5)$$

Considerando a forma matricial

$$h(t) = \alpha(t)Y(t),$$

em que $\alpha(t) = (\alpha_0(t), \alpha_1(t), \dots, \alpha_p(t))'$ é um vetor de funções do tempo desconhecidas, cujo primeiro elemento $\alpha_0(t)$ é interpretado como uma função de parâmetro básica, enquanto os $\alpha_j(t)$'s, $j = 1, \dots, p$, chamados de funções de regressão medem a influência das respectivas covariáveis. A matriz $Y(t)$ de ordem $n \times (p+1)$ é construída da seguinte maneira: se o evento considerado ainda não ocorreu para o i -ésimo indivíduo e ele não é censurado, então, a i -ésima linha de $Y(t)$ é o vetor

$$x_i(t) = (1, x_{i1}(t), x_{i2}(t), \dots, x_{ip}(t))'.$$

Caso contrário, se o indivíduo não está sob risco no tempo t , então, a linha correspondente de $Y(t)$ contém apenas zeros.

Este modelo é considerado não-paramétrico, pois nenhuma forma paramétrica particular é assumida para as funções de regressão. Como visto, estas funções podem variar arbitrariamente com o tempo, revelando mudanças na influência das covariáveis.

Esta é uma das vantagens do modelo (5), bem como a não exigência de tamanho de amostra extremamente grande.

O modelo de riscos proporcionais assume que os efeitos das covariáveis agem multiplicativamente na função de risco. Os coeficientes estimados da estrutura de regressão são constantes desconhecidas, cujos valores não mudam com o tempo. No modelo de Aalen assume-se que as covariáveis agem de maneira aditiva na função de risco e os coeficientes de riscos desconhecidos podem ser funções do tempo, ou seja, os efeitos das covariáveis podem variar durante o estudo. Dessa forma os estimadores dos parâmetros são baseados nas técnicas de mínimos quadrados. A derivação desses estimadores é similar a derivação do estimador de Nelson-Aalen da função de risco acumulada (Klein e Moeschberger, 1997).

A aproximação para estimação depende das suposições sobre a forma funcional das funções de regressão que, neste caso, são não-paramétricas. A estimação direta das funções de regressão é difícil na prática, sendo mais fácil a estimação da função de regressão acumulada. Isto ocorre pelo mesmo motivo que é mais fácil estimar a função de distribuição acumulada do que a função de densidade de probabilidade. Considera-se, então, a estimação do vetor coluna $A(t)$ com elementos $A_j(t)$ dados por

$$A_j(t) = \int_0^t \alpha_j(s) ds ,$$

$j = 1, \dots, p$. Sejam $t_1 < t_2 < \dots < t_k$ os tempos de falhas ordenados. Aalen considerou um estimador razoável de $A(t)$, denominado estimador de mínimos quadrados de Aalen, que é dado por

$$A^*(t) = \sum_{t_k \leq t} Z(t_k) I_k , \quad (6)$$

em que I_k é um vetor de zeros que assume o valor 1 para o indivíduo cujo evento ocorre no tempo t_k . Enquanto que $Z(t)$ é a inversa generalizada de $Y(t)$. Em princípio, $Z(t)$ pode ser qualquer inversa generalizada de $Y(t)$. Uma escolha simples pode ser baseada no princípio de mínimos quadrados local, ou seja,

$$Z(t) = [Y(t)' Y(t)]^{-1} Y(t)' .$$

Esta inversa usada comumente em modelos de regressão, em geral, pode não ser ótima. Uma escolha ótima dependerá do conhecimento dos verdadeiros valores dos parâmetros. Huffer e McKeague (1987) sugeriram o uso de uma outra inversa definindo assim o estimador de mínimos quadrados ponderados. Neste trabalho usamos a inversa de mínimos quadrados.

Os componentes de $A^*(t)$ convergem assintoticamente, sob condições apropriadas, para um processo gaussiano (Aalen 1989). Então, um estimador da matriz de covariância de $A^*(t)$ é dado por

$$\Omega^*(t) = \sum_{t_k \leq t} Z(t_k) I_k^D Z(t_k)',$$

em que I_k^D é uma matriz diagonal com I_k como diagonal.

Não é difícil verificar, como consequência dos resultados obtidos anteriormente, que se pode estimar o risco acumulado e a função de sobrevivência correspondentes, dados os valores das covariáveis. Seja $x(t) = (1, x_1(t), x_2(t), \dots, x_p(t))'$ o conjunto de valores das covariáveis no tempo t . O estimador do risco acumulado $H^*(t)$ é dado por

$$H^*(t) = A^*(t)' x(t).$$

A partir da função de risco acumulada, a função de sobrevivência pode, então, ser estimada por

$$S^*(t) = \exp(-H^*(t)). \quad (7)$$

A função de sobrevivência estimada não é necessariamente monótona sobre todo o período de observação. Ela pode aumentar para alguns valores de t e de acordo com a equação (7) decrescer para algum t .

É freqüentemente de interesse testar se uma covariável específica tem algum efeito na função de risco total. Para o modelo aditivo de Aalen isto corresponde a testar a hipótese nula de que não existe efeito da covariável no risco. A hipótese nula para algum $j \geq 1$ é estabelecida como

$$H_j : \alpha_j(t) = 0, \quad t \in [0, T],$$

em que T é o maior tempo observado.

É importante lembrar que no contexto não-paramétrico a hipótese nula acima pode apenas ser testada sobre intervalos de tempo onde $Y(t)$ tem posto completo. Dentro da estrutura do modelo, Aalen (1980, 1989) desenvolveu para todo tempo de falha uma estatística de teste para H_j dada pelo j -ésimo elemento U_j do vetor

$$U = \sum_{t_k} K(t_k) Z(t_k) I_k, \quad (8)$$

em que $K(t)$, uma função peso não negativa, é uma matriz diagonal $(p+1) \times (p+1)$. A estatística de teste da Equação (8) surge como uma combinação ponderada da soma do estimador de $A_j(t)$ apresentado na equação (6). Os elementos diagonais de $K(t)$ são funções pesos e suas escolhas podem depender das alternativas para a hipótese nula de interesse.

Uma escolha ótima da função peso necessitará do conhecimento das verdadeiras variâncias dos estimadores, entretanto isto dependerá de funções de parâmetros desconhecidas. Aalen (1989) considerou duas escolhas para a função peso. A primeira possibilidade é considerar cada função peso igual ao número de observações que permanecem no conjunto de risco em algum tempo dado. Uma segunda escolha, utilizada neste trabalho, é tomar $K(t) = \{diag[(Y(t)'Y(t))^{-1}]\}^{-1}$, em que $K(t)$ é dada como a inversa de uma matriz diagonal, tendo a mesma diagonal principal da matriz $(Y(t)'Y(t))^{-1}$. Este peso é escolhido por analogia ao problema da regressão de mínimos quadrados, em que as variâncias dos estimadores são proporcionais aos elementos diagonais da matriz $(Y'Y)^{-1}$, sendo Y o desenho da matriz. Um estimador da matriz de covariância de U dado pela Equação (8) é

$$V = \sum_{t_k} K(t_k) Z(t_k) I_k^D Z(t_k)' K(t_k)'$$

Suponha que se queira testar simultaneamente todos H_j para j em algum subconjunto A de $\{1, \dots, p\}$ consistindo de s elementos. Seja U_A definido como o subvetor correspondente de U e V_A a submatriz correspondente de V , isto é, V_A é a

matriz de covariâncias estimadas de U_A . A estatística de teste normalizada $U_A' V_A^{-1} U_A$ é assintoticamente distribuída como uma qui-quadrado com s graus de liberdade quando H_j vale para todo j em A . Se o interesse é testar apenas uma das hipóteses H_j , então, é usada a estatística de teste $U_j V_{jj}^{-1/2}$. Esta estatística tem uma distribuição assintótica normal padrão sob a hipótese nula.

Em estudos clínicos ou tratamentos médicos, a significância de uma covariável pode mudar durante o período de acompanhamento. Através do modelo de Aalen é possível estimar a contribuição das covariáveis para a função de risco em cada tempo de falha. O resumo desta contribuição sobre o tempo produz uma função de regressão para cada covariável que pode ser plotada contra o tempo. Ou seja, $A_j^*(t)$ pode ser considerada como uma função empírica descrevendo a influência da j -ésima covariável.

A inclinação do gráfico da função de regressão acumulada contra o tempo fornece informação sobre a influência de cada covariável, sendo possível verificar se uma covariável particular tem um efeito constante ou varia com o tempo ao longo do período de estudo. Por exemplo, se $\alpha_j(t)$ é constante, então, o gráfico deve aproximar-se de uma linha reta, paralela ao eixo do tempo. Inclinações positivas ocorrem durante períodos em que aumentos dos valores das covariáveis são associados com aumentos na função de risco. Por outro lado, inclinações negativas ocorrem em períodos quando crescimentos nos valores das covariáveis estão associados com decréscimos na função de risco. As funções de regressão acumuladas têm inclinações aproximadamente iguais a zero em períodos em que as covariáveis não influenciam a função risco. Portanto, o gráfico das funções de regressão do modelo linear de Aalen pode ser recomendado também como um instrumento para detectar efeitos de covariáveis dependentes do tempo, bem como uma técnica de diagnóstico que pode extrair informações adicionais úteis. Através de um estudo, Mau (Mau, 1986) mostrou que as funções de regressão podem fornecer informações importantes que devem ser perdidas quando apenas o modelo de Cox é utilizado.

4. Ilustração numérica: dados de aleitamento materno

A partir de pesquisas de âmbito nacional, feitas sobre a questão alimentar e nutricional do País, evidenciou-se no Brasil que havia uma tendência decrescente na prevalência da amamentação, conseqüência de mudanças de hábitos e comportamentos introduzidos pela vida moderna, tais como a participação da mulher no mercado de trabalho e a utilização do leite artificial. A zona da mata meridional de Pernambuco é uma área socialmente mais vulnerável em vários aspectos, entre eles na ocorrência da desnutrição, com características muito comuns no Nordeste brasileiro. A principal atividade econômica é a monocultura açucareira, imprimindo características próprias à organização política e social, no meio rural e urbano. A fragilidade dos sistemas produtivos alternativos (fruticultura, pequenas lavouras), a sazonalidade do desemprego, o subemprego em atividades subsidiárias de comércio urbano e da prestação de serviços avulsos contribuem ainda mais para que esta área seja mais exposta aos problemas de saúde e nutrição que resultam de condições desfavoráveis de vida. Uma das conseqüências decorrente dos fatos citados acima, e que ocorre nesta região, é a curta duração do aleitamento materno. Os dados utilizados neste artigo fazem parte de um projeto de pesquisa realizado pelos Departamentos de Nutrição, Materno-Infantil e de Fisiologia e Farmacologia da Universidade Federal de Pernambuco em colaboração com a Universidade de Londres (LSHTM) e a Universidade de Montpellier-França. Trata-se de uma coorte de 652 crianças selecionadas através de amostragem sistemática em seis maternidades dos municípios pernambucanos de Palmares, Catende, Água Preta e Joaquim Nabuco. O acompanhamento dessas crianças, selecionadas no período de setembro de 1997 a agosto de 1998, foi de setembro de 1997 a fevereiro de 2000. Neste trabalho, crianças que nunca mamaram foram excluídas, sendo utilizadas então as informações de 642 crianças. Foram avaliadas as condições socioeconômicas, ambientais e demográficas das famílias, assistência no pré-natal, características reprodutivas maternas, bem como o estado nutricional *post-partum* através de indicadores antropométricos (peso e altura). O acompanhamento foi realizado através de visitas domiciliares aos 2, 4, 6, 9, 12, 15 e 18 meses de vida, onde além das informações a respeito do aleitamento e das medidas de peso e comprimento registrou-

se o perímetro cefálico e o torácico. Os Quadros 1 e 2 apresentam as covariáveis utilizadas neste estudo.

Quadro 1 - Covariáveis categóricas utilizadas

Variáveis	Descrição	Codificação	Frequência (%)
SEXO	Sexo da criança	1 (masculino) 2 (feminino)	327 (50,9) 315 (49,1)
TRABGRAV	Mãe trabalhou durante a gravidez	1 (sim) 2 (não)	156 (24,3) 486 (75,7)
BEBEU	Mãe ingeriu bebida alcoólica na maioria dos dias da gravidez	1 (sim) 2 (não)	22 (3,4) 620 (96,6)
TIPOCASA	Regime de ocupação da residência	1 (própria/alugada) 2 (outros)	567 (88,3) 75 (11,7)
ÁGUA	Abastecimento de água	1 (rede geral) 2 (outros)	589 (91,7) 53 (8,3)
LIXO	Destino do lixo	1 (coleta direta) 2 (coleta indireta)	406 (63,2) 236 (36,8)
GELAD	Existe geladeira funcionando	1 (sim) 2 (não)	344 (53,6) 298 (46,4)

Fonte dos dados básicos: arquivos da Pesquisa

Quadro 2 - Covariáveis contínuas e discretas utilizadas

Variáveis	Descrição	Média (D.P.) ⁽²⁾
WAZ ⁽¹⁾	Estado nutricional: escore Z do peso observado sobre o peso normal, por idade	-0,0016 (1,0048)
HAZ ⁽¹⁾	Estado nutricional: escore Z da altura observada sobre a altura normal, por idade	-0,4997 (0,8941)
QUANCONS	Quantidade de consultas de pré-natal durante a gravidez	3 (2,293)
CIGARROS	Quantidade média de cigarros fumados por dia na gravidez	1,24 (4,697)
IDMAE	Idade da mãe em anos	22,75 (6,229)
ESTUDMAE	Anos de escolaridade da mãe	5,43 (3,560)
ESTUDPAI	Anos de escolaridade do pai	5,78 (3,713)
RPERCAPT	Renda <i>per capita</i> do domicílio (R\$)	80,9735 (94,8909)
IMCMAE	Índice de massa corporal da mãe	24,1594 (3,6728)

Fonte dos dados básicos: arquivos da Pesquisa.

(1) Usado o padrão do National Center Health Statistics - NCHS.

(2) Desvio Padrão.

Para identificar quais covariáveis dentre as pesquisadas que influenciam no tempo de aleitamento, foi utilizado inicialmente o modelo de regressão de Cox com covariáveis dependentes do tempo. O Quadro 3 apresenta os resultados deste ajuste. Pode-se observar no Quadro 3 uma grande quantidade de covariáveis não significativas. Um procedimento passa-a-passo foi utilizado para obter o modelo final. Neste procedimento, as covariáveis foram retiradas uma a uma a partir do maior valor-p no modelo. O modelo final foi obtido quando as covariáveis restantes tinham um valor-p inferior a 0,10. Este modelo está apresentado no Quadro 4. Observe que neste procedimento a covariável "tipocasa" ficou significativa e a "cigarros" perdeu significância quando comparado com seus valores no Quadro 3. O Quadro 4 mostra que há evidências estatísticas de que as covariáveis "waz" e "tipocasa" são fatores influentes no tempo de aleitamento. O risco relativo mostrado nesta tabela foi obtido após exponenciarmos os valores das estimativas dos parâmetros. Através dos dados apresentados no referido quadro, pode-se verificar que, se a criança mora em casa própria ou alugada (tipocasa = 1), o risco de ela parar de mamar é maior do que se ela mora em casa invadida ou cedida, o que indica que crianças com melhor condição financeira mamam menos. Quanto maior o valor da covariável "waz" melhores as condições nutricionais das crianças. Dessa forma pode-se verificar que quanto maior o valor de "waz" menor o risco da criança parar de mamar.

Quadro 3 - Resultados do primeiro ajuste do modelo de Cox com covariáveis dependentes do tempo

Covariável	Coefficiente	Erro padrão	P-valor	Risco Reativo (I.C. - 95%)
waz	-0,204	0,090	0,024	0,815(0,683;0,973)
haz	0,128	0,096	0,180	1,137(0,942;1,372)
sexo	0,0261	0,1084	0,810	1,026(0,830;1,269)
quancons	-0,0345	0,0258	0,180	0,966(0,918;1,016)
trabgrav	-0,0387	0,1311	0,770	0,962(0,744;1,244)
cigarros	0,0265	0,0125	0,033	1,027(1,002;1,052)
bebeu	-0,2953	0,3016	0,330	0,744(0,412;1,344)
idmae	-0,0004	0,0101	0,970	1,000(0,980;1,020)
estudmae	0,0309	0,0199	0,120	1,031(0,992;1,072)
estudpai	0,0011	0,0194	0,950	1,001(0,964;1,040)
rpercapt	-0,0005	0,0006	0,430	1,000(0,998;1,001)
imcmae	0,0152	0,0155	0,330	1,015(0,985;1,047)
tipocasa	-0,2726	0,1892	0,150	0,761(0,525;1,103)
agua	-0,0704	0,2242	0,750	0,932(0,601;1,446)
lixo	-0,1034	0,1293	0,420	0,902(0,700;1,162)
gelad	0,1717	0,1270	0,180	1,187(0,926;1,523)

Quadro 4 - Resultados do ajuste final do modelo de Cox com covariáveis dependentes do tempo

Covariável	Coefficiente	Erro padrão	P-valor	Risco Reativo (I.C. - 95%)
waz	-0,0976	0,0485	0,044	0,907(0,825;0,998)
tipocasa	-0,2819	0,1571	0,073	0,754(0,554;1,026)

Como uma alternativa ao modelo de Cox, foi ajustado o modelo aditivo de Aalen. Os resultados do ajuste, com todas as covariáveis, estão mostrados no Quadro 5, cujos resultados revelam que, tal como no modelo de Cox, a maioria das covariáveis não são significativas.

Quadro 5 - Resultados do primeiro ajuste do modelo aditivo de Aalen

Covariável	Coefficiente	Erro padrão	P-valor	(I.C. - 95%)
constante	1,673	1,451	0,116	(-1,171;4,517)
waz	-0,265	0,132	0,042	(-0,524;-0,005)
haz	0,148	0,147	0,215	(-0,140;0,437)
sexo	-0,007	0,176	0,722	(-0,352;0,338)
quancons	-0,042	0,038	0,135	(-0,115;0,032)
trabgrav	-0,206	0,255	0,740	(-0,705;0,293)
cigarros	0,075	0,054	0,169	(-0,030;0,180)
bebeu	-0,128	0,564	0,435	(-1,234;0,977)
idmae	-0,006	0,015	0,792	(-0,036;0,025)
estudmae	0,039	0,033	0,156	(-0,026;0,105)
estudpai	0,040	0,037	0,678	(-0,032;0,112)
rpercapt	-0,001	0,001	0,432	(-0,003;0,001)
imcmae	0,019	0,025	0,443	(-0,030;0,068)
tipocasa	-0,415	0,216	0,186	(-0,839;0,008)
agua	0,176	0,329	0,833	(-0,470;0,822)
lixo	-0,141	0,221	0,549	(-0,575;0,293)
gelad	0,255	0,216	0,210	(-0,168;0,677)

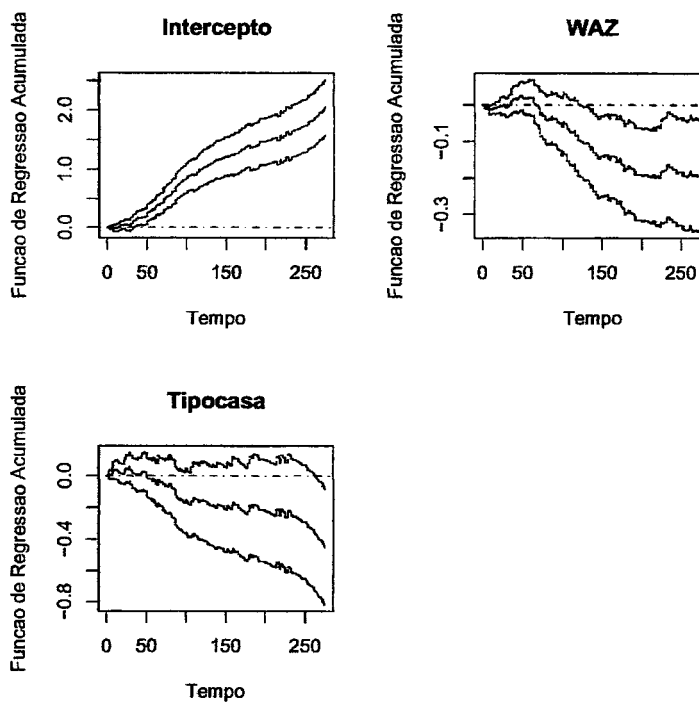
Após a retirada das covariáveis não significativas foi então definido um modelo, cujos resultados estão apresentados no Quadro 6, de onde constata-se que apenas as covariáveis "waz" e "tipocasa" apresentaram-se significativas. Pode-se observar que analogamente ao modelo de Cox, crianças que moram em casa própria ou alugada mamam menos do que crianças que moram em casa invadida ou cedida. Constata-se também que crianças desnutridas, menores valores de "waz", mamam menos do que as crianças que se encontram em condições nutricionais normais.

Quadro 6 - Resultados do ajuste final do modelo aditivo de Aalen

Covariável	Coefficiente	Erro padrão	P-valor	Risco Reativo (I.C. - 95%)
constante	2,056	0,240	0,000	(1,586;2,526)
waz	-0,188	0,079	0,049	(-0,343;-0,034)
tipocasa	-0,453	0,187	0,066	(-0,820;-0,085)

Uma vantagem do modelo aditivo de Aalen, conforme já referido, é que ele é capaz de fornecer informações detalhadas a respeito da influência temporal de cada covariável. A Figura 1 mostra as funções de regressão acumuladas e os respectivos intervalos, com 95%, de confiança para o intercepto e para as covariáveis significativas no modelo de Aalen. A função de regressão acumulada para a covariável "waz" é não linear e apresenta evidências de que o efeito desta covariável diminui com o tempo e parece desaparecer após aproximadamente seis meses de acompanhamento, quando a função tem um comportamento paralelo ao eixo do tempo. O mesmo acontece com a covariável "tipocasa" contudo, o seu efeito parece estabilizar em menos de quatro meses.

Figura 1 - Estimativas das funções de regressão acumuladas com intervalo de 95% de Confiança



5. Conclusões

Em estudos de sobrevivência, o modelo de riscos proporcionais de Cox é um método padrão de análise. O modelo aditivo de Aalen é uma alternativa ao de Cox, bastante útil principalmente na presença de covariáveis dependentes do tempo, pois permite que tanto as covariáveis como os parâmetros variem no tempo, fornecendo assim informações detalhadas a respeito da influência temporal de cada covariável. Neste trabalho foram apresentados estes dois modelos e uma aplicação envolvendo dados censurados na presença de covariáveis, cujos valores variam no tempo.

Foi utilizado o modelo de Cox generalizado para incluir o efeito de covariáveis dependentes do tempo e o modelo aditivo de Aalen. Os resultados obtidos a partir da análise estatística mostram que os resultados para os dois modelos são bastante similares. Inclusive, os modelos finais de Cox e de Aalen incluem as mesmas covariáveis. Uma diferença importante entre os dois modelos é que como no modelo de Aalen os parâmetros também variam no tempo, é possível construir um gráfico das funções de regressão acumuladas, onde é possível verificar o comportamento da influência de cada covariável no tempo. Foi verificado que tanto a situação nutricional quanto a variável tipo de casa diminuem sua influência no tempo de aleitamento materno com o passar do tempo. Este gráfico das funções de regressão acumuladas representa uma ferramenta importante na análise dos dados, pois covariáveis medidas no início do período de observação podem, freqüentemente, perder sua influência no tempo, sendo útil ter um método que revela isto.

Referências bibliográficas

- AALEN, O. O. (1978). Nonparametric Inference for a Family of Counting Processes. *Annals of Statistics*, 6, 701-726.
- AALEN O. O. (1980). A Model for Nonparametric Regression Analysis of Counting Processes. *Lecture Notes in Statistics* Springer, New York, 2, 1-25.
- AALEN, O. O. (1989). A Linear Regression Model for the Analysis of Life Times. *Statistics in Medicine*, 8, 907-925.
- AALEN, O. O. (1993). Further Results on the Non-Parametric Linear Regression Model in Survival Analysis. *Statist. Med.*, 12, 1569-1588.
- ANDERSEN, P. K. & GILL, R. (1982). Cox's Regression Model for a Counting Processes: A Large Sample Study. *Ann. Statistics*, 10, 1100-1200.

- COX, D.R.(1972). Regression Models and Life Tables(with discussion). *Journal Royal Statistics Soc. B*, 34, 187-220.
- COX, D.R.(1975). Partial Likelihood. *Biometrika*, 62, 269-276.
- CRIBARI-NETO, F. & ZARKOS, S. G. (1999). R Yet Another Econometric Programming Environment. *Journal Appl. Econ.*, 14, 319-329.
- HUFFER, F. W. & MCKEAGUE, I.W. (1987). *Survival analysis using additive risk models*. Technical Report 396, Department of Statistics, Stanford University.
- KLEIN, J. P. & MOESCHBERGER, M.L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag, New York.
- MAU, J. (1986). On a graphical method for the detection of time-dependent effects of covariates in survival data. *Applied Statistics*, 35, 245-255.

Abstract

In some real situations, covariates may be monitored and measured along all the follow-up period. These covariates are known as time-dependent covariates. Analyses that include these covariates can be more reliable and avoid possible estimation bias. Cox proportional hazards model is very flexible and can be extended to include time-dependent-covariates. Aalen (1989) proposed an additive risk model that is a very attractive alternative to the Cox model especially when the covariates effects varies in time. It uses some simple graphical techniques that allow to detect covariate time change effects. In this work these two models are presented in the presence of time-dependent covariates. A real data set related to mother breastfeeding is used to illustrate the fit of these two models.

Keywords: Aalen's model, additive risk model, Cox's model, proportional hazards model.

Dinâmica inflacionária brasileira: modelagem e previsão*

*Nilton José Neves Cordeiro**
Francisco Cribari-Neto****

Resumo

Com o principal objetivo de buscar boas previsões da dinâmica inflacionária brasileira, utilizamos neste artigo diferentes estratégias de modelagem/previsão. Dentre as abordagens consideradas, sugerimos uma nova proposta envolvendo o uso de decomposição não-paramétrica para estimar o núcleo da inflação onde adotamos esta componente em uma estrutura de regressão dentro de um modelo dinâmico na previsão de valores futuros. Procedemos ainda com alisamento exponencial, metodologia Box-Jenkins e combinação de previsões provenientes de mais de uma estratégia individual de previsão. Os resultados indicaram que a combinação de previsões foi o método mais eficaz dentre todos os outros.

1. Introdução

A evolução da inflação brasileira ao longo do tempo tem sido objeto de vários estudos, muitos dos quais dedicados à investigação de sua natureza e mecanismos de controle. Um aspecto relevante do ponto de vista de planejamento empresarial e de implementação de políticas públicas é a habilidade de prever satisfatoriamente o

* Os autores agradecem à Capes e ao CNPq pelo auxílio financeiro.

** Endereço para correspondências: Dept^o. de Matemática, Univ. Estadual Vale do Acaraú, Campus da Cidao - Sobral/CE, Cep. 62040-370, Brasil.

*** Depto. de Estatística, Univ. Federal de Pernambuco, Cidade Univ. - Recife/PE, Cep.50740-540, Brasil.

comportamento futuro da inflação. Neste sentido, empresas privadas (incluindo bancos) e formuladores de políticas econômicas buscam continuamente estratégias de previsão da dinâmica inflacionária de forma a planejarem ações futuras, quer de investimentos, quer de formulação e implementação de políticas econômicas. Quanto melhores forem as previsões obtidas, melhores serão, *ceteris paribus*, as decisões tomadas.

O presente artigo considera a construção de previsões do comportamento futuro da taxa de inflação. Diferentes estratégias são consideradas e avaliadas comparativamente. O enfoque mais simples envolve o uso de algoritmos de alisamento exponencial, enfoque este que não se revelou promissor. A metodologia Box-Jenkins também é considerada, tendo apresentado bons resultados. A análise desenvolvida aqui, todavia, não se restringe a estas duas estratégias de previsão, comumente utilizadas em aplicações práticas. Nós propomos uma estratégia de previsão que combina uma estrutura de modelagem dinâmica com uma estrutura de regressão que envolve uma estimativa do núcleo da inflação, estimativa esta que revela períodos de aceleração e desaceleração da dinâmica inflacionária. Para tanto, nós propomos uma nova forma de estimação do núcleo inflacionário utilizando uma decomposição não-paramétrica. Esta estimativa é de implementação simples e, contrariamente a outras estimativas propostas na literatura, não requer a modelagem de séries adicionais. A quarta e última estratégia de previsão que consideramos baseia-se na idéia de combinação de previsões provenientes de diferentes metodologias, explorando assim o que cada uma tem de melhor a oferecer. De fato, os resultados sugerem que esta é a estratégia mais eficaz.

O presente artigo encontra-se organizado da seguinte forma. A Seção 2 apresenta as diferentes estratégias de previsão consideradas. Os dados são apresentados e brevemente descritos na Seção 3. Os resultados de modelagem e previsão se encontram na Seção 4. A Seção 5, por sua vez, inclui uma análise comparativa dos resultados. Por fim, a Seção 6 conclui o artigo com comentários finais.

2. Metodologias de previsão empregadas

Um modelo bastante utilizado para a representação de uma série temporal X_t é o modelo SARIMA(p, d, q)(P, D, Q) $_s$, dado por

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps})(1 - B)^d (1 - B^s)^D X_t = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}) \varepsilon_t,$$

onde ϕ , θ , Φ e Θ são vetores de parâmetros auto-regressivos, de médias móveis, sazonais auto-regressivos e sazonais de médias móveis, respectivamente.

Adicionalmente, d e D são as ordens de integração e de integração sazonal, respectivamente, sendo s o período de sazonalidade, e B é o operador de defasagem. Já ε_t é um ruído branco, ou seja, uma seqüência de variáveis aleatórias não-correlacionadas, com média zero e variância constante. Os modelos ARMA e ARIMA, que também são bastante difundidos na literatura, podem ser obtidos como casos particulares do modelo acima.

É possível combinar modelos ARMA (e suas extensões) com uma estrutura de regressão, resultando em

$$X_t = \alpha + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \beta_0 Z_t + \beta_1 Z_{t-1} + \dots + \beta_r Z_{t-r} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (1)$$

onde Z_t é uma variável independente, ε_t é ruído branco e $\alpha, \beta_0, \dots, \beta_r, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ são parâmetros desconhecidos. O modelo (1) é comumente denominado ARMAX, o qual Newbold & Bos (1994) abordam de maneira sucinta. Este modelo pode ser estendido para incluir mais de uma variável explicativa.

A inclusão de uma estrutura de regressão, como em (1), possibilita verificar diferentes impactos ocorridos por mudanças nas variáveis independentes sobre a variável dependente. Quando os betas (β) são nulos, o modelo se reduz a uma representação ARMA(p, q) para a variável dependente X_t .

2.1. Método Box-Jenkins

Box & Jenkins (1976) popularizaram um método iterativo de construção de modelos para previsão conhecido como *método Box-Jenkins*, que consiste, basicamente, em encontrar um bom modelo estocástico linear na classe $ARIMA(p, d, q)$ para séries sem sazonalidade e na classe $SARIMA(p, d, q)(P, D, Q)$ para dados sazonais. A metodologia de modelagem é composta de cinco estágios, a saber: (i) seleção de uma classe de modelos adequada; (ii) identificação do modelo; (iii) estimação dos parâmetros do modelo; (iv) verificação da adequação do modelo; e (v) utilização do modelo para geração de previsões. As etapas (ii), (iii) e (iv) são realizadas de forma iterativa na busca de um modelo adequado.

Uma maneira de identificar o modelo é através da minimização de um critério de escolha de modelos, como o AIC (*Akaike Information Criterion*) ou o BIC (*Bayesian Information Criterion*). As propriedades destes critérios em amostras finitas podem ser encontradas em Mills & Prasad (1992).

Uma vez que o modelo foi selecionado, ou seja, tendo o valor d sido especificado e as ordens dos polinômios $\phi(B)$ e $\theta(B)$ determinadas (p e q , respectivamente), é necessário estimar os parâmetros desconhecidos, ou seja, estimar $\phi = (\phi_1, \phi_2, \dots, \phi_p)'$, $\theta = (\theta_1, \theta_2, \dots, \theta_q)'$ e a variância do processo ruído branco, σ_ε^2 . O método de estimação utilizado no presente trabalho foi o de máxima verossimilhança exata (EMLE); os resultados numéricos em Ansley & Newbold (1980) sugerem que este método possui melhor desempenho em amostras finitas que métodos alternativos.

A terceira etapa iterativa do método Box-Jenkins é a verificação de diagnóstico. Aqui busca-se verificar se o modelo estimado é estatisticamente adequado. Caso um modelo proposto seja adequado, espera-se que a dinâmica apresentada pelos dados haja sido capturada e ainda que os resíduos se comportem semelhantemente a um processo ruído branco. Usualmente, verifica-se se os valores das autocorrelações amostrais dos resíduos estão no intervalo (assintótico) $\left[\frac{-2}{\sqrt{T}}, \frac{2}{\sqrt{T}} \right]$ e usam-se as estatísticas Box-Pierce e Ljung-Box para testar a hipótese nula $H_0 : \rho_1 = \rho_2 = \rho_3 = \dots = \rho_h = 0$, para um dado h .

É razoável pensar em alguns critérios de eficiência do modelo que levem em conta o grau de precisão de suas previsões futuras. Assim, é comum reservar os dados mais recentes para avaliação da qualidade das previsões produzidas pelo modelo, baseando-se a estimação dos parâmetros nas observações restantes. O erro de previsão é definido como

$$e_j = \hat{X}_n(j) - X_{n+j},$$

onde $\hat{X}_n(j)$ denota a previsão de X_{n+j} produzida no instante n .

Há várias medidas que podem ser usadas para avaliação da qualidade das previsões, como, por exemplo, o erro quadrático médio (MSE), dado por $MSE = \frac{1}{M} \sum_{j=1}^M e_j^2$, e o erro total (TE), dado por $TE = \sum_{j=1}^M e_j$, onde M denota o horizonte de previsão.

2.2. Alisamento exponencial

Quando o interesse principal recai na previsão de valores futuros de uma série, sem a preocupação da formalização de um modelo, pode-se utilizar uma abordagem bastante difundida chamada de *alisamento exponencial*.

O algoritmo aditivo de alisamento exponencial de Holt-Winters é dado pelas fórmulas de recorrência

$$\begin{aligned} L_t &= \alpha(X_t - F_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1}), \quad 0 < \alpha < 1, \\ T_t &= \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}, \quad 0 < \beta < 1, \\ F_t &= \gamma(X_t - L_t) + (1-\gamma)F_{t-s}, \quad 0 < \gamma < 1, \end{aligned}$$

onde α, β, γ são constantes de alisamento, L_t, T_t, F_t são respectivamente o nível, a inclinação e o fator de sazonalidade estimados para o instante t e s é o período sazonal. As previsões são obtidas através de

$$\begin{aligned} \hat{X}_t(h) &= L_t + hT_t + F_{t+h-s}, \quad h = 1, 2, \dots, s \\ &= L_t + hT_t + F_{t+h-2s}, \quad h = s+1, s+2, \dots, 2s. \\ &\vdots \end{aligned}$$

Para inicializar o algoritmo são sugeridos como valores iniciais (Brockwell & Davis, 1996)

$$\begin{aligned}\hat{X}_s(I) &= X_{s+1}, \\ T_{s+1} &= (X_{s+1} - X_1)/s, \\ F_h &= X_h - (X_1 + T_{s+1}(h-1)), \quad h = 1, \dots, s.\end{aligned}$$

Quando o fator de sazonalidade F no algoritmo Holt-Winters é igual a zero, obtemos o algoritmo de tendência linear de Holt. Aqui, previsões futuras são obtidas a partir das mesmas fórmulas de recorrência apresentadas, porém desconsiderando o fator sazonal. Para o algoritmo ser inicializado utilizam-se como valores iniciais $L_2 = X_2$ e $T_2 = X_2 - X_1$ (Newbold & Bos, 1994). O nível estimado correspondente ao instante t é dado por $L_{t-1} + T_{t-1}$. Previsões futuras são obtidas através de

$$\hat{X}_t(h) = L_t + hT_t, \quad h = 1, 2, 3, \dots$$

Quando a série não apresenta tendência nem sazonalidade, naturalmente há de se desconsiderar os fatores T e F do algoritmo Holt-Winters. Assim, tem-se uma fórmula de recorrência altamente simplificada. Neste caso, obtém-se o algoritmo de alisamento exponencial simples, onde todas as previsões futuras serão iguais ao valor do último nível estimado, ou seja,

$$\hat{X}_t(h) = L_t, \quad h = 1, 2, 3, \dots,$$

onde $\hat{X}_t(h)$ representa o valor previsto de X_{t+h} no instante t , chamado previsão h passos à frente no instante t . Para inicializar este algoritmo, uma escolha óbvia é tomar o valor do primeiro nível igual ao valor da primeira observação.

Para escolha das constantes de alisamento, o critério de minimização da soma dos erros quadrados de previsão um passo à frente é empregado, ou seja, minimizamos $S = \sum e_t^2$, onde e_t é igual a diferença entre o valor predito e o observado.

2.3. Combinação de Previsões

Granger & Newbold (1986), Newbold & Bos (1994), Newbold & Granger (1974) e Reid (1969), entre outros, argumentam que previsões podem ser melhoradas através da combinação de previsões obtidas de mais de uma técnica. A possível melhoria da qualidade da previsão pode ser explicada devido ao fato de cada previsão individual poder conter informações relevantes que talvez estejam ausentes nas outras previsões. Assim, a combinação de previsões busca colher, de cada previsão, sua parcela mais representativa para compor uma nova previsão potencialmente superior. A forma mais usual de geração de uma previsão combinada, $\hat{X}_{c,t}$, é através do cálculo de uma média ponderada das previsões individuais,

$$\hat{X}_{c,t} = w_1 \hat{X}_{1,t} + w_2 \hat{X}_{2,t} + w_k \hat{X}_{k,t}, \quad (2)$$

onde $\hat{X}_{i,t}$ e w_i com $i = 1, 2, \dots, k$ são, respectivamente, as k previsões alternativas e os pesos empregados, tal que $0 \leq w_i \leq 1$, \forall_i e $\sum_{i=1}^k w_i = 1$. Note que estamos simplificando a notação usada para denotar previsões. Um caso especial de (2) é a combinação de apenas duas previsões como

$$\hat{X}_{c,t} = w_1 \hat{X}_{1,t} + w_2 \hat{X}_{2,t} \quad \text{ou} \quad \hat{X}_{c,t} = w_1 \hat{X}_{1,t} + (1 - w_1) \hat{X}_{2,t}, \quad (3)$$

fazendo com que w_1 e w_2 assumam valores entre 0 e 1 e $w_1 + w_2$ seja igual a 1

2.3.1. Combinação de previsões através da média aritmética simples (pesos iguais)

Dadas as previsões individuais $\hat{X}_{1,t}, \hat{X}_{2,t}, \dots, \hat{X}_{k,t}$ para o valor desconhecido X_t , obtidas através de k métodos diferentes de previsão, a maneira mais simples de se fazer combinação de previsões é através da média aritmética simples, ou seja,

$$\hat{X}_{c,t} = \frac{\sum_{i=1}^k \hat{X}_{i,t}}{k}, \quad i = 1, 2, \dots, k, \quad (4)$$

onde $\hat{X}_{c,t}$ é uma previsão combinada para X_t .

Barnard (1963) constatou, empiricamente, melhoria em previsões utilizando (4) para combinar previsões oriundas de modelagem Box-Jenkins com previsões provenientes de alisamento exponencial.

2.3.2. Combinação de previsões através de pesos inversamente proporcionais à soma de erros quadrados

Este procedimento procura atribuir pesos maiores a métodos de previsão mais eficazes e pesos menores aos de desempenho mais pobre. A forma de se atribuir pesos maiores a métodos de maior acurácia e pesos menores a métodos que rendem previsões menos precisas é baseada na soma dos erros quadrados de previsão de forma que o método que apresentar baixo valor para esta soma receberá um peso mais elevado. Assim, para obter $\hat{X}_{c,t}$ como sendo uma combinação de previsões como em (2), é necessário calcular, para cada método, a soma de erros quadrados das n previsões um passo à frente mais recentes:

$$\sum_{j=1}^n (\hat{X}_{i,t-j} - X_{t-j})^2 = \sum_{j=1}^n e_{i,t-j}^2, \quad i = 1, 2, \dots, k;$$

os pesos w_i são dados por

$$w_i = \frac{\left(\sum_{j=1}^n e_{i,t-j}^2\right)^{-1}}{\left(\sum_{j=1}^n e_{i,t-j}^2\right)^{-1} + \dots + \left(\sum_{j=1}^n e_{k,t-j}^2\right)^{-1}}, \quad i = 1, 2, \dots, k.$$

De maneira simplificada, uma previsão combinada de duas previsões alternativas é da forma de (3), onde $\hat{X}_{1,t}$ e $\hat{X}_{2,t}$ são previsões um passo à frente obtidas de duas abordagens distintas, e w_1 e w_2 são os pesos calculados como

$$w_1 = \frac{\sum_{j=1}^n e_{2,t-j}^2}{\sum_{j=1}^n e_{1,t-j}^2 + \sum_{j=1}^n e_{2,t-j}^2} \quad (5)$$

e $w_2 = 1 - w_1$.

Logo, para a utilização desse método é necessário que estejam disponíveis as n previsões mais recentes, as quais podem ser previsões um passo à frente, para que possam ser avaliados os desempenhos passados de cada abordagem utilizada na combinação. Newbold & Bos (1994) sugerem que n esteja entre 6 e 12.

2.3.3. Combinação de previsões feitas através de pesos baseados em regressão

Suponha que desejamos combinar duas previsões com pesos w_1 e w_2 como em (3). Podemos considerar a observação X_t como sendo uma média ponderada de duas previsões mais um erro, da forma

$$X_t = w_1 \hat{X}_{1,t} + (1 - w_1) \hat{X}_{2,t} + e_{c,t}$$

ou, de maneira equivalente,

$$(X_t - \hat{X}_{2,t}) = w_1 (\hat{X}_{1,t} - \hat{X}_{2,t}) + e_{c,t}. \quad (6)$$

Nota-se que (6) é uma representação de um modelo de regressão linear onde $(X_t - \hat{X}_{2,t})$ é a variável dependente, $(\hat{X}_{1,t} - \hat{X}_{2,t})$ é a variável independente, w_1 é o coeficiente angular da reta e $e_{c,t}$ é um erro. Sejam $\hat{X}_{i,t-j}$, $i = 1, 2$ e $j = 1, 2, \dots, n$, dois conjuntos das n mais recentes previsões um passo à frente provenientes de dois procedimentos alternativos de previsão. Assim,

$$(X_{t-j} - \hat{X}_{2,t-j}) = w_1 (\hat{X}_{1,t-j} - \hat{X}_{2,t-j}) + e_{c,t-j}, \quad j = 1, 2, \dots, n,$$

são usados para encontrar uma estimativa para o parâmetro de regressão w_1 por mínimos quadrados:

$$\hat{w}_1 = \frac{\sum_{j=1}^n e_{2,t-j}^2 - \sum_{j=1}^n e_{1,t-j} e_{2,t-j}}{\sum_{j=1}^n e_{1,t-j}^2 + \sum_{j=1}^n e_{2,t-j}^2 - 2 \sum_{j=1}^n e_{1,t-j} e_{2,t-j}}. \quad (7)$$

Conseqüentemente, uma estimativa de w_2 é $\hat{w}_2 = 1 - \hat{w}_1$. Vale notar que a única diferença entre os métodos (5) e (7) é a presença de termos de correlação entre os erros em (7), revelada pelos produtos cruzados, que estão ausentes em (5).

2.3.4. Combinação de previsões feitas através de pesos alternativos

Bates & Granger (1969) propuseram estimadores que são variações de (6) e obtiveram sucesso aplicando-os ao conjunto de dados utilizado por Barnard (1963) considerando pares de previsões um passo à frente. Alguns desses estimadores são:

$$\hat{w}_1 = \frac{\sum_{j=1}^n e_{2,t-j}^2}{\sum_{j=1}^n (e_{1,t-j}^2 + e_{2,t-j}^2)}, \quad (8)$$

$$\hat{w}_{1,t} = \alpha \hat{w}_{1,t-1} + \frac{(1-\alpha) \sum_{j=1}^n e_{2,t-j}^2}{\sum_{j=1}^n (e_{1,t-j}^2 + e_{2,t-j}^2)}, \quad 0 < \alpha < 1, \quad (9)$$

$$\hat{w}_1 = \frac{\sum_{j=1}^{t-1} K^j e_{2,j}^2}{\sum_{j=1}^{t-1} K^j (e_{1,j}^2 + e_{2,j}^2)}, \quad K \geq 1, \quad (10)$$

$$\hat{w}_1 = \frac{\sum_{j=1}^{t-1} K^j (e_{2,j}^2 - e_{1,j} e_{2,j})}{\sum_{j=1}^{t-1} K^j (e_{1,j}^2 + e_{2,j}^2 - 2e_{1,j} e_{2,j})}, \quad K \geq 1. \quad (11)$$

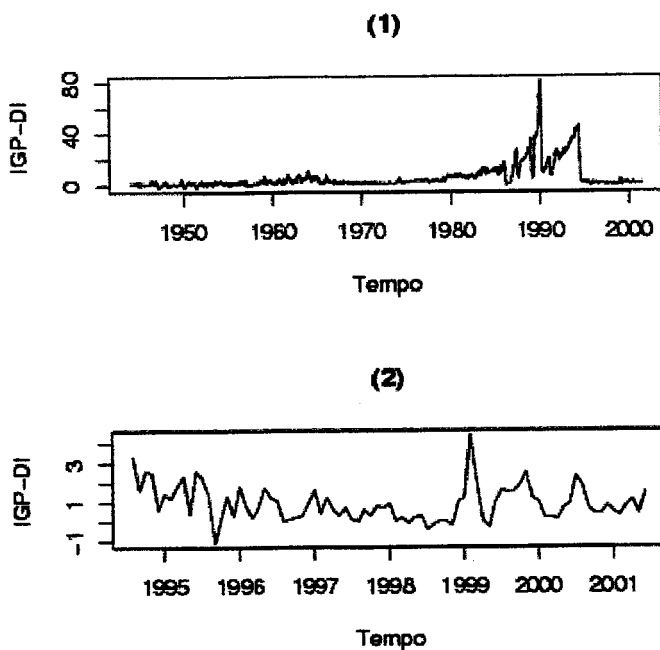
Newbold & Granger (1974) combinaram previsões um passo à frente obtidas por Box-Jenkins e Holt-Winters para 80 séries temporais utilizando (7) e (8) com $n = 1, 3, 6, 9, 12$, (9) com $\alpha = 0,5; 0,7; 0,9$ e $n = 1, 3, 6, 9, 12$, e (10) e (11) com $K = 1,0; 1,5; 2,0; 2,5$. Os pesos w_1 e w_2 ficaram restritos de forma a apenas assumirem valores entre 0 e 1. Os procedimentos (8), (9) e (10) mostraram-se superiores aos demais, indicando assim que os procedimentos que desprezam a correlação entre os erros de previsão (que está presente, por exemplo, em (7) e (11), facilmente constatada pelos produtos cruzados) podem vir a ser mais eficientes, sendo a eficiência aqui definida em relação ao critério de erro quadrático médio.

Foi ainda adotado pelos autores um critério de ordenamento, onde (8) com $n = 12$ apresentou o melhor desempenho, de forma que em 56,25% das 80 séries trabalhadas, o MSE obtido pela combinação de previsões de Box-Jenkins e Holt-Winters, mostrou-se inferior ao MSE fornecido através das previsões individuais por Box-Jenkins.

3. Os dados

Utilizamos o Índice Geral de Preços-Disponibilidade Interna (IGP-DI), que é um índice tradicional, calculado pela Fundação Getúlio Vargas. Este índice engloba apenas produtos disponíveis para o mercado interno, sendo utilizado nas contas nacionais, em todas as suas desagregações, como salienta o *Relatório de Gestão 2001*, da Secretaria de Política Econômica do Ministério da Fazenda do Brasil. O índice é uma ponderação de outros três índices: preços no atacado (peso 0,6), preços ao consumidor (peso 0,3) e preços da construção civil (peso 0,1).

Figura 1 - Séries IGP-DI

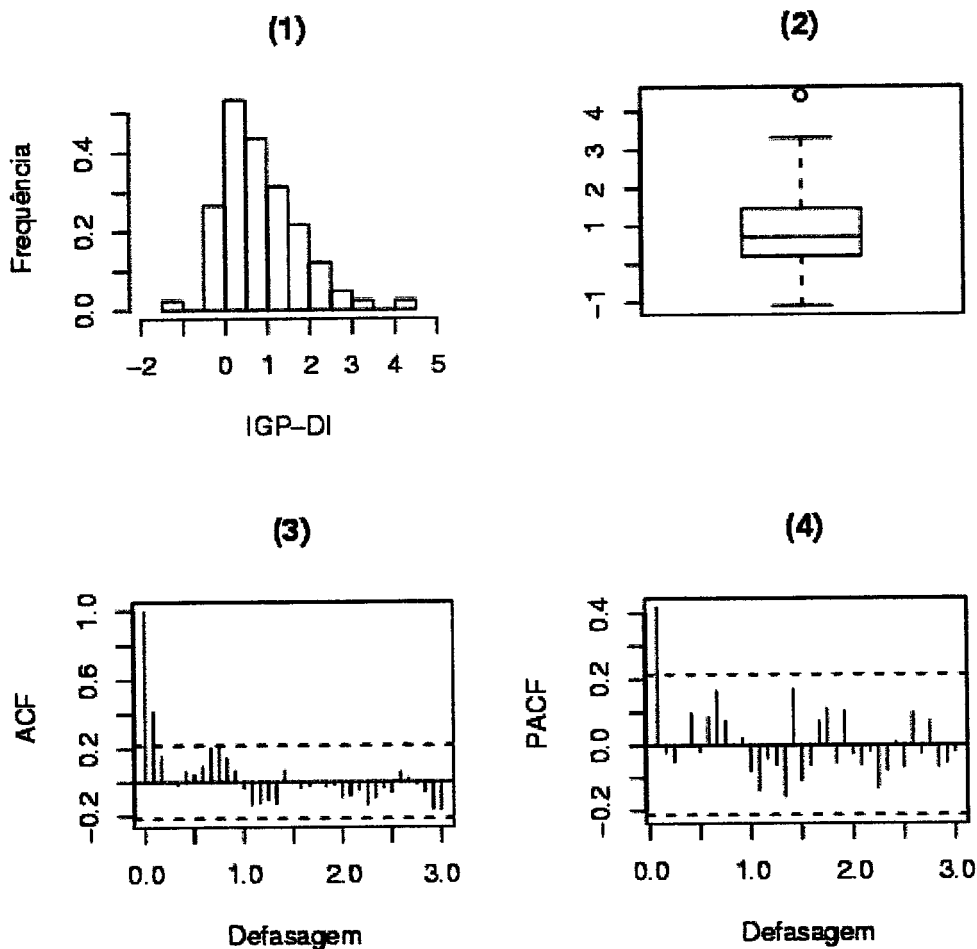


O painel superior da Figura 1 apresenta a série IGP-DI de fevereiro de 1944 a junho de 2001, ao passo que no painel inferior temos a evolução da série entre agosto de 1994 e junho de 2001. Na primeira série observamos uma variação de padrões, possivelmente devida a períodos marcados por características distintas (hiperinflação, planos de estabilização, etc.). Já a segunda série apresenta-se mais homogênea e está inteiramente compreendida em um único período: o do Plano Real. Com o intuito de encontrar modelos que proporcionem boas previsões, consideraremos apenas a segunda série, denominada IGP-DI₉₄₋₀₁.

A característica de simetria apresentada por esta série mais curta pode ser observada tanto no histograma quanto no 'box-plot' apresentados na Figura 2 (painéis superiores). Os dados são ligeiramente assimétricos à direita, havendo apenas uma observação distoante (4,4; fevereiro de 1999, fruto de uma desvalorização cambial). Os coeficientes de assimetria e curtose são aproximadamente iguais a 1,0 e 1,8, respectivamente. A média e o desvio padrão são ambos iguais a 0,9. Nos painéis inferiores da Figura 2 encontram-se o correlograma e o correlograma parcial da série, que apresentam rápido decaimento para zero¹.

¹ Nesses correlogramas, as defasagens vão de 0 a 36, com 36 correspondendo a 3 no eixo horizontal.

Figura 2 - Características da série IGP-DI₉₄₋₀₁



4. Modelagem e previsão

A seguir consideraremos diferentes estratégias de previsão para a série de interesse, nelas incluídas alisamento exponencial, metodologia Box-Jenkins e combinação de previsões, além de uma quarta estratégia que utiliza uma estimativa do núcleo da inflação. É importante ressaltar que no processo de modelagem e previsão nós utilizamos apenas os dados referentes ao período que se estende de agosto de 1994 a dezembro de 2000, reservando, assim, as observações referentes aos meses de janeiro a junho de 2001 para avaliação comparativa das capacidades preditivas das diferentes

estratégias de previsão. Salienta-se, ainda, que todo o processo computacional utilizado fora feito através dos recursos do pacote estatístico R.

4.1. Modelagem e previsão I: alisamento exponencial

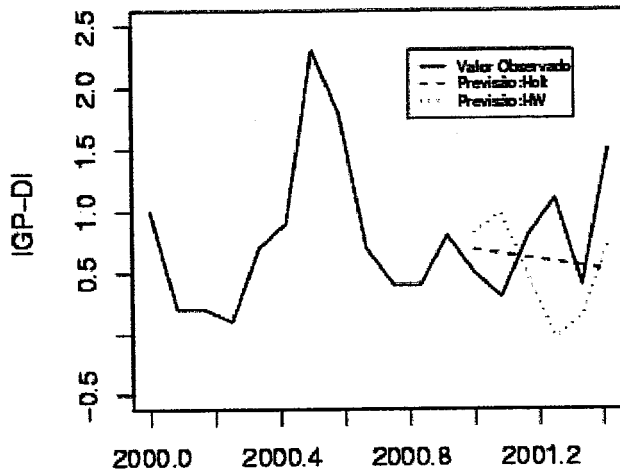
Com a aplicação dos algoritmos de tendência linear de Holt e Holt-Winters (adotando sazonalidade anual, denotada por Holt-Winters₁₂) obtivemos os resultados dispostos no Quadro 1, onde estão expostas as previsões referentes aos meses de janeiro a junho de 2001.

Fazendo-se uma comparação individual mês a mês, é possível notar que somente em junho a previsão do algoritmo de Holt é de qualidade inferior à do algoritmo de Holt-Winters. Essa maior precisão de Holt é evidenciada tanto pelo erro quadrático médio (MSE) quanto pelo erro total (TE). É possível ilustrar a maior precisão das previsões de Holt apresentando-se no mesmo gráfico os dois conjuntos de previsões (Figura 3).

Quadro 1 - Previsões de janeiro a junho de 2001 através dos algoritmos de Holt e Holt-Winters₁₂ aplicados à série IGP-DI₉₄₋₀₁

Tempo	Parâmetro	Holt	Holt-Winters ₁₂
	α	0,84	0,27
	β	0,17	0,06
	γ	-	0,51
	Observado	Previsões	
Jan./2001	0,5	0,6848	0,8219
Fev./2001	0,3	0,6515	0,9790
Mar./2001	0,8	0,6182	0,4711
Abr./2001	1,1	0,5850	-0,0335
Mai./2001	0,4	0,5517	0,1427
Jun./2001	1,5	0,5184	0,7300
-	MSE	0,2404	0,4361
-	TE	-0,9905	-1,4888

Figura 3 - Valores observados e previsões por Holt e Holt-Winters¹² de janeiro a junho de 2001 da série IGP-DI₉₄₋₀₁



4.2. Modelagem e previsão II: método Box-Jenkins

Com o intuito de propor boa(s) representação(ões) para a série em questão, utilizamos o método Box-Jenkins anteriormente comentado. Uma tentativa de identificar um modelo que descreva adequadamente a série IGP-DI poderia incluir o exame das autocorrelações amostrais e parciais amostrais (Figura 2). Todavia, uma forma mais objetiva de identificação envolve a minimização de um critério de escolha, como, por exemplo, os critérios AIC e BIC. Aqui a identificação será feita calculando-se o(s) valor(es) do(s) critério(s) para vários modelos diferentes e escolhendo-se o que apresente o menor valor do critério.

Nós estimamos vários modelos diferentes, associando valores para as ordens p e q de 0 a 4, $d = 0$ ou 1, P e Q variando de 0 a 2, e $D = 0$ ou 1. Essas combinações proporcionam cerca de 900 modelos diferentes. Seis modelos, para cada critério, foram escolhidos de acordo com as seguintes características:

- dois modelos não-sazonais, um com $d = 0$, o outro com $d = 1$, isto é, os modelos $ARIMA(p, 0, q)$ e $ARIMA(p, 1, q)$ com menores AIC e BIC;

- dois modelos sazonais com $d = 0$, variando o grau de diferenciação sazonal D ($D = 0$ ou 1), isto é, os modelos SARIMA($p, 0, q$)($P, 0, Q$)₁₂ e SARIMA($p, 0, q$)($P, 1, Q$)₁₂ com menores AIC e BIC; e
- ainda, dois modelos sazonais com $d = 1$, novamente variando D ($D = 0$ ou 1), isto é, os modelos SARIMA($p, 1, q$)($P, 0, Q$)₁₂ e SARIMA($p, 1, q$)($P, 1, Q$)₁₂ com menores AIC e BIC.

No Quadro 2 estão os seis modelos escolhidos pelos critérios AIC e BIC. Pode-se verificar que AIC e BIC divergiram apenas na seleção do modelo não-sazonal com $d = 1$.

Quadro 2 - Modelos selecionados para a série IGP-DI₉₄₋₀₁ e respectivos valores dos critérios AIC e BIC

Modelo	AIC	BIC
(1,0,0)(0,0,0)	195,0087	199,6963
(3,1,3)(0,0,0)	195,1175	–
(0,1,2)(0,0,0)	–	200,1290
(1,0,0)(0,0,1) ₁₂	196,9843	204,0157
(1,0,0)(0,1,1) ₁₂	190,0043	194,3530
(0,1,2)(0,0,1) ₁₂	197,4533	204,4455
(1,1,1)(0,1,1) ₁₂	191,4896	197,9663

A seguir estão apresentados os sete modelos sugeridos pelos dois critérios:

- modelo ARIMA(1, 0, 0):

$$X_t = 0,9402 + 0,4610X_{t-1} + \varepsilon_t;$$

- modelo ARIMA(3, 1, 3):

$$(1 - B)X_t = 1,897(1 - B)X_{t-1} - 1,396(1 - B)X_{t-2} + 0,285(1 - B)X_{t-3} + \varepsilon_t - 3,086\varepsilon_{t-1} + 3,345\varepsilon_{t-2} - 1,334\varepsilon_{t-3};$$

- modelo ARIMA(0, 1, 2):

$$(1 - B)X_t = \varepsilon_t - 0,798\varepsilon_{t-1} - 0,410\varepsilon_{t-2};$$

- modelo SARIMA(1, 0, 0)(0, 0, 1)₁₂:

$$(1 - 0,4597B)X_t = 0,9362 + (1 + 0,0208B^{12})\varepsilon_t;$$

- modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂:

$$(1 - 0,507B)(1 - B^{12})X_t = (1 + 0,707B^{12})\varepsilon_t;$$

- modelo SARIMA(1, 1, 1)(0, 1, 1)₁₂:

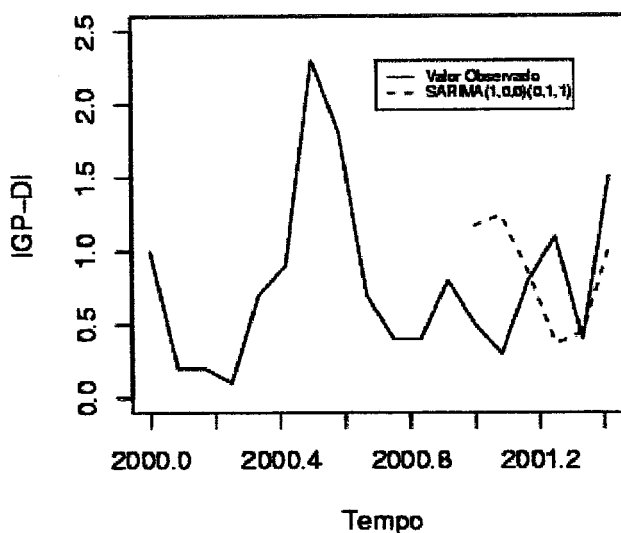
$$(1 - 0,3993B)(1 - B)(1 - B^{12})X_t = (1 + 0,8799B)(1 + 0,7668B^{12})\varepsilon_t; e$$

- modelo SARIMA(0, 1, 2)(0, 0, 1)₁₂:

$$(1 - B)X_t = (1 + 0,798B + 0,410B^2)(1 + 0,016B^{12})\varepsilon_t.$$

O Quadro 3 mostra os modelos selecionados pelos critérios AIC e BIC, com suas respectivas previsões de janeiro a junho de 2001 e duas medidas de precisão: MSE e TE. Como fora mostrado na Seção 2.1, tanto o MSE quanto o TE captam as diferenças entre cada observação da série e a sua respectiva previsão. Assim, quanto mais próximas de 0 (zero) estiverem tais medidas, implica menor divergência entre o que é observado e o que é previsto, conseqüentemente teríamos boas previsões. Logo, com base no MSE e no TE, o modelo com a melhor capacidade preditiva é o modelo ARIMA(3, 1, 3), uma vez que apresentou os valores para ambas as medidas de precisão mais próximos de zero (MSE = 0,179 e TE = 0,148). Dentre todos os modelos considerados, o que apresentou os menores valores dos critérios AIC e BIC foi o modelo SARIMA(1,0,0)(0,1,1), cujas previsões encontram-se apresentadas, graficamente, na Figura 4. Na Figura 5 é apresentada uma comparação da série observada (de agosto de 1994 até dezembro de 2000) com os valores ajustados oriundos do modelo estimado.

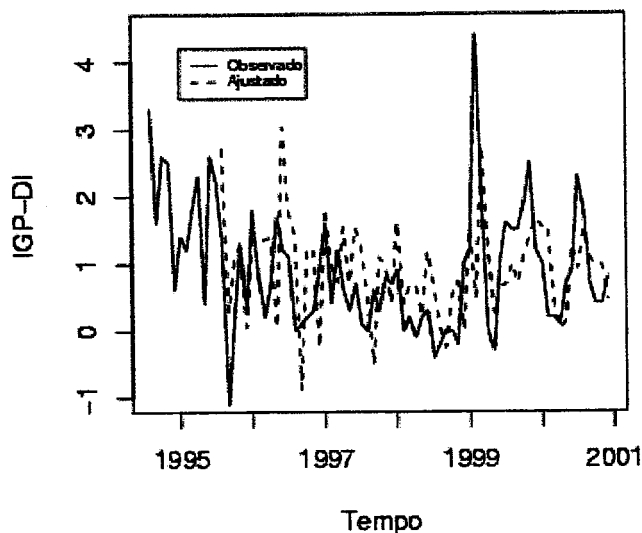
Figura 4 - Valores observados e previsões por SARIMA(1, 0, 0)(0, 1, 1)₁₂ de janeiro a junho de 2001 da série IGP-DI₉₄₋₀₁



Quadro 3 - Previsões de janeiro a junho de 2001 da série IGP-DI_{94.01} através dos modelos selecionados pelos critérios AIC e BIC

Tempo	Modelo Observado	Sazonal						Não-sazonal		
		$d = 0$		$d = 1$		$d = 0$	$d = 1$	$d = 0$	$d = 1$	
		$D = 0$	$D = 1$	$D = 0$	$D = 1$					
		(1,0,0)(0,0,1)	(1,0,0)(0,1,1)	(0,1,2)(0,0,1)	(1,1,1)(0,1,1)	(1,0,0)(0,0,0)	(3,1,3)(0,0,0)	(0,1,2)(0,0,0)		
Previsões										
Jan./2001	0,5	0,8747	1,1720	0,9182	1,1881	0,8756	0,8019	0,9163		
Fev./2001	0,3	0,9224	1,2440	0,9237	1,1725	0,9104	0,7917	0,9077		
Mar./2001	0,8	0,9383	0,8561	0,9243	0,8228	0,9265	0,7832	0,9077		
Abr./2001	1,1	0,9482	0,3699	0,9263	0,3766	0,9339	0,7821	0,9077		
Mai./2001	0,4	0,9389	0,4418	0,9170	0,3874	0,9373	0,7888	0,9077		
Jun./2001	1,5	0,9357	1,0234	0,9136	0,9769	0,9389	0,8006	0,9077		
-	MSE	0,1965	0,3513	0,2034	0,3387	0,1935	0,1791	0,2000		
-	TE	0,9582	0,5072	0,9231	0,3242	0,9226	0,1483	0,8546		

Figura 5 - Valores observados e valores ajustados por SARIMA(1, 0, 0)(0, 1, 1)₁₂ de agosto de 1994 a dezembro de 2000 da série IGP-DI₉₄₋₀₁



Um modelo é considerado adequado a partir da análise dos resíduos padronizados, se os valores destes resíduos estiverem, via de regra, entre -2 e 2 . Nesse sentido, a adequabilidade do modelo considerado é aceita após inspeção da Figura 6, já que apenas um resíduo ultrapassa estas bandas; observa-se ainda a não-significância das autocorrelações amostrais dos resíduos, o que também é desejável. A Figura 7 apresenta os p -valores dos testes *port-manteau* de Ljung-Box e Box-Pierce de aderência do modelo em consideração. Nenhum dos testes apresenta evidências para rejeitar a hipótese nula de que o modelo é adequado.

Figura 6 - Resíduos padronizados e respectivo correlograma do modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂ para a série IGP-DI₉₄₋₀₁

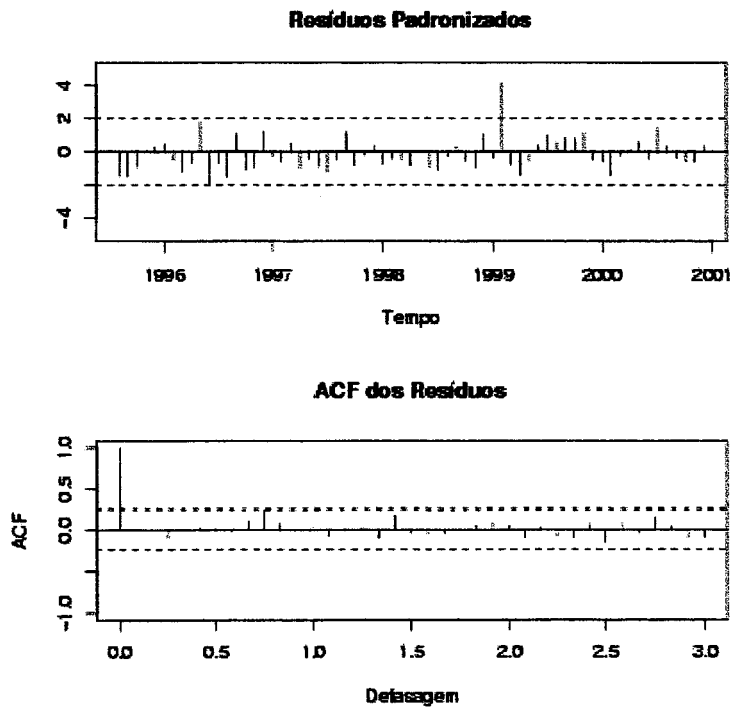
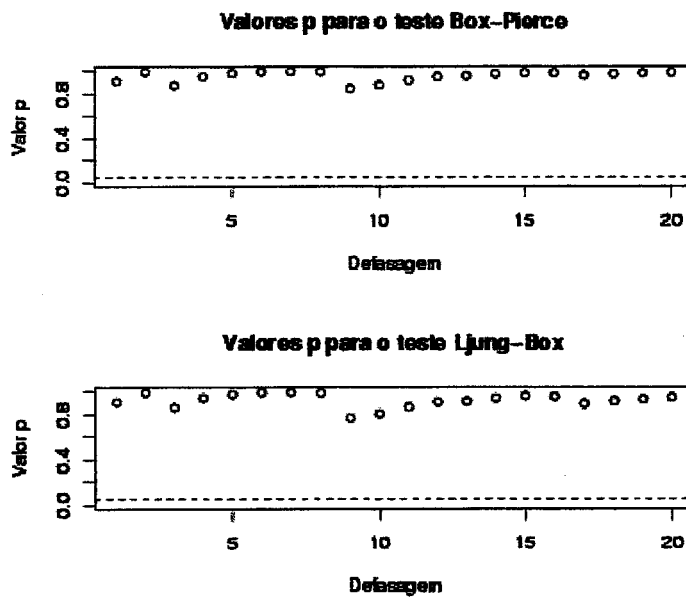


Figura 7- Testes Box-Pierce e Ljung-Box do modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂ para a série IGP-DI₉₄₋₀₁



4.3. Modelagem e previsão III: combinação de previsões

Motivados por resultados empíricos obtidos por vários autores a partir da combinação de previsões, nós combinamos previsões de modelos ARIMA com aquelas obtidas a partir do algoritmo de Holt, e combinamos, também, previsões de modelos SARIMA com as provenientes do algoritmo de Holt-Winters. Foram utilizadas as abordagens abaixo, onde o número entre parênteses indica o método alternativo apresentado em seções anteriores:

- (4): média aritmética simples;
- (8): pesos alternativos, sem constante e sem erros correlacionados;
- (10): pesos alternativos, com constante e sem erros correlacionados; e
- (11): pesos alternativos, com constante e com erros correlacionados.

É importante ressaltar que foram adotados os mesmos valores de $n = 1, 3, 6, 9$ ou 12 para as equações (8), (10) e (11), porém para (10) e (11) os valores utilizados para a constante K não foram somente 1,0; 1,5; 2,0 e 2,5, que são os valores originalmente propostos por Newbold & Granger (1974), mas sim valores entre 1,0 e 3,0 (ambos inclusive) variando de 0,1 em 0,1, ou seja, $K = \{1,0; 1,1; \dots ; 2,9; 3,0\}$. Desta forma, o procedimento de busca implementado foi consideravelmente mais exaustivo do que o utilizado por aqueles autores.

A qualidade das previsões combinadas é avaliada pelo MSE e pelo TE. Foram escolhidas combinações através de quatro critérios, a saber:

- (i) combinação que apresenta o menor MSE, sendo este inferior ao MSE de cada método de previsão individual;
- (ii) combinação que apresenta o menor TE, sendo este inferior ao TE de cada método de previsão individual;
- (iii) combinação que apresenta o menor TE, sendo este inferior ao TE de cada método de previsão individual e sendo o MSE também inferior ao MSE dos métodos de previsão individuais; e

- (iv) combinação que apresenta o menor TE, sendo este inferior ao TE de cada método de previsão individual e com a constante K utilizada variando de acordo com os valores sugeridos por Newbold & Granger (1974).

Em relação aos resultados obtidos, para todas as previsões provenientes dos sete modelos selecionados pelo método Box-Jenkins, bem como todas as previsões oriundas de alisamento exponencial sempre conseguimos encontrar uma previsão combinada com melhor precisão do que as previsões individuais. Adicionalmente, as melhores combinações resultantes foram aquelas provenientes de nossa sugestão de pesos alternativos, ora considerando erros correlacionados e ora não, com K variando de 0,1 em 0,1, ou seja, $K = \{1,0; 1,1; \dots; 2,9; 3,0\}$.

O Quadro 4, a seguir, apresenta as previsões combinadas a partir do modelo selecionado SARIMA(1, 0, 0)(0, 1, 1)₁₂ e de Holt-Winters. Pode-se notar que através da combinação de previsões foi obtida uma ponderação entre os dois métodos que proporcionou tanto um valor inferior do MSE quanto do TE relativamente às previsões individuais dando-se peso 0,8006 para previsões SARIMA e peso 0,1994 para previsões Holt-Winters. Outra combinação que merece destaque é a do modelo SARIMA(0, 1, 2)(0, 0, 1)₁₂ com Holt-Winters, que chegou a fornecer um erro agregado total (TE) de apenas - 0,0003, como pode ser visto no Quadro 5.

4.4. Modelagem e Previsão IV: Previsões com Base no Núcleo da Inflação

Esta quarta abordagem de modelagem e previsão utiliza o conceito de núcleo da inflação, o qual é apresentado a seguir. Após indicarmos uma nova medida de núcleo da inflação, utilizaremos esta como uma variável explicativa em um modelo que combina uma estrutura de regressão com um modelo ARIMA para produzir previsões, ou seja, será sugerido um modelo do tipo ARMAX contendo em sua estrutura de regressão a estimativa do núcleo da inflação que propusemos.

Uma outra forma de mensurar a inflação é através de uma medida chamada de 'núcleo da inflação' ou '*core inflation*'. Não há uma definição consensual do conceito de núcleo da inflação. Aqui, definimos núcleo da inflação como uma medida que procura minimizar os efeitos das componentes mais voláteis, bem como desconsidera

sazonalidades e acidentalidades, procurando revelar um comportamento mais estável que os tradicionais índices de inflação. Vários países e seus respectivos bancos centrais têm utilizado medidas de núcleo da inflação no delineamento de políticas monetárias; ver por exemplo, Hogan, Johnson & Laflèche (2001), Laflèche (1999) e Wynne (1999). No Brasil, a discussão sobre medidas de núcleo da inflação é recente e vem recebendo maior atenção desde junho de 1999, quando foi introduzido o regime de metas inflacionárias.

Nossa proposta é a de produzir uma estimativa do núcleo da inflação através do expurgo de movimentos sazonais e erráticos. Para tanto, os três efeitos devem ser separados, ou seja, é necessário separar o comportamento tendencial da série de possíveis oscilações sazonais e de movimentos erráticos. Para realizar esta decomposição, nós utilizamos o procedimento STL ('Seasonal Decomposition of Time Series by Loess'), que decompõe a série temporal observada nas seguintes componentes não-observadas: sazonal, tendência e irregular. A decomposição é realizada utilizando o método LOESS, que consiste de um ajuste de regressão polinomial local; para detalhes, Cleveland *et al.* (1990). Note que um aspecto desejável da estimação que propomos para o núcleo da inflação, por decomposição não-paramétrica, é que ela não requer a utilização de séries adicionais, necessitando-se apenas das observações temporais da série de interesse.

O resultado da decomposição STL aplicada à série IGP-DI₉₄₋₀₁, que se estende de agosto de 1994 a dezembro de 2000, encontra-se apresentado, graficamente, na Figura 8. O painel superior apresenta o comportamento temporal da série, ao passo que o segundo painel fornece a estimativa da parte sazonal da inflação, o terceiro painel contém o gráfico de nossa estimativa de núcleo da inflação que corresponde à componente não observável de tendência e, finalizando, o painel inferior apresenta a estimativa da componente irregular.

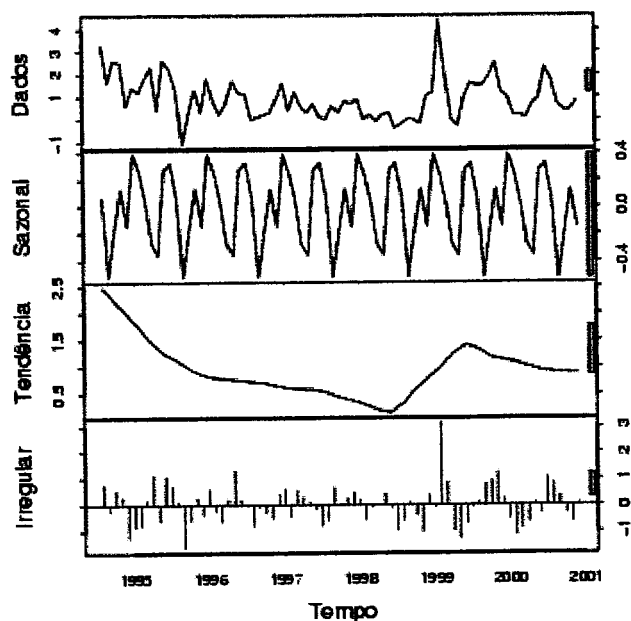
Quadro 4 - Combinação de previsões de janeiro a junho de 2001 do modelo SARIMA (1, 0, 0)(0, 1, 1)₁₂ Holt-Winters₁₂ para a série IGP-DI₁₉₄₋₀₁

Tempo	Observado	Previsões Individuais Utilizadas		Previsão Combinada			
		SARIMA	Holt-Winters ₁₂	Menor MSE	Menor TE	MSE e TE	N & G (1974)*
		Método de Combinação					
		(1,0,0)(0,1,1) ₁₂	$\alpha = 0,27$ $\beta = 0,06$ $\gamma = 0,51$	(11)	(10)	(11)	(10)
				$n = 6, K = 1,5$ $w_1 = 0,8006$ $w_2 = 0,1994$	$n = 3, K = 3$ $w_1 = 0,7143$ $w_2 = 0,2857$	$n = 6, K = 1,5$ $w_1 = 0,8006$ $w_2 = 0,1994$	$n = 3, K = 2,5$ $w_1 = 0,7123$ $w_2 = 0,2877$
Jan./2001	0,5	1,1720	0,8219	1,1022	1,0720	1,1022	1,0713
Fev./2001	0,3	1,2440	0,9790	1,1911	1,1682	1,1911	1,1677
Mar./2001	0,8	0,8561	0,4711	0,7793	0,7461	0,7793	0,7453
Abr./2001	1,1	0,3699	-0,0335	0,2895	0,2547	0,2895	0,2539
Maió/2001	0,4	0,4418	0,1427	0,3821	0,3563	0,3821	0,3557
Jun./2001	1,5	1,0234	0,7300	0,9650	0,9396	0,9650	0,9390
-	MSE	0,3513	0,4361	0,3501	0,3524	0,3501	0,3525
-	TE	0,5072	-1,4888	0,1092	-0,0631	0,1092	-0,0671

Quadro 5 - Combinação de previsões de janeiro a junho de 2001 do modelo SARIMA (0, 1, 2)|(0, 0, 1)₁₂ com Holt-Winters₁₂ para a série IGP-DI₆₄₋₀₁

Tempo	Observado	Previsões Individuais Utilizadas		Previsão Combinada				
		SARIMA	Holt-Winters ₁₂	Menor MSE	Menor TE	MSE e TE	N & G (1974)*	
		Método de Combinação						
				(11)	(11)	(10)	(8)	
			$\alpha = 0,27$ $\beta = 0,06$ $\gamma = 0,51$	$n = 6, K = 1,3$	$n = 6, K = 1,4$	$n = 3, K = 1,7$	$n = 9, K = 1$	
		(0,1,2) (0,0,1) ₁₂		$w_1 = 0,8348$ $w_2 = 0,1652$	$w_1 = 0,6171$ $w_2 = 0,3829$	$w_1 = 0,7878$ $w_2 = 0,2122$	$w_1 = 0,5885$ $w_2 = 0,4115$	
Jan./2001	0,5	0,9182	0,8219	0,9023	0,8814	0,8978	0,8786	
Fev./2001	0,3	0,9237	0,9790	0,9328	0,9448	0,9354	0,9464	
Mar./2001	0,8	0,9243	0,4711	0,8494	0,7508	0,8281	0,7378	
Abr./2001	1,1	0,9263	-0,0335	0,7678	0,5589	0,7226	0,5314	
Maió/2001	0,4	0,9170	0,1427	0,7891	0,6206	0,7527	0,5984	
Jun./2001	1,5	0,9136	0,7300	0,8833	0,8433	0,8746	0,8381	
-	MSE	0,2034	0,4361	0,2011	0,2227	0,2034	0,2277	
-	TE	0,9231	-1,4888	0,5248	-0,0003	0,4112	-0,0693	

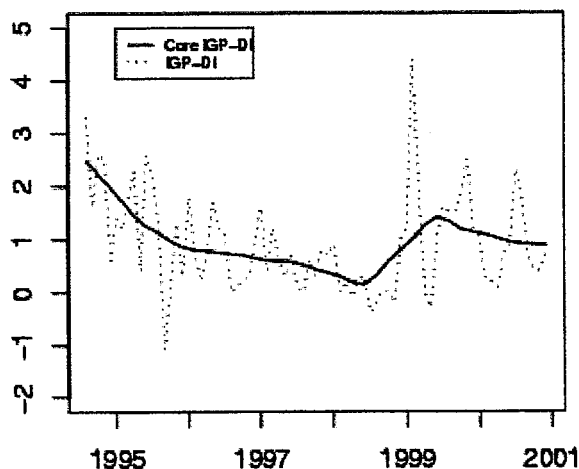
Figura 8 - Decomposição STL da série IGP-DI₉₄₋₀₁ entre agosto de 1994 e dezembro de 2000.



É interessante notar, inspecionando a estimativa do núcleo inflacionário, que a inflação apresenta tendência de continuada queda até julho de 1998 (0,23), quando há uma inflexão e a tendência passa a ser de aceleração, o que se mantém até julho de 1999 (1,41), quando o comportamento de desaceleração é retomado.

Na Figura 9 comparamos nossa estimativa de núcleo da inflação à série IGP-DI₉₄₋₀₁ no período entre agosto de 1994 e dezembro de 2000. Notamos que a medida do núcleo da inflação apresenta uma dinâmica mais suave e menos sujeita a flutuações temporárias do que a medida usual de inflação, como desejado.

Figura 9 - Comparação entre o núcleo da inflação e a série IGP-DI₉₄₋₀₁ entre agosto de 1994 e dezembro de 2000



Na tentativa de gerar boas previsões, propomos uma nova estratégia de previsão: a utilização da estimativa de núcleo da inflação como regressor em uma estrutura dinâmica de modelagem. Aqui, nós utilizamos o núcleo da inflação como uma variável independente, denominada Z_t , dentro de uma estrutura SARIMAX para representação dos dados. A motivação é que a medida de núcleo da inflação deve conter sinalização de períodos onde a tendência da inflação é de aceleração e de períodos onde há desaceleração e esta informação pode ser eficientemente utilizada dentro de uma estrutura dinâmica de modelagem a fim de direcionar as previsões geradas.

A busca de modelo(s) SARIMAX que proporcione(m) boa representatividade dos dados foi feita de forma semelhante à utilizada na metodologia Box-Jenkins. Recorremos aos critérios AIC e BIC para selecionar modelo(s) SARIMAX que represente(m) adequadamente a série de interesse. O modelo SARIMAX que apresentou o menor valor do critério AIC com $d=0$ foi selecionado. O procedimento foi repetido, porém com o uso do critério BIC. Além de selecionados dois modelos ARIMAX, um por cada critério, foram selecionados também dois modelos SARIMAX, sendo o primeiro através do critério AIC com $d=0$, $D=0$ e apresentando termos sazonais e o outro através do critério BIC com $d=0$, $D=0$ e apresentando termos sazonais. O Quadro 6 apresenta os modelos selecionados pelos critérios.

Quadro 6 - Modelos ARIMAX e SARIMAX para a série IGP-DI₉₄₋₀₁ selecionados pelos critérios AIC e BIC

Modelo	Critério	Valor do critério
(2, 0, 2)	AIC	160,0710
(2, 0, 1)	BIC	171,8844
(2, 0, 1)(1, 0, 0) ₁₂	AIC	162,1215
(2, 0, 1)(1, 0, 0) ₁₂	BIC	176,1843

Notamos que foram selecionados três modelos distintos pelos critérios AIC e BIC, havendo divergência somente na escolha do modelo não-sazonal. Após a estimação dos parâmetros tem-se, por exemplo, o seguinte modelo ARIMAX(2,0,2) estimado:

$$X_t = -0,019 + 1,286X_{t-1} + -0,675X_{t-2} + \varepsilon_t - 1,375\varepsilon_{t-1} + 0,384\varepsilon_{t-2} + 1,009Z_t$$

Escolhidos os modelos ARIMAX/SARIMAX, necessitamos de valores referentes ao comportamento futuro do núcleo da inflação a fim de produzir previsões da dinâmica inflacionária utilizando a estratégia proposta. Uma vez que o núcleo da inflação é livre de oscilações sazonais, que foram expurgadas no processo de decomposição não-paramétrica, o procedimento que sugerimos é a utilização de previsões para os valores futuros do núcleo por Holt, que é o algoritmo de alisamento exponencial indicado para séries com tendência local.

Com a aplicação do algoritmo de Holt à estimativa do núcleo da inflação, obtemos previsões tais como as dispostas no Quadro 7, onde vemos os valores selecionados para α e β , bem como as previsões de janeiro a junho de 2001, valores estes que serão utilizados pelos modelos ARIMAX e SARIMAX na previsão do IGP-DI.

Quadro 7 - Previsões de janeiro a junho de 2001 do núcleo da inflação através do algoritmo de Holt

Parâmetro	Holt
α	0,95
β	1,00
Tempo	Previsões
Jan./2001	0,87509
Fev./2001	0,87096
Mar./2001	0,86682
Abr./2001	0,86269
Mai./2001	0,85855
Jun./2001	0,85442

Finalmente, com os modelos ARIMAX/SARIMAX estimados e de posse das previsões de janeiro a junho do núcleo da inflação obtidas através do algoritmo de Holt, podemos gerar previsões para a série IGP-DI de janeiro a junho de 2001. O Quadro 8 apresenta as previsões dos três modelos ARIMAX/SARIMAX selecionados e ainda os graus de precisão destas mensurados pelo MSE e pelo TE.

Quadro 8 - Previsões de janeiro a junho de 2001 da série IGP-DI₉₄₋₀₁ através dos modelos ARIMAX/SARIMAX selecionados

Tempo	Modelo	(2, 0, 2)	(2, 0, 1)	(2, 0, 1)(1, 0, 0) ₁₂
	Observado	Previsões		
Jan./2001	0,5	0,7807	0,8584	0,8768
Fev./2001	0,3	0,8393	0,8875	0,8868
Mar./2001	0,8	0,9193	0,8864	0,8766
Abr./2001	1,1	0,9473	0,8691	0,8470
Mai./2001	0,4	0,9276	0,8501	0,8388
Jun./2001	1,5	0,8817	0,8376	0,8315
-	MSE	0,1780	0,1960	0,1993
-	TE	0,6959	0,5891	0,5575

Analisando os resultados apresentados no Quadro 8, vemos que os três modelos superestimam o IGP-DI entre janeiro e junho de 2001, em média em 0,6. O modelo que apresentou melhor desempenho de acordo o MSE foi o modelo ARIMAX(2, 0, 2); já em relação ao TE, o melhor modelo foi o SARIMAX(2, 0, 1)(1, 0, 0)₁₂.

5. Comparação de resultados

Através da aplicação do método de alisamento exponencial verificou-se que tanto o algoritmo de Holt quanto o de Holt-Winters conduzem a previsões que subestimam os valores do IGP-DI. O algoritmo de Holt produziu previsões mais precisas que o algoritmo de Holt-Winters (vide Quadro 1).

Em relação à metodologia Box-Jenkins, os critérios AIC e BIC foram utilizados no processo de identificação e foram selecionados seis modelos por cada um desses critérios. Houve coincidência em cinco deles, ou seja, ao todo foram escolhidos sete modelos. Tanto o AIC quanto o BIC de menor valor apontaram o modelo SARIMA(1,0,0)(0,1,1)₁₂ como o de maior capacidade preditiva. Diferentemente de Holt, a metodologia Box-Jenkins sugeriu um padrão de sazonalidade. O erro total de previsão

(TE = 0,5072, que equivale a 11,03% em termos relativos) correspondente ao modelo SARIMA(1,0,0)(0,1,1)₁₂ é positivo, indicando assim superestimação dos valores observados, mas em menor grau (em termos absolutos) do que o erro obtido de Holt (TE = - 0,9905, que equivale a 21,53% em termos relativos).

O modelo ARIMA(3, 1, 3), mesmo não sendo o indicado pelos critérios AIC e BIC, foi o modelo que apresentou melhores medidas de precisão de previsão (TE = 0,1483, que equivale a 3,22% em termos relativos, e MSE = 0,1791).

Utilizando o método de combinação de previsões, foram obtidos resultados satisfatórios combinando-se previsões de modelos ARIMA com Holt e previsões de modelos SARIMA com Holt-Winters. A combinação que proporcionou a melhor precisão das previsões, considerando o MSE, foi a combinação ARIMA(3, 1, 3) com Holt (MSE = 0,1760), sendo este o menor MSE encontrado. Levando-se em conta o erro total de previsão (TE), a combinação mais satisfatória foi a que utilizou o modelo SARIMA(0,1,2)(0,0,1)₁₂ juntamente com Holt-Winters, proporcionando TE = - 0,0003 (uma subestimação insignificante). Também o modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂, que foi o modelo que apresentou o menor valor dos critérios AIC e BIC, quando combinado com Holt-Winters apresentou boas previsões. Por exemplo, uma das combinações resultantes revelou um TE de - 0,0631 (Quadro 4), que equivale a apenas 1,37% em termos relativos, ou seja, uma pequena subestimação. Vale ressaltar que a grande maioria das combinações apresentou algum tipo de melhoria em relação à precisão de previsões individuais.

A quarta abordagem utilizada buscou selecionar modelos ARIMAX/SARIMAX usando uma estimativa do núcleo da inflação em uma estrutura de regressão. Esta estimativa foi construída de forma não-paramétrica. Três modelos foram selecionados através dos critérios AIC e BIC: ARIMAX(2, 0, 2), ARIMAX(2, 0, 1) e SARIMAX(2, 0, 1) (1, 0, 0)₁₂ (Quadro 6). Verificamos que estes modelos apresentaram resultados satisfatórios principalmente, em relação ao MSE; o modelo ARIMAX(2, 0, 2), por exemplo, apresentou MSE de 0,1780, sendo este um dos menores valores encontrados. Já em relação ao TE, todos os três modelos superestimaram, de forma agregada, o IGP-DI. Dentre os três modelos, o que conduziu ao menor TE foi o modelo SARIMAX(2,0,1) (1,

0, 0)₁₂, sendo este 0,5575, ou seja, um erro agregado da ordem de 12,12% do valor observado.

Com base nessa breve discussão das previsões provenientes das quatro estratégias de modelagem e previsão, nota-se que os algoritmos de alisamento exponencial não fornecem boas previsões da dinâmica inflacionária.

Já pela abordagem mais formal da metodologia Box-Jenkins, foram examinados cerca de 900 modelos e foram utilizados os critérios AIC e BIC para seleção de modelos que representam adequadamente a série IGP-DI₉₄₋₀₁. Encontrou-se boa qualidade nas previsões; por exemplo, com o modelo ARIMA(3, 1, 3), tanto o MSE quanto o TE são os menores em relação aos sete modelos ARIMA/SARIMA selecionados. O modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂ foi o que apresentou o menor valor dos critérios AIC e BIC, conduzindo também a boas previsões.

Na combinação de previsões buscamos unir a simplicidade da metodologia de alisamento exponencial com a formalidade da metodologia Box-Jenkins, esperando que cada método pudesse contribuir com parcelas representativas na construção de uma previsão combinada para o comportamento dinâmico da inflação brasileira. De fato, obtivemos êxito, com, por exemplo, com a combinação das previsões do modelo SARIMA(0, 1, 2)(0, 0, 1)₁₂ com aquelas obtidas por Holt-Winters, que rendeu uma previsão agregada de janeiro a junho praticamente sem erro, uma vez que o erro total foi de apenas - 0,0003. O modelo de menor valor dos critérios, o SARIMA(1, 0, 0)(0, 1, 1)₁₂, combinado com Holt-Winters também proporcionou resultados bons: erro total de - 0,0631, o que corresponde a 1,37% em termos relativos; o MSE dessa previsão combinada foi 0,3524.

A quarta metodologia empregada envolve modelos ARIMAX e SARIMAX. A novidade aqui foi uma nova proposta que fizemos para estimação do núcleo da inflação através de uma decomposição não-paramétrica. Utilizamos esta medida como variável regressora em modelos ARIMAX/SARIMAX. Obtivemos bons resultados, principalmente em relação à medida MSE; em relação ao TE não obtivemos o mesmo grau de sucesso alcançado com combinação de previsões. Assim, acreditamos que a combinação de previsões de alisamento exponencial com Box-Jenkins (modelos ARIMA/SARIMA) é uma

boa estratégia de previsão para a dinâmica inflacionária, pois proporciona resultados aceitáveis do ponto de vista de precisão. Deve-se utilizar, na combinação de previsões, o modelo ARIMA/SARIMA selecionado através de um critério de escolha de modelos; o modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂ foi o modelo indicado pelos dois critérios utilizados neste trabalho. Portanto, a combinação sugerida é a combinação do modelo SARIMA(1,0,0)(0,1,1)₁₂ com o algoritmo de Holt-Winters; como vimos, esta combinação proporciona previsões de boa qualidade. Enfim, sugerimos o uso da combinação do algoritmo Holt-Winters, com peso 0,2857, conjuntamente com modelo SARIMA(1,0,0)(0,1,1)₁₂, com peso 0,7143, onde este modelo SARIMA(1,0,0)(0,1,1)₁₂ tem como parâmetros $\phi_1 = 0,507$ e $\Theta_1 = -0,707$; o algoritmo Holt-Winters indicado tem como parâmetros $\alpha = 0,27$; $\beta = 0,06$ e $\gamma = 0,51$.

6. Conclusão

Com o objetivo principal de buscar previsões suficientemente precisas para a dinâmica inflacionária brasileira, consideramos quatro estratégias de modelagem/previsão. As abordagens utilizadas foram alisamento exponencial, com os algoritmos de Holt e Holt-Winters, a metodologia Box-Jenkins, com o uso dos modelos ARIMA e SARIMA, combinação de previsões com a utilização dos algoritmos de Holt e Holt-Winters conjuntamente com os modelos ARIMA e SARIMA e, como quarta abordagem, o uso de modelos ARIMAX e SARIMAX.

As previsões obtidas pelos algoritmos de Holt e Holt-Winters foram as menos precisas. Com o emprego da metodologia Box-Jenkins, foram obtidos resultados melhores que os proporcionados pelo método anterior. Com o uso de modelos ARIMAX e SARIMAX, foram encontradas previsões satisfatórias, porém sugeriu-se o uso da combinação de previsões de Holt-Winters e SARIMA, pois com o modelo recomendado pelos critérios AIC e BIC combinado com Holt-Winters as previsões obtidas foram mais precisas do que as previsões oriundas dos dois modelos individuais e ainda das provenientes de modelos ARIMAX/SARIMAX.

Os resultados contidos no presente artigo abrem espaço para pesquisas futuras.

Primeiro, seria interessante investigar a capacidade preditiva de modelos de integração fracionária, onde as ordens de integração e integração sazonal fossem estimadas a partir dos dados. Segundo, seria igualmente interessante investigar o desempenho de previsões obtidas por modelagem de redes neurais. Por fim, trabalhos futuros devem investigar o desempenho, em outras situações, da metodologia de previsão proposta no presente artigo envolvendo a combinação de modelagem dinâmica com uma estrutura de regressão que engloba uma medida do comportamento tendencial da série.

Referências bibliográficas

- ANSLEY, C.F. & NEWBOLD, P. (1980). Finite sample properties of estimators for autoregressive-moving average models. *Journal of Econometrics*, 13, 159-183.
- BARNARD, G.A. (1963). New methods of quality control. *Journal of the Royal Statistical Society A*, 126, 255-259.
- BATES, J.M. & GRANGER, C.W.J. (1969). The combination of forecasts. *Operations Research Quarterly*, 20, 451-468.
- BOX, G.E.P. & JENKINS, G.M. (1976). *Time Series Analysis: Forecasting and Control*, Revised Edition. San Francisco: Holden-Day.
- BRASIL, Ministério da Fazenda, Secretaria de Política Econômica, *Relatório de Gestão 2001*. (2002).
- BROCKWELL, P.J. & DAVIS, R.A. (1996). *Introduction to Time Series and Forecasting*. New York: Springer-Verlag.
- CLEVELAND, R.B., CLEVELAND, W.S., MCRAE, J.E. & TERPENNING, I. (1990). STL: a seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6, 3-73.
- GRANGER, C.W.J. & NEWBOLD, P. (1986). *Forecasting Economic Time Series*, 2ª ed. New York: Academic Press.
- HOGAN, S., JOHNSON, M. & LAFLÈCHE T. (2001). Core inflation. *Bank of Canada Research Department Technical Report*, 89.
- LAFLÈCHE A. (1999). Statistical measures of the trend of inflation. *Bank of Canada Research Department Working Paper*.
- MILLS, J.A. & PRASAD, K. (1992). A comparison of model selection criteria. *Econometric Reviews*, 11, 201-233.
- NEWBOLD, P. & BOS, T. (1994). *Introductory Business and Economic Forecasting*, 2ª ed. Cincinnati: South-Western Publishing.
- NEWBOLD, P. & GRANGER, C.W.J. (1974). Experience with forecasting univariate time series and combination of forecasts. *Journal of the Royal Statistical Society A*, 137, 131-146.
- REID, D.J. (1969). *A Comparative Study of Time Series Prediction Techniques on Economic Data*. Ph.D. Thesis, Department of Mathematics, University of Nottingham.
- WYNNE M.A. (1999). Core inflation: a review of some conceptual issues. Federal Reserve Bank of Dallas: Research Department, *mimeo*.

Abstract

We consider different forecasting strategies for producing in forecasts using Brazilian data. A new proposal involves the use a nonparametric decomposition to estimate the core inflation component, and then its usage in a regression structure within a dynamic model used to forecast the future path of inflation. We also consider combination of forecasts. The results favor the approach based on combining forecasts obtained from different forecasting strategies.

Avaliação de critérios de convergência univariados para o método de Monte Carlo via Cadeias de Markov

*Denismar Alves Nogueira**

*Thelma Sáfadi***

*Daniel Furtado Ferreira****

Resumo

Os métodos de Monte Carlo via Cadeias de Markov têm sido estudados com aplicações em diversas áreas, mas uma das maiores dificuldades é determinar o tamanho adequado da amostra, ou seja, a convergência do processo para assim inferir sobre os parâmetros da distribuição de interesse. Apesar de existirem vários critérios na literatura, os objetivos foram: propor uma forma alternativa de aplicação combinada de quatro critérios univariados de convergência das seqüências de Markov (Gelman e Rubin, 1992; Raftery e Lewis, 1992a; Geweke, 1992; Heidelberger e Welch, 1983), otimizando suas propriedades. Para isso dados referentes a dois modelos foram utilizados, sendo um de séries temporais com duas intervenções e erro ARMA(2,2) e o outro uma normal trivariada, considerando três situações distintas para a matriz de covariâncias. Em ambos os casos foram utilizados o Amostrador de Gibbs e aplicados os critérios para monitorar a convergência. Os resultados obtidos permitiram que se concluísse que: os critérios, por sugestão, devem ser aplicados no monitoramento da convergência de forma combinada, para aproveitar suas melhores características.

Palavras-chave: Amostrador de Gibbs, correlação e séries temporais.

* Aluno de pós-graduação do curso de Estatística e Experimentação Agropecuária do Deptº. de Ciências Exatas da Univ. de Federal de Lavras, Lavras, MG, Cep.37200-000. e-mail: denisnog@yahoo.com.br. Bolsista CAPES.

** Professora Adjunta IV do Deptº. de Ciências Exatas da Univ. Federal de Lavras, Lavras, MG, Cep. 37200-000. e-mail: safadi@ufla.br.

*** Professor Adjunto IV do Deptº. de Ciências Exatas da Univ. Federal de Lavras, Lavras, MG, Cep. 37200-000. e-mail: danielff@ufla.br, Bolsista CNPq.

1. Introdução

Um assunto de grande interesse na atualidade é a inferência Bayesiana, referindo-se a um modo de se fazer estatística criada muito antes da análise que é utilizada atualmente. A inferência Bayesiana ficou durante um grande período esquecida devido à necessidade de se calcular analiticamente integrais complicadas. Muitas vezes estas eram inviáveis de serem feitas de forma algébrica. Os métodos de Monte Carlo via Cadeias de Markov (MCMC) têm sido aplicados na obtenção das distribuições marginais de forma iterativa, nas situações em que a forma algébrica é inviável. Estudos com aplicações em diversas áreas utilizando MCMC têm sido realizados, mas uma das maiores dificuldades é a verificação empírica da convergência da seqüência à distribuição de equilíbrio. Esta definição é a mais geral noção de convergência para um processo iterativo. A palavra "convergência" pode ter várias interpretações, pois não significa um simples número ou toda uma distribuição, mas sim uma amostra de uma distribuição. O processo Markoviano sustenta uma memória curta, sendo a variável gerada só dependente da variável gerada anteriormente. Desta forma, a seqüência apresenta uma autocorrelação. O problema em se utilizar amostras com presença de autocorrelação é que a estimativa do erro padrão da média será viesada, pois a variância da média é viesada. Estimadores e detalhes para esta variância foram dados em Carlin e Louis (2000). Para a utilização do Amostrador de Gibbs é necessário que se tenham as condicionais completas. A correlação cruzada refere-se à correlação existente entre as condicionais completas. Se a correlação for alta, ou seja, se parâmetros são altamente correlacionados, o passo do amostrador pode ser muito curto, acarretando lentidão no passeio pela distribuição. Com isto, a convergência fica lenta. Este fato está relacionado diretamente com a aceleração do Amostrador de Gibbs. Segundo Gamerman (1996), a reparametrização é uma forma de resolver este tipo de situação, reduzindo o tempo de convergência. Outra maneira seria construir blocos com as variáveis altamente correlacionadas, formando uma única condicional completa para um bloco, estimando-as de maneira multivariada. Liu *et al.* (1994) mostraram que estimativas obtidas para componentes de blocos foram sempre mais precisas que quando as componentes foram tratadas isoladamente, pois se podem considerar as covariâncias entre as variáveis.

Existem várias técnicas que auxiliam na identificação e monitoração da convergência da seqüência. Alguns trabalhos abordam o tema e afirmam que todas estas técnicas pecam, umas por serem inviáveis de se implementar na prática, outras, em alguns casos específicos de uma dada situação, apresentam resultados contraditórios (Cowles e Carlin, 1996 e Brooks e Roberts, 1998).

Dos critérios propostos na literatura optou-se pelos critérios mais populares. Os critérios utilizados foram: Gelman e Rubin, 1992; Geweke, 1992; Raftery e Lewis, 1992a; Heidelberger e Welch, 1983. O critério de Gelman e Rubin utiliza seqüências em paralelo, os demais utilizam seqüências simples. O objetivo deste trabalho foi propor uma forma alternativa de aplicação combinada de quatro critérios univariados de convergência das seqüências de Markov, otimizando suas propriedades.

Na aplicação dos testes usar-se-á dois modelos com dados simulados e uma aplicação com dados reais.

Os softwares utilizados foram o SAS e o software R com algumas rotinas do pacote BOA "Bayesian Output Analysis".

2. Verificação formal da convergência

2.1. Critério de Gelman e Rubin (1992)

O algoritmo MCMC tem produzido um significativo impacto no alcance de problemas para os quais a análise Bayesiana pode ser aplicada (Gilks, Richardson e Spiegelhalter; 1996). O método envolve simulações provindas de uma complexa e, geralmente, multivariada distribuição de interesse, $P(\theta)$, indiretamente gerado de uma Cadeia de Markov com a densidade de interesse, uma densidade estacionária. Geralmente são simuladas $m \geq 2$ seqüências de simulações, cada uma de tamanho n , $(\theta_{j1}, \theta_{j2}, \dots, \theta_{jn})$, para $j = 1 \dots m$ (sendo cada θ_{jt} um vetor). Se $m > 1$, as m seqüências são usualmente, mas nem sempre, simuladas independentemente.

O método que utiliza múltiplas repetições de seqüências para decidir se a estacionariedade foi alcançada dentro da segunda metade de cada amostra foi proposta

por Gelman e Rubin (1992). O método pressupõe que m seqüências tenham sido simuladas em paralelo, cada uma partindo de diferentes pontos iniciais. Tendo obtido os pontos iniciais satisfatórios, ou seja, pertencentes ao espaço paramétrico da *posteriori*, as seqüências são geradas para $2n$ iterações, das quais as primeiras n são descartadas para evitar o período de aquecimento (*burn-in*), ou seja, para evitar a influência dos valores iniciais. As m seqüências rendem m possíveis inferências. Se estas inferências são bastante similares, é um indicativo de que a convergência foi alcançada ou está próxima. Gelman e Rubin (1992) sugerem comparar estas inferências produzidas com as obtidas com a união das seqüências, ou seja, com mn valores.

Considere uma variável aleatória θ que tem média μ e variância σ^2 , sob a distribuição de interesse; considere também que se tem um estimador $\hat{\mu}$ não viesado para μ . A variação entre seqüência E/n e a variação dentro de seqüência D , são definidas por:

$$E/n = \frac{1}{m-1} \sum_{j=1}^m (\bar{\theta}_{j\cdot} - \bar{\theta}_{\cdot\cdot})^2 \quad (2.1)$$

$$D = \frac{1}{m(n-1)} \sum_{j=1}^m \sum_{i=1}^n (\theta_{ji} - \bar{\theta}_{j\cdot})^2, \quad (2.2)$$

sendo θ_{ji} o i -ésimo valor de um total de n de θ na seqüência j , $\bar{\theta}_{j\cdot}$ a média da seqüência j e $\bar{\theta}_{\cdot\cdot}$ a média geral.

Note que a análise de variância (ANOVA) assume um *pool* de variâncias total que pode ser produzido porque, sobre convergência, a variação total das seqüências é igual, de fato, à variação dentro de seqüência, pois a variação entre seqüências se anula.

Depois de calculadas as estimativas anteriores, pode-se estimar $\hat{\sigma}^2$ pela média de E e D :

$$\hat{\sigma}^2 = \frac{n-1}{n} D + \frac{E}{n}, \quad (2.3)$$

a qual pode ser uma estimativa não viesada da verdadeira variância se os valores iniciais das seqüências provirem da distribuição de interesse. O $\hat{\sigma}^2$ é uma ferramenta que pode levar a falsos diagnósticos de convergência.

Considerando a variância do estimador um *pool* de variâncias:

$$\hat{V} = \hat{\sigma}^2 + \frac{E}{mn}, \quad (2.4)$$

a comparação entre as variâncias é expressa por:

$$R = \frac{\hat{V}}{\sigma^2}, \quad (2.5)$$

o qual chama-se fator de redução de escala (da sigla inglesa SRF). O denominador de R não é conhecido, devendo ser estimado dos dados; pode-se, por isso, ganhar uma superestimativa de R , pois σ^2 será subestimada pelo D . Por isso, utiliza-se uma superestimativa de R :

$$R = \frac{\hat{V}}{D} = \frac{m+1}{m} \left(\frac{\hat{\sigma}^2}{D} \right) - \frac{n-1}{mn}, \quad (2.6)$$

o qual é conhecido como fator de redução de escala potencial (da sigla inglesa PSRF) e pode ser interpretado como um fator de diagnóstico da convergência. Quando o \hat{R} for grande, sugere-se que quaisquer das estimativas das variâncias podem ser reduzidas com a utilização de mais iterações, desde que as seqüências simuladas ainda não tenham feito um passeio por toda a distribuição de interesse. Se PSRF aproxima-se de 1, pode-se concluir que cada m conjunto de n iterações se aproxima da distribuição de interesse. A estimativa para D parece ser razoável desde que não estimada com n pequeno, mas será mais acurada à medida que n aumenta. O uso do E também não será muito satisfatório quando se utiliza $m-1$ graus de liberdade, o qual é, em geral, muito pequeno. O critério é de fácil interpretação e geralmente fidedigno quando aplicado em situações de aproximada simetria e unimodais, mas pode ser muito falho se aplicado a problemas multimodais (Brooks *et al.*, 1995).

A limitação principal para o método é a suposição de normalidade da distribuição marginal de cada parâmetro. Uma transformação apropriada é sugerida, mas esta não é muito atrativa, pois métodos MCMC utilizam freqüentemente distribuições não normais (Gelman e Rubin, 1992).

No trabalho inicial de Gelman e Rubin (1992), é adotado incorretamente o fator de correção aplicado à variância para a variabilidade amostral, isto é, $d/(d-2)$, em que d representa os números de graus de liberdade (Brooks e Gelman, 1998). O uso deste fator conduziu a um número de problemas, em que o corrigido pode ser infinito ou até mesmo negativo. Então, seguindo o método de Fisher (1935), baseado em amostras independentes provindas da distribuição normal, o parâmetro-chave para este ajustamento são os graus de liberdade d . Para informar corretamente a variabilidade amostral o fator correto é:

$$\hat{R} = \left(\frac{d+3}{d+1} \right) \frac{\hat{V}}{D}, \quad (2.7)$$

esta correção será importante para pequenas amostras.

Caso o processo indique convergência, deve-se lembrar que os 50% finais da seqüência são valores amostrados da marginal desejada. Gelman (1995) sugere aceitar como garantia de convergência valores de \hat{R} menores que 1,1 ou 1,2.

2.2. Critério de Raftery e Lewis (1992a)

O critério sugerido por Raftery e Lewis (1992a) é um método que estima quantas iterações são necessárias para que o Amostrador de Gibbs apresente convergência à distribuição estacionária, propondo especificações de quantidades necessárias para isto. Deve-se ter uma seqüência do Amostrador de Gibbs com um N mínimo, que é o número mínimo de iterações requeridas para se obter a precisão necessária de estimação. O método fornece as estimativas do *burn-in* (B), que é o número de iterações que devem ser descartadas, o número de iterações que devem ser computadas N (total) e o k , que é a distância mínima de uma iteração à outra para se obter a

subamostra aproximadamente independente (*thin*). Esses valores serão calculados mediante especificações para garantir que um quantil u de uma determinada função $f(\theta)$ seja estimado com precisão pré-definida. Seja um exemplo no qual se deseja obter $q = P(f(\theta) \leq u | y) = r = 0,025$, sendo u o limite inferior do intervalo de confiança com probabilidade $s = 95\%$. Pode-se desejar que o valor dessa probabilidade seja calculado com 0,01 e probabilidade 0,99. Desta forma, o intervalo de confiança de 95% está sendo calculado realmente *a posteriori* entre 93% e 97%, com 99% de confiança.

Uma cadeia binária (0 ou 1) é formada segundo $Z_i = I(f(\theta)_{(i)} \leq u)$, sendo $f(\theta)_{(i)}$ o valor gerado da i -ésima iteração. Pode-se construir novas seqüências tomando-se valores de Z_i a cada k iterações, sendo $Z_i^{(k)} = Z_{i+(i-1)k}$, supondo que as dependências decaem rapidamente com a distância. A aproximação é baseada no modelo de Markov de segunda ordem, ou seja, a dependência é em dois estágios passados. Toma-se k como o menor valor para o qual a dependência em apenas um estágio é preferida.

Seja também

$$\begin{bmatrix} 1-\alpha & \beta \\ \beta & 1-\alpha \end{bmatrix},$$

a matriz de transição de $Z_i^{(k)}$. Com a definição do k , determina-se $B = mk$, sendo,

$$m = \frac{\log\left(\frac{\varepsilon(\alpha + \beta)}{\max(\alpha, \beta)}\right)}{\log(1 - \alpha - \beta)},$$

em que ε é a distância requerida entre $p(Z_B^{(k)} = t | Z_0^{(k)} = j)$ e a probabilidade de equilíbrio. Esse valor é escolhido de maneira a garantir a probabilidade de que a seqüência esteja próxima da probabilidade de equilíbrio em um determinado estágio. Então, se escolhe $N = nk$ de maneira a garantir que \hat{q} esteja próximo de q de acordo

com a probabilidade de cobertura definida. O estimador de q é $\bar{Z}_{k,n} = \frac{1}{n} \sum_{i=1}^n Z_k^{(m+i)}$.

Então,

$$n = \frac{\alpha\beta(2-\alpha-\beta)}{(\alpha+\beta)^3} \left[\frac{\Phi\left(\frac{1}{2}(1+s)\right)}{r} \right]^2,$$

sendo $\Phi(\cdot)$ a normal padrão. Para se utilizar o critério de Raftery e Lewis (1992a), basta especificar o quantil q a ser calculado (ex.: 0,025), a precisão desejada r (ex.: $\pm 0,005$), a probabilidade exigida s de atingir a precisão especificada e uma tolerância em sua estimação. Para se ter uma idéia das ordens de grandezas envolvidas para o caso mais simples de $k = 1$, e desejando-se estimar o quantil 0,025 com erro máximo de 0,0125 a um nível de 95% de confiança, obtém-se $N = n = 600$ (Gamerman, 1996).

A presença de inúmeras variáveis possibilita resultados dessa análise para cada uma delas; assim, devem-se tomar os maiores valores de B , N e k , pois a convergência só se dará quando todas as variáveis convergirem.

Outra saída importante é o Fator de Dependência que é responsável pelo acréscimo multiplicativo ao número de iterações necessárias para se alcançar a convergência, devido à autocorrelação dentro de seqüência. Para situações em que o fator for maior que 5,0, pode-se dizer que a convergência ainda não foi obtida, necessitando reparametrizar o modelo em estudo (Raftery e Lewis, 1992b).

2.3. Análise espectral

MCMC são especiais casos de séries temporais estacionárias, então é natural usar métodos de séries temporais para testar a convergência. Dois métodos existentes na literatura utilizam técnicas padrão de análise espectral para estimar a verdadeira variância via densidade espectral, $S(w)$.

2.3.1. Critério de Geweke (1992)

A ausência de independência entre os dados amostrais é considerada um problema. Na literatura tem-se tratado este problema de modo informal, denotando-o a

um problema computacional. Usando técnicas padrão de análise espectral, promove-se um erro padrão para a aproximação $E[g(\theta)]$ das médias amostrais correspondentes; desta maneira, leva-se em conta qualquer autocorrelação em que $g(\theta)$ é uma função real. Com base nestes resultados, um diagnóstico para a ausência de convergência do Amostrador de Gibbs é construído (Geweke, 1992).

Considerando a função real $g(\theta)$ e sua trajetória $g^{(1)}, g^{(2)}, \dots$, construída a partir de $g^{(i)} = g(\theta^{(i)})$, $i = 1, 2, \dots, n$, essa trajetória define uma série temporal. Pode-se, portanto, estimar a média desta série utilizando o seguinte estimador:

$$\bar{g}_i = \left(\frac{1}{n} \right) \sum_{i=1}^n g^{(i)}.$$

Ela fornece um estimador não-viesado de $E[g(\theta)]$ cuja variância assintótica é dada por $S_g(0)/n$, sendo $S_g(w)$ a densidade espectral da série g de frequência 0. Em geral, S_g é desconhecido sendo estimado por \hat{S}_g com base em análise espectral. Após um número suficiente grande, n , de iterações, determina-se as médias \bar{g}_A e \bar{g}_B , sendo \bar{g}_A a média das primeiras n_A e \bar{g}_B , a média das n_B últimas iterações. Considerando, também, os estimadores independentes $\hat{S}_g^A(0)$ e $\hat{S}_g^B(0)$ das variâncias assintóticas de $\{g^{(i)} : i = 1, \dots, n_A\}$ e $\{g^{(i)} : i = n^*, \dots, n_B\}$, respectivamente, sendo $n^* = n - n_B + 1$, pode-se mostrar que se n_A/n e n_B/n são fixos, com $(n_A + n_B)/n < 1$, quando $n \rightarrow \infty$, tem-se a diferença padronizada entre as médias,

$$\frac{\bar{g}_A - \bar{g}_B}{\sqrt{(\hat{S}_g^A(0)/n_A) + (\hat{S}_g^B(0)/n_B)}} \sim N(0,1), \quad (2.8)$$

se a seqüência $\{g^{(i)}\}$ for estacionária. Assim, um teste para uma seqüência pode ser construído. Se a diferença padronizada entre as médias for grande, existe indicação de ausência de convergência. As médias devem ser construídas após algumas iterações iniciais terem sido descartadas, e que sejam usados os valores de $n_A = 0,1n$ e $n_B = 0,5n$ (Geweke, 1992).

Na prática, deve-se notar que esta tentativa de diagnóstico para verificar uma necessária, mas não suficiente condição de convergência, informa ao pesquisador somente se a convergência não foi alcançada, e não se realmente convergiu. Para outras discussões do problema associando análise espectral, veja Ripley (1987).

2.2.2. Critério de Heidelberger e Welch (1983)

O método é indicado para determinar o valor inicial de uma seqüência discreta de eventos, que é também apropriada para ser utilizada para o diagnóstico de convergência do Amostrador de Gibbs (Heidelberger e Welch, 1983). Inicialmente assume-se que no estado estável tem-se uma covariância ou um processo estacionário fraco. É razoável assumir isto desde que se esteja interessado em seqüências provindas de cadeias de Markov, as quais, conseqüentemente, terão estacionariedade completa. Para testar a hipótese nula de estacionariedade, a partir de uma seqüência de dados $\{g^{(i)} : i = 1, \dots, n\}$ oriundos de um processo não estacionário de desconhecida densidade espectral $S(w)$, utiliza-se a teoria Browniana para fazer esta ligação (ver maiores detalhes em Brooks e Roberts, 1998), e assim avaliar possível superdispersão na presença dos valores iniciais. Pode-se, portanto, utilizar vários testes estatísticos para testar a hipótese de nulidade, como Cramer-von Mises (Von Mises, 1931), a estatística de Kolmogorov-Smirnov (Kolmogorov, 1933) ou a estatística de Schruben's (Schruben *et al.*, 1983), que testam se os valores amostrados para cada variável provêm de uma distribuição estacionária. Se a hipótese nula é rejeitada para um dado valor, o teste é repetido depois de descartados os primeiros 10% das iterações. Se a hipótese é novamente rejeitada, outros 10% são descartados após o descarte dos 10% primeiros. Este processo é repetido até se ter uma proporção de 50% (10%, 20%, ..., 50%). Se a hipótese for novamente rejeitada ou o teste não conseguir ser realizado, isto indica falha da estacionariedade, implicando que é necessário um número maior de iterações. Se o teste for satisfatório, o número inicial de iterações descartadas é indicado como o tamanho do *burn-in*.

O teste de *half-width* é utilizado para verificar se a média estimada está sendo

calculada com uma acurácia pré-especificada, sendo utilizada na porção da seqüência que passou no teste de estacionariedade para cada variável em estudo. Calcula-se um intervalo de confiança com 95% para a média, utilizando o desvio padrão assintótico via métodos de séries temporais, ou seja, utilizando densidade espectral. Pelo teste pode-se verificar que a razão entre o $(\text{limite superior} - \text{limite inferior})/2$ e a estimativa da média é menor que a fração específica pré-definida (acurácia), a qual deverá ser de 0,1 (Heidelberger e Welch, 1983). O resultado passará no teste se isto for verdadeiro, indicando que a média amostral é estimada com um erro relativo aceitável, portanto julgada ser a média *a posteriori*, sendo a amostra realizada da distribuição de interesse. Se o teste falhou, é necessário um número maior de iterações. Para evitar problemas de dados seqüenciais, não se deve repetir este teste muitas vezes. Uma maneira de amenizar o problema é acrescentar à nova seqüência no mínimo 50% referente a seu tamanho, ou seja, a seqüência de dados deverá ser multiplicada por um fator $I > 1,5$ (Heidelberger e Welch, 1983).

Heidelberger e Welch (1983) sugerem este processo iterativo como estimativa do tamanho do *burn-in*.

O diagnóstico tem muito pouco poder para detectar a falta de convergência se a amostra é de tamanho muito pequeno, podendo ser problema também em obter estimativas fidedignas da densidade espectral, o que ocorre com o método de Geweke. Na Tabela 1 resumiram-se os métodos de diagnósticos utilizados, de acordo com as suas características.

Tabela 1 - Resumo de alguns critérios de convergência encontrados na literatura

Critério	Densidade Conjunta?	Seqüências múltiplas?	Aplicável para quais amostradores?	Problema específico?	Computação extensiva?	Interpretação	Fácil de usar
Raftery e Lewis (1992)	Não	Não	Qualquer (GS, MH)	Não	Rápida	Fácil	1
Geweke (1992)	Não*	Não*	Qualquer	Não	Rápida	Fácil	1
Heidelberger e Welch (1983)	Não*	Não*	Qualquer	Não	Rápida	Fácil	1
Gelman e Rubin (1992)	Não	Sim	Qualquer	Não	Rápida	Subjetiva	1

*Pode ser generalizado; Facilidade de usar: 1- Código genérico simples está disponível para a implementação; Gibbs Samples (GS) e Metropolis-Hastings (MH). Fonte: Cowles e Carlin (1996) e Brooks e Roberts (1998).

3. Metodologia

Para avaliar os critérios de convergência foram considerados de maneira geral e sem perda de generalidades dois modelos.

Para o modelo 1, considerou-se um modelo de séries temporais com duas intervenções e erros auto-regressivos e de médias móveis, ARMA. Os critérios foram avaliados considerando dados simulados e a série de Índice Nacional de Preços ao Consumidor – INPC, produzida pelo IBGE (IBGE, 2005). Para maiores esclarecimentos sobre séries temporais e modelos de intervenção ver Morettin e Tolo (2004).

Para o modelo 2 foi utilizado a normal trivariada, considerando três situações de correlação entre as variáveis e dados simulados.

O Amostrador de Gibbs foi usado em cada caso para se obter uma amostra da distribuição de cada parâmetro.

O modelo de intervenção com erros ARMA(p, q), é dado por:

$$y_t = X_t^T \beta + N_t, \quad t = 1, \dots, n,$$

sendo $X_t^T = [\xi_{t,1} \quad \xi_{t,2} \quad \dots \quad \xi_{t,w}]$ é uma matriz ($n \times w$) de variáveis binárias, na qual cada elemento é um vetor e w o número de intervenções, e $\beta^T = [\beta_1 \quad \beta_2 \quad \dots \quad \beta_w]$ um vetor de parâmetros da intervenção, $N_t = \frac{\theta(B)}{\phi(B)} a_t$, a_t uma seqüência de variáveis aleatórias $i.i.d. \sim N(0, \tau^{-1})$, sendo τ a precisão e $\tau^{-1} = \sigma^2$ a variância, $\phi(B)$ e $\theta(B)$ são operadores em B descritos por $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$.

Os resíduos no instante t podem ser dados em termos das observações y_1, y_2, \dots, y_n pela equação:

$$a_t = \phi(B)y_t - \sum_{i=1}^q \theta_i B^i a_t - \phi(B)X_t^T \beta.$$

E, portanto, a função de verossimilhança condicionada nas r -primeiras observações dada por:

$$L(\phi, \theta, \beta, \tau | Y) \propto \tau^{\frac{n-r}{2}} \exp\left\{-\frac{\tau}{2} \sum_{i=r+1}^n a_i^2\right\},$$

sendo $\phi^T = (\phi_1, \phi_2, \dots, \phi_p)$, $\theta^T = (\theta_1, \theta_2, \dots, \theta_q)$, $\beta^T = (\beta_1, \beta_2, \dots, \beta_w)$, $Y = \{y_1, y_2, \dots, y_n\}$ e $r = \max\{p, q\}$.

Para a análise considerou-se *a priori* não informativa $P(\phi, \theta, \beta, \tau) \propto 1/\tau$. Este modelo foi caracterizado por Diaz (1988). A análise Bayesiana juntamente com as distribuições *a priori* e as distribuições *a posteriori* condicionais completas para cada parâmetro foi desenvolvida por Milani (2000) e são apresentadas no Anexo.

Valores usados para a simulação foram $p=2$; $q=2$, $\theta_1 = -0,3$; $\theta_2 = 0,5$; $\beta_1 = -30$; $\beta_2 = 20$; $\phi_1 = 0,5$; $\phi_2 = -0,3$ e $\tau = 1$.

O segundo modelo utilizado foi o da normal trivariada. O Amostrador de Gibbs foi usado para gerar amostras de Monte Carlo das três variáveis. Para isso utilizou-se uma propriedade da distribuição normal multivariada de que todos os subconjuntos de \underline{X} têm distribuição também normal multivariada. Tomando-se uma partição

$${}_p \underline{X}_1 = \begin{bmatrix} {}_q \underline{X}_1 \\ \underline{X}_1 \\ (p-q) \end{bmatrix} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix}$$

e suas correspondentes partições no vetor de média e covariância, tem-se:

$${}_p \underline{\mu}_1 = \begin{bmatrix} {}_q \underline{\mu}_1 \\ \underline{\mu}_1 \\ (p-q) \end{bmatrix} = \begin{bmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{bmatrix} \text{ e } \Sigma = \begin{bmatrix} {}_q \Sigma_{11q} & {}_q \Sigma_{12(p-q)} \\ (p-q) \Sigma_{21q} & (p-q) \Sigma_{22(p-q)} \end{bmatrix},$$

sendo $\underline{X}_1 \sim N_p(\underline{\mu}_1, \Sigma_{11})$, $\underline{X}_2 \sim N_p(\underline{\mu}_2, \Sigma_{22})$ para $q < p$. ${}_q \Sigma_{11q}$ e $(p-q) \Sigma_{22(p-q)}$ são matrizes de covariâncias de \underline{X}_1 e \underline{X}_2 , respectivamente.

Com esta mesma partição a distribuição condicional de $\underline{X}_1 | \underline{X}_2$ pode ser dada por:

$$\underline{X}_1 | \underline{X}_2 \sim N_q(\underline{\mu}_c, \Sigma_c), \quad (2.9)$$

sendo o vetor de média $\underline{\mu}_c = \underline{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underline{x}_2 - \underline{\mu}_2)$ e a matriz de covariâncias dada por

$\Sigma_c = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, retratados em Bock (1975).

O modelo normal trivariado foi simulado considerando três situações de correlação entre as variáveis e estas foram simuladas com variâncias distintas ($S = 1$, $M = 10$ e $A = 100$) e as matrizes de correlações, ρ_i , e covariâncias, Σ_i , respectivamente, são:

$$\rho_S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ e } \Sigma_S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 100 \end{bmatrix},$$

$$\rho_M = \begin{bmatrix} 1 & 0,50 & 0,52 \\ 0,50 & 1 & 0,48 \\ 0,52 & 0,48 & 1 \end{bmatrix} \text{ e } \Sigma_M = \begin{bmatrix} 1 & 1,58 & 5,20 \\ 1,58 & 10 & 15,18 \\ 5,20 & 15,18 & 100 \end{bmatrix},$$

$$\rho_A = \begin{bmatrix} 1 & 0,98 & 0,98 \\ 0,98 & 1 & 0,99 \\ 0,97 & 0,99 & 1 \end{bmatrix} \text{ e } \Sigma_A = \begin{bmatrix} 1 & 3,10 & 9,70 \\ 3,10 & 10 & 31,31 \\ 9,70 & 31,31 & 100 \end{bmatrix},$$

sendo $\Sigma_i = V_i^{1/2}\rho_iV_i^{1/2}$, em que $V_i = \text{Diagonal}(\Sigma_i)$ e $i = S, M$ e A . Esta diferenciação permite-nos simular situações de presença de alta correlação entre as condicionais, média e ausência completa de correlação, coincidindo com o caso extremo, onde se tem a própria marginal.

Os critérios de Gelman e Rubin e Geweke foram calculados com um *burn-in* de 50%. Os critérios foram calculados a primeira vez com 20 iterações e depois recalculados a cada 2 até 30 000 iterações para o modelo de séries temporais e 5 000 iterações para o modelo normal trivariado. Devido o critério de Geweke ter sido calculado múltiplas vezes o nível de significância global não mais será mantido e por isso propõe-se diminuir o nível de significância para 1%, na tentativa de realizar uma proteção. Em seguida estes foram apresentados de forma gráfica para melhor visualização da convergência. O critério de Heidelberger e Welch foi calculado a cada 100 até as 30 000 iterações para o modelo de séries temporais e 5 000 para o modelo normal trivariado. O critério de Raftery e Lewis estima o tamanho ideal da amostra, a

partir de uma amostra-piloto. Utilizou-se os critérios de Geweke e Heidelberger e Welch na amostra considerada ideal pelo critério de Raftery e Lewis, nas situações estudadas e assim confirmar se a convergência ocorreu.

4. Resultados e discussão

Inicialmente serão apresentados os resultados para os modelos 1 e 2, considerando dados simulados e finalmente a série de Índice Nacional de Preços ao Consumidor – INPC.

Na Tabela 2 estão dispostas as análises descritivas para todos os parâmetros do modelo 1, considerando dados simulados e diferentes pontos de convergência. Pode-se perceber que à medida que aumentamos o número de iterações a qualidade da inferência melhora. Quando se utilizou um $R=1,05$ obteve-se uma inferência parecida com 30 000 iterações.

Tabela 2 - Análise descritiva para os sete parâmetros do modelo 1 em três diferentes níveis de convergência para o PRFS

Iterações	Par.	Real	Média	Erro Padrão da Média ("Naive")	Erro de Monte Carlo	HPD*(95%)		C.I.**
						L.I.	L.S.	
380 iter. ($R = 1,2$)	θ_1	-0,3	-0,16599	0,009623	0,01727	-0,4279	0,0574	0,485
	θ_2	0,5	0,56533	0,008155	0,01261	0,3680	0,7905	0,423
	β_1	-30	-29,8548	0,003734	0,00438	-29,945	-29,743	0,202
	β_2	20	19,74254	0,003536	0,00401	19,660	19,839	0,179
	ϕ_1	0,5	0,33354	0,009216	0,01625	0,1022	0,5566	0,454
	ϕ_2	-0,3	-0,25443	0,007474	0,01197	-0,4364	-0,0435	0,393
	τ	1	0,97746	0,004805	0,00577	0,8731	1,1215	0,248
2 156 iter. ($R = 1,05$)	θ_1	-0,3	-0,12314	0,004590	0,01210	-0,4211	0,1569	0,578
	θ_2	0,5	0,52020	0,003886	0,00941	0,2626	0,7612	0,499
	β_1	-30	-29,8585	0,001446	0,00166	-29,953	-29,768	0,185
	β_2	20	19,74804	0,001445	0,00166	19,662	19,842	0,180
	ϕ_1	0,5	0,28859	0,004447	0,01173	-0,0107	0,5541	0,565
	ϕ_2	-0,3	-0,20407	0,003937	0,00994	-0,4247	0,0840	0,509
	τ	1	0,97436	0,001905	0,00169	0,8561	1,1011	0,245
30 000 iter.	θ_1	-0,3	-0,10441	0,000836	0,00300	-0,3852	0,1812	0,566
	θ_2	0,5	0,50796	0,000750	0,00252	0,2445	0,7469	0,502
	β_1	-30	-29,8601	0,000283	0,00035	-29,956	-29,763	0,193
	β_2	20	19,74731	0,000282	0,00034	19,654	19,845	0,191
	ϕ_1	0,5	0,27146	0,000802	0,00288	-0,0105	0,5349	0,545
	ϕ_2	-0,3	-0,18854	0,000757	0,00270	-0,4337	0,0770	0,511
	τ	1	0,97429	0,000360	0,00039	0,8543	1,0989	0,245

*Highest Probability Density Intervals; ** Comprimento do intervalo.

O comprimento do intervalo de maior densidade (HPD) apresentou uma maior variação quando comparado o resultado obtido para um $R = 1,2$ com o de 30 000 iterações, mas não é muito diferente o de $R = 1,05$ do de 30 000 iterações. Pode-se realmente afirmar que as marginais da primeira comparação são diferentes e que as obtidas com um $R = 1,05$ já se encontram com uma forma mais parecida com a de 30 000. Não se necessita de muitas iterações para se ter a estabilização das Cadeias de Markov, mas para o processo de Monte Carlo necessita-se de mais iterações, pois o Erro de Monte Carlo deverá ser aceitável. As estimativas obtidas para as médias estão próximas aos valores reais, entretanto o importante a ser verificado é o formato da distribuição.

Para o Modelo 2 da normal trivariada na condição de nenhuma correlação entre as variáveis, tem-se o caso extremo de se obterem as próprias marginais e, assim, o processo de amostragem do Amostrador de Gibbs se comportar como um processo de Monte Carlo. O processo de convergência é independente, ou seja, pode ou não ter boas estimativas, pois a convergência pode ter ocorrido, mas se a amostra for pequena, o erro de Monte Carlo será grande. O tamanho da amostra é importante para reduzir este erro e, assim, obterem-se estimativas mais precisas (Tabela 3).

Tabela 3 - Análise descritiva para os três parâmetros do modelo 2 em dois diferentes estágios de convergência, considerando ausência de correlação

Iterações	Variáveis	\bar{X}	$S_{\bar{X}}^2$	Quantil 2,5	Quantil 97,5	S^2	Erro de Monte Carlo
90* iter. (R = 1,05)	X1	2,40	0,178	0,00	4,87	1,43	0,087
	X2	199,69	0,465	191,12	204,00	9,72	0,319
	X3	2,18	1,156	-13,62	14,00	60,19	0,858
5 000** iter.	X1	1,98	0,021	0,00	4,00	1,13	0,013
	X2	200,03	0,063	194,00	206,00	9,86	0,042
	X3	-0,17	0,199	-20,00	20,00	98,67	0,128

* Contempla uma amostra da iteração 46 a 90. ** Amostra de 2 501 a 5 000. ¹ Estimador viesado. As médias reais foram 2,0; 200,0 e 0,2 e as variâncias reais foram 1; 10 e 100, respectivamente.

A amostra converge com 90 iterações, mas as estimativas das médias só estabilizaram com aproximadamente 2 000 iterações para as variáveis 1 e 2, e não estabiliza para a variável 3, necessitando de um maior número de iterações (maior amostragem). É evidente que quando diagnosticada a convergência utilizando o critério

de Gelman e Rubin, não necessariamente ter-se-á maior precisão. É uma condição necessária, mas não suficiente. Observa-se, também, que a estimativa de S^2 é viesada quando a amostra não apresentou um tamanho mínimo de iterações suficiente, só melhorando com o aumento do número de iterações, mas o viés foi maior para a variável de maior variabilidade (variável 3).

Considerando correlação média entre as variáveis, pôde-se perceber que o critério de Gelman e Rubin convergiu com um número maior de iterações, isto também foi verificado com a correlação alta, verificado também que a convergência se apresentou de maneira lenta, pois o processo de amostragem percorre mais lentamente o espaço paramétrico.

Apesar de a caracterização da convergência para o caso de alta correlação ter sido verificada, observa-se na Tabela 4 que a média estimada da variável 3 apresentou estimativa muito distante da verdadeira média ($\mu_3 = 0,2$), o que pode ser devido à alta variabilidade que esta variável apresenta. Mesmo com um número alto de iterações pode-se verificar que as estimativas das variâncias são viesadas.

O emprego deste critério deve ser feito com cuidado, exigindo um conhecimento sobre a variabilidade do parâmetro de interesse.

Gelman e Rubin (1992) é um critério que se preocupa em caracterizar a convergência, mas que sozinho não garante a qualidade da inferência, pois é fundamental que a amostra seja representativa. Por isso um número maior de iterações é necessário, além daquela quantidade atingida quando o fator $R = 1,2$ já for alcançado.

Tabela 4 - Análise descritiva para os três parâmetros do modelo 2 em dois diferentes estágios de convergência, considerando correlação alta

Iterações	Variáveis	\bar{X}	$S_{\bar{X}}^2$	Quantil 2,5	Quantil 97,5	S^2	Erro de Monte Carlo
1 408 iter.* ($R = 1,2$)	X1	2,529	0,034	0,2809	3,9826	0,825	0,07732
	X2	201,7	0,108	194,88	206,49	8,269	0,24853
	X3	5,607	0,343	-15,721	20,124	83,133	0,78351
3 108 iter.** ($R = 1,05$)	X1	2,193	0,021	0,601	3,758	0,716	0,06336
	X2	200,6	0,067	195,58	205,61	7,052	0,20195
	X3	2,160	0,213	-13,767	18,745	70,791	0,63718
5 000*** iter.	X1	2,353	0,018	0,5115	4,3297	0,8334	0,06027
	X2	201,1	0,058	195,43	207,40	8,2945	0,19236
	X3	3,820	0,182	-14,562	23,896	83,292	0,60509

* Contempla uma amostra da iteração 705 a 1 408.

** Contempla uma amostra da iteração 1 555 a 3 108.

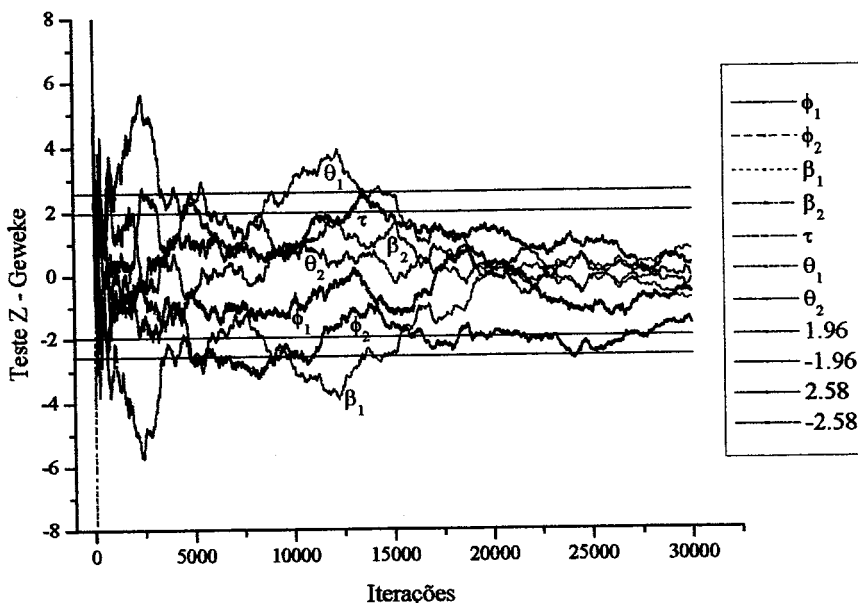
*** Amostra de 2 501 a 5 000. ¹ Estimador viesado.

O critério de Geweke (1992) se mostrou mais rigoroso, caracterizando a convergência com um número grande de iterações se comparado ao critério de Gelman e Rubin (1992). A utilização do nível de significância de 1% mostrou uma boa alternativa para o teste de hipótese realizado pelo critério de Geweke, caracterizando a convergência muito antes do que era verificado quando se utiliza 5%.

Utilizou-se o acompanhamento do critério para verificar seu comportamento e o resultado foi bastante interessante. A estatística do teste deve ser construída concomitantemente com o processo de amostragem, sendo atualizada a cada x iterações. O uso de um número pequeno para x se mostrou desnecessário, aumentando o esforço computacional. A utilização de x em torno de 10 a 20 e a utilização de métodos de suavização seriam os indicados para a construção do gráfico. O critério se comportou de maneira crítica quando aplicado pontualmente, pois não se tem a idéia de como o processo se comportará quando aumentado o número de iterações. Com isso o emprego de gráfico para o critério de Geweke para acompanhar a convergência foi interessante e se mostrou extremamente importante. Este método não deve ser aplicado sozinho, devendo ser combinado com outro método, pois seu resultado favorável à convergência é uma condição necessária, mas não suficiente para garantir a convergência.

Para o caso de se estudar a convergência de maneira pontual, ou seja, fazendo, por exemplo, 5 000 iterações e depois se aplicando o teste, verifica-se que ele fornecerá um resultado que caracteriza a convergência. No entanto, caso as iterações sejam aumentadas para 12 500 e aplique-se novamente o teste, ter-se-á um resultado que caracteriza a ausência de convergência. Estudando o modelo de séries temporais, verificou-se o problema e que pode ilustrar este caso, Figura 1; para este exemplo o critério caracterizou convergência e ao aumentar as iterações, esta não mais foi verificada. Outro ponto que merece destaque é que este método caracteriza a convergência somente quando ocorre a estabilização da média, sendo, portanto, diferente do critério de Gelman e Rubin.

Figura1 - Representação gráfica das estatísticas do teste Z do critério de Geweke para os sete parâmetros até 30 000 iterações



Ao se analisar mais detalhadamente a Figura 1, observa-se que as linhas que caracterizam os limites do intervalo de confiança, para a diferença de duas médias foram ultrapassadas várias vezes no decorrer das iterações. As quatro linhas retas caracterizam os limites para aceitação da hipótese de confirmação da convergência. Para afirmar que a convergência ocorreu, as seqüências dos sete parâmetros teriam que estar dentro destes limites. Verifica-se, pelo gráfico, que esta só ocorreu próximo das 27 500 iterações. Se for utilizado um rigor maior para rejeição (1%), haverá as duas linhas externas, caracterizando a convergência muito antes, próximo das 15 000 iterações.

O critério de Geweke também se mostrou rigoroso para o modelo normal trivariado, mesmo para a situação de ausência de correlação.

Para as situações de ausência de correlação e correlação média, observa-se que as variáveis apresentaram comportamento similar, apresentando convergência para 5% próxima das 4 600 iterações. Para a situação de correlação alta pode-se verificar a tendência do critério em permanecer dentro dos limites do intervalo de confiança, ou

seja, convergir com o aumento do número de iterações, mas que serão necessárias mais de 5 000 iterações.

O método de Heidelberger e Welch (1983) apresentou a mesma característica que o método de Geweke, devendo também ser evitado o seu uso de maneira pontual. Este fato pode, também, ocorrer com o critério de Gelman e Rubin.

À medida que o número de iterações aumentou, este critério não mais rejeitou nenhuma das três hipóteses (Estacionaridade, Precisão da média e Tamanho do *burn-in*). Segundo este critério um número grande de iterações implica que não é necessário retirar o *burn-in*.

Para o modelo da normal trivariada na ausência de correlação, o teste de estacionariedade não rejeitou a hipótese para as três variáveis em todas as iterações testadas. O teste de *half-width* não rejeitou a acurácia para as variáveis 1 e 2. Para a variável 3, este rejeitou-a em todo o processo iterativo. Cabe aqui mencionar que a variável 3 foi simulada com maior variância, ou seja, 100 vezes maior que a variável 1 e 10 vezes maior que a variável 2. O teste não foi sensível a esta diferença de variabilidade, rejeitando a hipótese de que a média foi estimada com acurácia pré-especificada. Talvez essa hipótese deixe de ser rejeitada para esta situação de alta variabilidade somente com amostras de tamanho muito grande, pois a acurácia é pré-especificada para todos os parâmetros (variáveis) simultaneamente, quando se utilizam softwares como BOA e CODA (ambos os pacotes do software R). O teste do *burn-in* apresentou total contradição, indicando necessidade de *burn-in* em certos momentos e em outros não, mesmo com poucas iterações.

O critério de Raftery e Lewis (1992a), para a determinação do tamanho adequado da amostra, se mostrou extremamente importante, fornecendo bons resultados para o modelo de séries temporais.

Para o modelo normal trivariado, o critério de Raftery e Lewis apresentou bons resultados na situação de ausência de correlação e de correlação média, sendo os resultados coerentes. Chamaram a atenção os resultados desse critério quando as condicionais completas do modelo normal trivariado, apresentavam correlação alta. Para este caso específico o critério de Raftery e Lewis apresentou resultados de convergência

inconsistentes com a qualidade da amostra obtida, devendo-se tomar cuidado com o tamanho da amostra nesta situação, não dispensando o conhecimento do pesquisador sobre o parâmetro de interesse.

Para se obter uma maior eficiência no monitoramento da convergência utilizando os critérios estudados nesse trabalho, alguns procedimentos são recomendados. A princípio, é necessário gerar uma amostra-piloto e aplicar o critério de Raftery e Lewis. Esse critério fornecerá um tamanho de seqüência a ser obtido. O Amostrador de Gibbs é implementado e na proximidade do fim do tamanho da seqüência recomendada pelo critério de Raftery e Lewis são aplicados os critérios de Gelman e Rubin e Geweke. O objetivo desse passo é monitorar a seqüência a cada x iterações para verificar se a convergência foi alcançada. Caso esse objetivo tenha sido verificado, o processo é interrompido. Por outro lado, se a convergência não foi alcançada, o Amostrador de Gibbs deve ser acionado para se obter novos x componentes da seqüência e novamente o processo é repetido até que seja atingida a convergência. Espera-se que a convergência seja atingida próxima ao tamanho de seqüência apontado pelo critério de Raftery e Lewis. Finalmente, pode-se obter o tamanho do *burn-in* utilizando-se o critério de Heidelberger e Welch.

É importante enfatizar que o procedimento de verificação da convergência dependerá muito da opinião do pesquisador, por isso nunca se deve realizar um único procedimento de amostragem. A qualidade da amostra é fundamental para se realizar inferência, por isso, o erro de Monte Carlo deve ser averiguado.

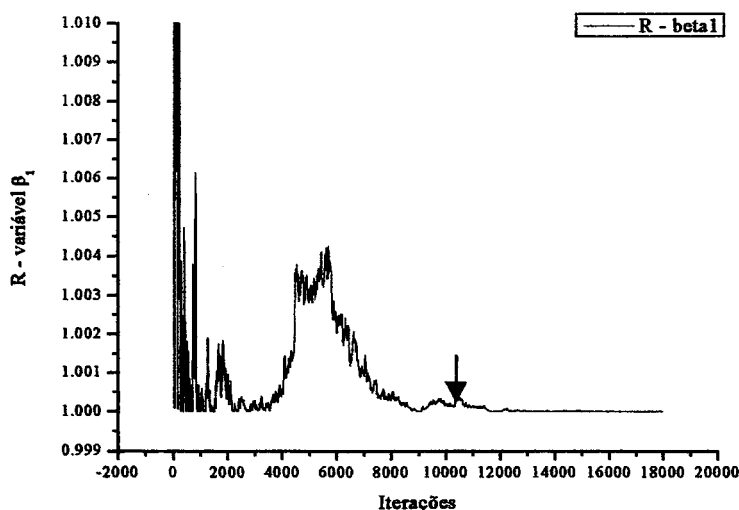
Série de índice nacional de preços ao consumidor – INPC

Para exemplificar a aplicação dos critérios com dados reais, utilizou-se o trabalho de Sáfadi e Milani (2004), em que os autores fazem o estudo da série de Índice Nacional de Preços ao Consumidor – INPC, produzida pelo IBGE (IBGE, 2005) e obtida no *site* www.ipea.gov.br, no período de Abril de 1979 a Outubro de 2002. Durante o período em estudo foi considerado três pontos de intervenções, sendo os fatores possíveis: Plano Cruzado (Março de 1983 – obs. 91); Plano Collor I (Março de 1990 – obs. 131) e Plano Real (Julho de 1994 – obs. 184). O objetivo deste exemplo é estudar a

convergência do Amostrador de Gibbs com a utilização de duas seqüências, iniciando a partir de valores arbitrários diferentes e a estimação dos parâmetros do modelo. O modelo ajustado pelos autores continha três intervenções ($\beta_1, \beta_2, \beta_3$) e componente residual AR(1).

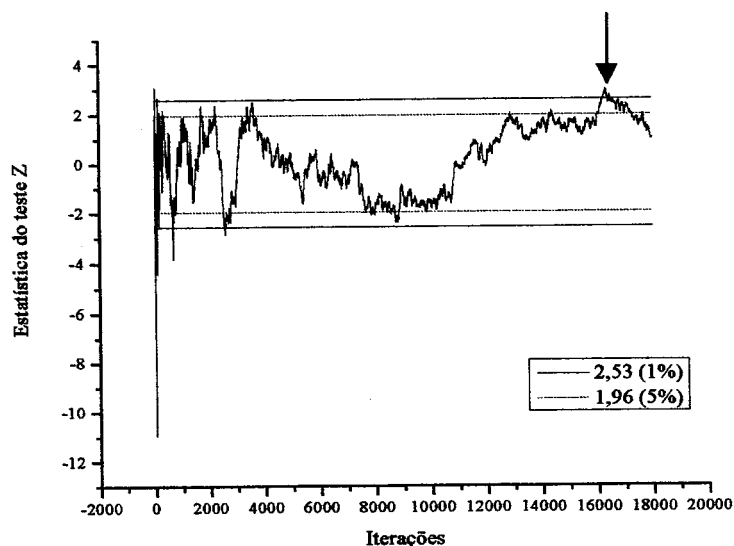
Iniciar-se-á o estudo simulando uma amostra piloto de 4 000 iterações para ser testada com o critério de Raftery e Lewis. O resultado indicou que a variável mais problemática, que demorará mais a convergir é a β_1 . A indicação do teste foi a de realizar um pulo de 3 e um total de 17 478 iterações. Devido a esta saída optou-se por realizar 18 000 iterações. Concomitantemente serão realizados de maneira, também, iterativa os testes de Gelman e Rubin, Geweke e Heidelberger e Welch. Este último teste nos fornecerá o valor do *Burn-in*. O resultado do Critério de Gelman e Rubin mostrou que valores de $R < 1,05$ ocorreram bem próximos do início do processo. A estabilização do critério, próximo de 1, ocorreu por volta das 12 000 iterações para a variável de maior variabilidade β_1 (Figura 2), coincidindo com a apresentada pelo critério de Raftery e Lewis.

Figura 2 - Representação gráfica da estatística do teste de Gelman e Rubin para o parâmetro β_1



Segundo o Critério de Geweke, o parâmetro β_1 apresentou um resultado significativo na interação 16 230 para um nível de significância de 1%. Na Figura 3 é possível visualizar seu comportamento. Foi verificado que ocorreram várias vezes a rejeição da hipótese para um nível de significância de 5%. Os outros parâmetros não apresentaram resultado significativo a partir de 4 000 iterações para 1% de significância.

Figura 3 - Representação gráfica das estatísticas do teste Z do critério de Geweke para o parâmetro β_1



Como o critério de Geweke é um método para rejeitar a hipótese de não convergência, pode-se dizer que esta não ocorreu antes das 16 230 iterações. Para o critério de Heidelberger e Welch foi verificado que os resultados para o *burn-in* caracterizaram sua presença até a interação 6 700 na variável β_2 , a partir desta não se verificou mais a necessidade de *burn-in*. Com isso e devido aos resultados apresentados pelos testes de Gelman e Rubin e Geweke é recomendado adotar um *burn-in* de 50% e um pulo de 3 para a realização da amostragem. Segundo Sáfadi e Milani (2004) foram utilizadas 8 000 iterações para cada seqüência, um *burn-in* de 40% e um pulo de 15. Os resultados para as inferências dos parâmetros obtidos por Sáfadi e Milani (2004) e neste trabalho, se encontram na Tabela 5.

Tabela 5 - Análise descritiva para os cinco parâmetros obtidos por Sáfadi e Milani e neste trabalho

Parâmetros	Média	Desvio Padrão	Média	Desvio Padrão
	Por: Sáfadi e Milani (2004)			
ϕ	0,901	0,041	0,902	0,055
β_1	12,369	5,113	11,965	5,172
β_2	13,619	4,251	13,414	4,485
β_3	-33,425	5,030	-33,481	4,966
τ	0,0404	0,035	0,0403	0,003

A Tabela 5 apresenta os resultados obtidos por Sáfadi e Milani (2004) e o deste trabalho. Pode-se perceber que os valores não diferiram muito, maiores diferenças foram encontradas para o parâmetro β_1 que também é o parâmetro de maior variabilidade e para o parâmetro da precisão τ , que apresentou um desvio padrão diferente do encontrado por Sáfadi e Milani (2004).

5. Conclusões

Sugere-se a utilização dos critérios univariados do monitoramento da convergência de forma combinada. O seguinte procedimento é recomendado:

- aplicar Raftery e Lewis em uma amostra-piloto e determinar o tamanho ideal da seqüência;
- monitorar a convergência das seqüências nas proximidades do tamanho ideal indicado pelo critério de Raftery e Lewis por meio dos critérios de Gelman e Rubin e Geweke; e
- determinar o tamanho do *burn-in* pelo critério de Heidelberger e Welch.

Referências bibliográficas

- BOCK, R. D., (1975). *Multivariate statistical methods in behavioral research*. New York: McGraw Hill, 623 p.
- BROOKS, S. P.; GELMAN, A., (1998). General methods for monitoring convergence of iterative simulations, *Journal of Computational and Graphical Statistics*, Alexandria, v. 7, n. 4, p. 434-455.
- BROOKS, S. P.; ROBERTS, G. O., (1995). Some L^2 results for diagnosing convergence fo MCMC algorithms. *Tech. Rep.*, University of Cambridge, In Preparation.

- BROOKS, S. P.; ROBERTS, G. O., (1998). Assessment convergence of markov chain monte carlo algorithms. *Statistics and Computing*, Dordrecht, v. 8, n. 4, p. 319-335.
- CARLIN, B. P.; LOUIS, T. A., *Bayes and empirical bayes methods for data analysis*. 2. ed. London: Chapman and Hall, 2000. 419 p.
- COWLES, M. K.; CARLIN, B. P., (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Associations*, Alexandria, v. 91, n. 434, p. 883-904.
- DIAZ, J., (1988). Bayesian analysis of regression models with time series erros. In SPALL, J. C. (Ed.). *Bayesian analysis of Time Series and Dynamic Models*. New York: M. Dekker, p. 347-358.
- FISHER, R. A., (1935). *The desing of experiments*, Edinburgh: Oliver and Boyd.
- GAMERMAN, D., (1996). *Simulação estocástica via cadeias de Markov*. São Paulo: ABE, 196 p.
- GELMAN, A., (1995). Inference and monitoring convergence. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (ED.). *Markov Chain Monte Carlo in Practice*. New York: Chapman and Hall, chap. 8.
- GELMAN, A.; RUBIN, D.B., (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, Hayward, v.7, n.4, p.457-511.
- GEWEKE, J., (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Bayesian Statistics 4* (eds. Bernardo, J.M.; Berger, J.O.; Dawid, A.P.; Smith, A.F.M.), New York: Oxford University Press, p.625-631.
- GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J., (1996). *Markov Chain Monte Carlo in Practice*, London: Chapman and Hall.
- HEIDELBERGER, P.; WELCH, P., (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, v.31, p.1109-44.
- KOLMOGOROV, A. N.; (1933). Sulla determinazione Empirica delle leggi di probabilita. *Giornal Istituto Italia Attuari*, v. 4, p. 1-11.
- LIU, C.; WONG, W. H.; KONG, A., (Apr.1994). Correlation structure and convergence rate of the Gibbs sampler: applications to the comparison of estimators and augmentation schemes. *Biometrika*, London, v. 81, n. 1, p. 27-40.
- MILANI, L. L., (2000). *Análise bayesiana do modelo de intervenção com erro arma*. 142 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.
- RAFTERY, A. E.; LEWIS, S., (1992a). Comment: One long run with diagnostics: implementation strategies for markov chain monte carlo. *Statistical Science*, Hayward, v. 7, n. 4, p. 493-497.
- RAFTERY, A. E.; LEWIS, S., (1992b). How many iterations in the Gibbs sampler? Em *Bayesian Statistics 4* (eds.: J. M. Bernardo et al.), Oxford: University Press, p.763-773.
- RIPLEY, B. D., (1987). *Stochastic Simulation*. New York: John Wiley.
- SCHURUBEN, L.; SINGH H.; TIERNEY, L., (1993). Optimal tests for initialization bias in simulation output. *Operations Research*, Landing, v. 31, n. 6. p. 1167-1178.
- VON MISES, R., (1931). *Wahrscheinlichkeitsrechnung*. Vienna: Deutiche.
- IBGE(2005) Sistema Nacional de Índices de Preços ao Consumidor: Estruturas de Ponderação a partir da Pesquisa de Orçamentos Familiares 2002-2003. *Série Relatórios Metodológicos*, volume 34, Rio de Janeiro,RJ.

ANEXO

No que segue é apresentado a análise Bayesiana para o Modelo 1, determinando as distribuições *a posteriori* condicionais completas para cada parâmetro.

A distribuição *a posteriori* conjunta para os parâmetros é obtida através do Teorema de Bayes:

$$P(\phi, \theta, \beta, \tau | Y) \propto \tau^{\frac{n-r}{2}-1} \exp\left\{-\frac{\tau}{2} \sum_{t=r+1}^n (\mathbf{Z} - \mathbf{U}^T \beta)^T (\mathbf{Z} - \mathbf{U}^T \beta)\right\},$$

sendo n o número de observações, p o número de parâmetros auto-regressivos, β o vetor de parâmetros da intervenção e os componentes do vetor \mathbf{Z} e as linhas da matriz \mathbf{U} dadas, respectivamente, por:

$$Z_t = y_t - \mathbf{y}^T \phi - \hat{\mathbf{a}}^T \theta, \quad \mathbf{U}_t^T = \mathbf{X}_H^T - \phi^T \mathbf{X}_p,$$

em que $\mathbf{y}^T = [y_{t-1} \ y_{t-2} \ \dots \ y_{t-p}]$ é o vetor de dados, $\hat{\mathbf{a}}^T = [\hat{a}_{t-1} \ \hat{a}_{t-2} \ \dots \ \hat{a}_{t-q}]$ o vetor de resíduos estimados, $\mathbf{X}_H^T = [\xi_{t,1} \ \xi_{t,2} \ \dots \ \xi_{t,w}]$ a matriz de incidência das intervenções e \mathbf{X}_p uma matriz dada por:

$$\mathbf{X}_p = \begin{bmatrix} \xi_{t-1,1} & \xi_{t-1,2} & \dots & \xi_{t-1,m} \\ \xi_{t-2,1} & \xi_{t-2,2} & \dots & \xi_{t-2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{t-p,1} & \xi_{t-p,2} & \dots & \xi_{t-p,m} \end{bmatrix}.$$

Distribuição condicional completa para ϕ :

$$\phi | \theta, \beta, \tau, Y \sim \text{Normal}(N^{-1}M, \tau^{-1}N^{-1})$$

sendo $N = N_1 - N_2 - N_3 + N_4$, em que: $N_1(i, j) = \sum_{t=r+1}^n y_{t-i} y_{t-j}$,

$$N_2(i, j) = \sum_{t=r+1}^n \sum_{k=1}^w \beta_k y_{t-i} \xi_{t-j,k}, \quad N_3 = N_2^T \quad \text{e} \quad N_4(i, j) = \sum_{t=r+1}^n \sum_{k=1}^w \sum_{l=1}^w \beta_k \beta_l \xi_{t-i,k} \xi_{t-j,l}.$$

Para $M = M_1 - M_2 - M_3 - M_4 + M_5 + M_6$, tem-se:

$$M_1 = \begin{bmatrix} \sum_t y_{t-1} \xi_{t-1} \\ \sum_t y_{t-2} \xi_{t-2} \\ \vdots \\ \sum_t y_{t-p} \xi_{t-p} \end{bmatrix}, \quad M_2 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^q \theta_i y_{t-1} \hat{a}_{t-i} \\ \sum_{t=r+1}^n \sum_{i=1}^q \theta_i y_{t-2} \hat{a}_{t-i} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^q \theta_i y_{t-p} \hat{a}_{t-i} \end{bmatrix}$$

$$M_3 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^w \beta_i y_t \xi_{t-1,i} \\ \sum_{t=r+1}^n \sum_{i=1}^w \beta_i y_t \xi_{t-2,i} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^w \beta_i y_t \xi_{t-p,i} \end{bmatrix}, \quad M_4 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^w \beta_i y_{t-1} \xi_{t,i} \\ \sum_{t=r+1}^n \sum_{i=1}^w \beta_i y_{t-2} \xi_{t,i} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^w \beta_i y_{t-p} \xi_{t,i} \end{bmatrix}$$

$$M_5 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^q \beta_i \theta_j \xi_{t-1,i} \hat{a}_{t-j} \\ \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^q \beta_i \theta_j \xi_{t-2,i} \hat{a}_{t-j} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^q \beta_i \theta_j \xi_{t-p,i} \hat{a}_{t-j} \end{bmatrix} \quad \text{e} \quad M_6 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^w \beta_i \beta_j \xi_{t,i} \xi_{t-1,j} \\ \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^w \beta_i \beta_j \xi_{t,i} \xi_{t-2,j} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^w \beta_i \beta_j \xi_{t,i} \xi_{t-p,j} \end{bmatrix}.$$

Distribuição condicional completa para ϕ :

$$\theta | \phi, \beta, \tau, Y \sim \text{Normal}(C^{-1}P, \tau^{-1}C^{-1}).$$

sendo $C(i, j) = \sum_{t=r+1}^n \hat{a}_{t-i} \hat{a}_{t-j}$ e $P = P_1 - P_2 - P_3 + P_4$, em que:

$$P_1 = \begin{bmatrix} \sum_{t=r+1}^n y_t \hat{a}_{t-1} \\ \sum_{t=r+1}^n y_t \hat{a}_{t-2} \\ \vdots \\ \sum_{t=r+1}^n y_t \hat{a}_{t-q} \end{bmatrix}, \quad P_2 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^p \phi_i \hat{a}_{t-1} y_{t-i} \\ \sum_{t=r+1}^n \sum_{i=1}^p \phi_i \hat{a}_{t-2} y_{t-i} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^p \phi_i \hat{a}_{t-p} y_{t-i} \end{bmatrix},$$

$$P_3 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^w \beta_i \hat{a}_{t-1} \xi_{t,i} \\ \sum_{t=r+1}^n \sum_{i=1}^w \beta_i \hat{a}_{t-2} \xi_{t,i} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^w \beta_i \hat{a}_{t-p} \xi_{t,i} \end{bmatrix} \quad \text{e} \quad P_4 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^p \beta_i \phi_j \hat{a}_{t-1} \xi_{t-j,i} \\ \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^p \beta_i \phi_j \hat{a}_{t-2} \xi_{t-j,i} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^w \sum_{j=1}^p \beta_i \phi_j \hat{a}_{t-q} \xi_{t-j,i} \end{bmatrix}$$

Distribuição condicional completa para β :

$$\beta | \phi, \theta, \tau, Y \sim \text{Normal}(B^{-1}A, \tau^{-1}B^{-1}),$$

sendo $A = A_1 - A_2 - A_3 + A_4 - A_5 + A_6$, em que:

$$A_1 = \begin{bmatrix} \sum_{t=r+1}^n y_t \xi_{t,1} \\ \sum_{t=r+1}^n y_t \xi_{t,2} \\ \vdots \\ \sum_{t=r+1}^n y_t \xi_{t,w} \end{bmatrix}, \quad A_2 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^p \phi_i y_t \xi_{t-i,1} \\ \sum_{t=r+1}^n \sum_{i=1}^p \phi_i y_t \xi_{t-i,2} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^p \phi_i y_t \xi_{t-i,w} \end{bmatrix},$$

$$A_3 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^p \phi_i y_{t-i} \xi_{t,1} \\ \sum_{t=r+1}^n \sum_{i=1}^p \phi_i y_{t-i} \xi_{t,2} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^p \phi_i y_{t-i} \xi_{t,w} \end{bmatrix}, \quad A_4 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^p \sum_{j=1}^p \phi_i \phi_j y_{t-i} \xi_{t-j,1} \\ \sum_{t=r+1}^n \sum_{i=1}^p \sum_{j=1}^p \phi_i \phi_j y_{t-i} \xi_{t-j,2} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^p \sum_{j=1}^p \phi_i \phi_j y_{t-i} \xi_{t-j,w} \end{bmatrix},$$

$$A_5 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^q \theta_i \hat{a}_{t-i} \xi_{t,1} \\ \sum_{t=r+1}^n \sum_{i=1}^q \theta_i \hat{a}_{t-i} \xi_{t,2} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^q \theta_i \hat{a}_{t-i} \xi_{t,w} \end{bmatrix} \quad \text{e} \quad A_6 = \begin{bmatrix} \sum_{t=r+1}^n \sum_{i=1}^q \sum_{j=1}^p \theta_i \phi_j \hat{a}_{t-i} \xi_{t-j,1} \\ \sum_{t=r+1}^n \sum_{i=1}^q \sum_{j=1}^p \theta_i \phi_j \hat{a}_{t-i} \xi_{t-j,2} \\ \vdots \\ \sum_{t=r+1}^n \sum_{i=1}^q \sum_{j=1}^p \theta_i \phi_j \hat{a}_{t-i} \xi_{t-j,w} \end{bmatrix}.$$

Para $B = B_1 - B_2 - B_3 + B_4$ tem-se $B_1(i, j) = \sum_{t=r+1}^n \xi_{t,i} \xi_{t,j}$, $B_2(i, j) = \sum_{t=r+1}^n \sum_{k=1}^w \phi_k \xi_{t,i} \xi_{t-k,j}$.

$$B_3 = B_2^T \quad \text{e} \quad B_4(i, j) = \sum_{t=r+1}^n \sum_{k=1}^p \sum_{l=1}^p \phi_k \phi_l \xi_{t-k,i} \xi_{t-l,j}.$$

Distribuição condicional completa para τ :

$$\tau | \phi, \theta, \beta, Y \sim \text{Gama} \left(\frac{n-p}{2}, \frac{\sum_{t=r+1}^n (Z_t - U_t^T \beta)' (Z_t - U_t^T \beta)}{2} \right).$$

Abstract

Markov Chain Monte Carlo Methods have been studied in several areas, but one of the largest difficulties is to determine the appropriate sample size, i.e., to determine the convergence of the process and then make inference over parameters of the target distribution. In spite of several accessible convergence diagnostic procedures in literature we have opted to study the most used and simplest to implement diagnostic procedures. The objectives of this work were: to propose an alternative form of combined application of 4 univariate Markov chain convergence diagnostics (Gelman e Rubin, 1992; Raftery e Lewis, 1992a; Geweke, 1992; Heidelberger e Welch, 1983) and optimize their properties; Simulated data from two models was used. The first a time series with two intervention parameters with ARMA (2,2) error and the second a normal trivariate, utilizing three different structures for the covariance matrix. In both cases the Gibbs sampler was used. The following conclusions were obtained: it was suggest that univariate diagnostic procedures should be applied in a combined way.

Keywords: Correlation; Gibbs sampler, time series.

Uso de modelos de espaços de estados para a estimação do efeito de vício de grupos de rotação na PME/IBGE

*Marcelo Trindade Pitta**
*Denise Britz do Nascimento Silva***

Resumo

Várias das séries temporais utilizadas por diversos setores da sociedade são provenientes de pesquisas repetidas no tempo. A Pesquisa Mensal de Emprego do IBGE é uma pesquisa mensal com painéis rotativos na qual encontra-se evidência da ocorrência de um efeito denominado vício de grupos de rotação devido ao esquema de rotação adotado no plano amostral. Este artigo apresenta uma metodologia que permite a estimação do efeito de vício de grupos de rotação via modelos de espaço de estados, considerando o erro amostral e o desenho amostral da pesquisa. Os resultados obtidos caracterizam-se pela suavização das séries das taxas de variação anual (diferença entre as taxas de um mesmo mês em anos consecutivos), quando comparadas com as variações calculadas a partir da série divulgada pelo instituto.

Palavras-chave: Pesquisa Mensal de Emprego, pesquisas repetidas, vício de grupos de rotação, modelos de espaço de estados.

* Fundação SEADE.

** Endereço para correspondência: IBGE - Escola Nacional de Ciências Estatísticas - ENCE/IBGE - Denise Britz do Nascimento Silva - Rua André Cavalcanti, 106 - Cep. 20231-050 - Rio de Janeiro - RJ.

1. Introdução

Este artigo apresenta um modelo de séries temporais para a estimação de um efeito de tendência encontrado nas estimativas de algumas pesquisas repetidas, realizadas por meio de amostragem probabilística. Tal efeito, denominado efeito de vício de grupos de rotação, está associado ao processo de investigação que é adotado e é observado nas pesquisas repetidas que são realizadas, utilizando painéis rotativos (Bailar, 1975; Kumar e Lee, 1983, Pfeffermann 1991). As possíveis causas para a ocorrência de tal efeito estão relacionadas, principalmente, à mudança na forma de coleta de informações de um instante para outro da pesquisa e a erros não-amostrais. Os erros não-amostrais mais comuns são o erro de classificação (por exemplo, considerar como ocupada uma pessoa desocupada, no caso de uma pesquisa sobre força de trabalho) e a ocorrência de não-resposta (falta de informação).

Além do efeito de vício de grupos de rotação, as pesquisas repetidas têm uma outra componente que deve ser considerada na análise de suas séries temporais: o erro amostral. Como, em pesquisas repetidas, uma mesma unidade informante é observada em instantes distintos ao longo do tempo, tal procedimento induz uma autocorrelação entre os erros amostrais e, conseqüentemente, entre as estimativas (sujeitas a erro amostral) que não estão relacionadas ao fenômeno de interesse da pesquisa.

Estes dois efeitos correspondem a componentes associados ao processo de medição/observação e não à descrição do processo gerador da série temporal referente ao fenômeno em estudo. Ambos os efeitos são considerados explicitamente no modelo proposto neste artigo.

Em diversos países, pesquisas repetidas são usadas para acompanhamento da evolução de características socioeconômicas da população ao longo do tempo. Em todas elas, a autocorrelação induzida pelo erro amostral e o efeito de vício de grupos de rotação ocorrem. Estes efeitos foram investigados por diversos autores (Bailar, 1975; Pfeffermann, Feder e Signorelli, 1998; Bell e Carolan, 1998; etc.), principalmente na análise de resultados de pesquisas sobre mercado de trabalho. Na maioria das vezes, os resultados publicados são estimativas usuais do plano amostral e não consideram os efeitos anteriormente mencionados.

Os procedimentos de análise propostos neste artigo serão aplicados aos resultados da Pesquisa Mensal de Emprego do IBGE (PME/IBGE), mais especificamente à taxa de desemprego aberto da Região Metropolitana de São Paulo no período de 1994-2002. Esta taxa é publicada mensalmente pelo IBGE e não considera qualquer tratamento para o problema do efeito de vício de grupos de rotação ou de autocorrelação induzida pelo erro amostral, ambos associados à utilização de pesquisa repetida com painéis rotativos.

Na segunda seção, é apresentada uma breve descrição da PME, a terceira seção mostra alguns resultados encontrados que evidenciam a ocorrência do efeito de vício de grupos de rotação na PME. A quarta seção descreve, em linhas gerais, a metodologia utilizada. A quinta seção contém os resultados encontrados, e na sexta seção são apresentadas considerações finais.

2. Pesquisa Mensal de Emprego - PME

A PME é uma pesquisa domiciliar cuja população-alvo compreende as pessoas com 10 anos ou mais de idade (população em idade ativa). O plano amostral consiste de dois estágios de seleção em cada um dos municípios que compõem as seis regiões metropolitanas no âmbito da pesquisa¹: no primeiro estágio são selecionados setores censitários (unidades primárias de amostragem) e no segundo estágio são selecionadas unidades domiciliares (unidades secundárias de amostragem), dentro de cada setor censitário.

Para a seleção das unidades primárias e secundárias de amostragem é utilizado o cadastro obtido no recenseamento decenal realizado pelo IBGE. A seleção das unidades primárias de amostragem (setores) é realizada sistematicamente com probabilidade proporcional a uma medida de tamanho que corresponde ao número de unidades domiciliares do setor. As unidades secundárias de amostragem (unidades domiciliares) são selecionadas através de amostragem sistemática simples. Este tipo de plano amostral é denominado autoponderado (Série Relatórios Metodológicos Vol. 2 – Metodologia Pesquisa Mensal de Emprego IBGE/1983).

¹ Rio de Janeiro, São Paulo, Belo Horizonte, Recife, Salvador e Porto Alegre.

2.1. Esquema de rotação

A amostra de setores censitários (unidade primária de amostragem) é dividida em quatro conjuntos de setores mutuamente exclusivos, denominados grupos de rotação. Esta divisão é feita de forma que cada conjunto se espalhe por todo o território no âmbito da pesquisa, de modo que cada um seja representativo do universo das regiões metropolitanas. A cada instante da pesquisa, quatro amostras² de domicílios, uma em cada conjunto, são investigadas. Estas amostras espalham-se por todos os setores que compõem os quatro conjuntos e são denominadas painéis de domicílios.

O esquema de rotação de painéis da PME/IBGE é do tipo 4-8-4, com substituição de um painel de domicílios a cada instante. Quando um novo painel entra na amostra, permanece nela por quatro meses, ou seja, este painel é visitado durante quatro meses seguidos - visitas 1 a 4. Depois da quarta visita, o painel sai temporariamente da pesquisa durante oito meses, retornando no 13º mês a contar de sua inclusão na pesquisa. O painel permanece mais 4 meses (visitas 5 a 8) na pesquisa, sendo então retirado definitivamente da amostra (não será mais visitado).

O padrão de rotação adotado na PME ocasiona a ocorrência de meses com composições distintas de visitas. Essa não regularidade na composição da amostra da PME pode ser facilmente visualizada na Figura 1 a seguir. Nesta figura, o eixo vertical corresponde ao número da visita (1 a 8) em que se encontra o painel, cujo grupo de rotação (GR) a que pertence, por sua vez, está identificado no interior do quadro com os números I a IV. O eixo horizontal apresenta o mês e ano ao qual se refere a composição da amostra. Por exemplo: no mês de março (3) de 1994, o painel do grupo de rotação IV encontrava-se na visita 2, enquanto o mesmo painel, no mês de agosto (8) de 1995, encontrava-se na visita 7.

² Este esquema de rotação foi mantido na pesquisa até o ano de 2002, quando a pesquisa foi reformulada e a amostra passou a ser dividida em oito painéis rotativos.

Figura 1 - Composição da amostra mês a mês - PME/IBGE 1994 e 1995 RM São Paulo

Visita	1994												1995												
Mês	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	
8																									
7																									
6																									
5																									
4	IV	III	II	I	IV	III	II	I	IV	III	II	I													
3	III	II	I	IV	III	II	I	IV	III	II	I													IV	
2	II	I	IV	III	II	I	IV	III	II	I														IV	III
1	I	IV	III	II	I	IV	III	II	I														IV	III	II

Observando-se a Figura 1, que apresenta a composição de visitas nas amostras de janeiro de 1994 a dezembro de 1995, nota-se que nos anos com final par as amostras são formadas por painéis com as composições de visitas: 1-2-3-4, 4-1-2-3, 3-4-1-2, 2-3-4-1, 2-3-4-5, 3-4-5-6 e 4-5-6-7. Nos anos ímpares, as composições de visitas são: 5-6-7-8, 6-7-8-5, 7-8-5-6, 8-5-6-7, 6-7-8-1, 7-8-1-2 e 8-1-2-3. Esta irregularidade na composição de visitas da amostra é função do padrão de rotação de painéis que é adotado. Dizemos que na PME existe uma irregularidade no que se refere à composição mensal de visitas, porque as pesquisas sobre a força de trabalho, realizadas em outros países (como EUA, Canadá, Austrália, Israel, Inglaterra) possuem composição de visitas idêntica em todos os meses.

O entendimento do padrão de rotação da PME/IBGE é fundamental para o desenvolvimento da metodologia que será proposta neste trabalho para o tratamento do efeito de grupos de rotação. De fato, como foi dito acima, o padrão de rotação da pesquisa mensal de emprego - PME/IBGE - é bastante particular e diferente de qualquer outro padrão de rotação adotado em pesquisas por painéis de outros países.

A composição da amostra a cada instante na PME/IBGE promove a ocorrência de alguns eventos bem singulares. A Tabela 1 apresenta a composição de visitas da amostra para cada mês dos anos pares e dos anos ímpares. Quando o padrão de rotação utilizado em uma pesquisa é do tipo irregular, tal como na PME/IBGE, as comparações anuais das taxas de desemprego são obtidas pela análise de amostras compostas de painéis em visitas distintas. Por exemplo: a variação da taxa de desemprego de janeiro de 1994 para janeiro de 1995 é calculada com base no resultado

de duas amostras – a de janeiro de 1994 com visitas 1-2-3-4 e a de janeiro de 1995 composta de visitas 5-6-7-8. Este tipo de comparação, dada a ocorrência do efeito de vício de grupos de rotação, implica obtenção de resultados que estão contaminados pela diferença no conjunto de visitas que compõem a amostra dos mesmos meses de anos consecutivos.

Tabela 1 - Composição de visitas mês a mês nos painéis I, II, III e IV

Mês	Anos							
	Pares				Ímpares			
	Painéis				Painéis			
	I	II	III	IV	I	II	III	IV
Janeiro	1	2	3	4	5	6	7	8
Fevereiro	2	3	4	1	6	7	8	5
Março	3	4	1	2	7	8	5	6
Abril	4	1	2	3	8	5	6	7
Maiο	1	2	3	4	5	6	7	8
Junho	2	3	4	1	6	7	8	5
Julho	3	4	1	2	7	8	5	6
Agosto	4	1	2	3	8	5	6	7
Setembro	1	2	3	4	5	6	7	8
Outubro	2	3	4	5	6	7	8	1
Novembro	3	4	5	6	7	8	1	2
Dezembro	4	5	6	7	8	1	2	3

Essas peculiaridades da PME/IBGE são consideradas no desenvolvimento do modelo de séries temporais para a estimação do efeito do vício de grupos de rotação.

2.2. A reformulação PME/IBGE

No ano de 2002, ocorreu uma reformulação na PME na qual, com base no cadastro atualizado de setores censitários no Censo de 2000, o padrão de rotação de painéis da PME foi modificado. A partir deste ano, a PME adotou o mesmo padrão de rotação que é utilizado na CPS americana.

Atualmente, a amostra é constituída mensalmente por oito painéis. A cada mês, dois painéis entram na amostra – um na primeira visita e um na quinta visita. Saem os painéis que estão na quarta visita e na oitava visita. O painel que se retirou na quarta visita permanece oito meses fora da amostra retornando para a quinta visita no 13^o

mês, a contar de sua inclusão na amostra, ficando na amostra por mais quatro meses (visitas 5, 6, 7 e 8), após o que se retira da pesquisa definitivamente. O painel que entrou na primeira visita, permanece na amostra por quatro meses, é substituído após a quarta visita, repetindo o processo de retorno para a quinta visita já descrito anteriormente. Neste padrão de rotação, a cada instante existem na amostra oito painéis em visitas distintas, sempre com a composição completa de visitas - 1 a 8.

A pesquisa foi realizada simultaneamente durante o ano de 2002 de acordo com as duas metodologias (4 e 8 painéis a cada mês), obtendo-se estimativas para as mesmas características de interesse. Durante este ano foram publicadas, mensalmente, duas taxas de desemprego aberto na semana: a taxa obtida segundo a metodologia nova e aquela calculada com base na metodologia adotada até dezembro de 2001.

A modificação na metodologia da PME/IBGE implica uma mudança no padrão de rotação de painéis da amostra e, conseqüentemente, na modificação da estrutura de correlação entre erros amostrais e do efeito de vício de grupos de rotação. Como é de interesse dos analistas a estimação dos componentes de tendência, sazonalidade e ciclo da série da taxa de desocupação, deve-se considerar a estimação dos efeitos relativos à utilização de amostragem (erro amostral e efeito de vício de grupos de rotação) nestas análises. Para comparação entre as duas metodologias (encadeamento da série) deve-se levar em conta as diferenças entre metodologias e suas implicações. Este artigo apresenta um modelo completo de análise da série em sua metodologia antiga, sendo que a metodologia adotada pode ser adaptada para a análise da série em sua nova fase. A metodologia aqui apresentada permite também estimar o componente de efeito de vício de rotação que constitui um efeito de confundimento que prejudica a interpretação da evolução do fenômeno ao longo do tempo.

Neste artigo os dados utilizados referem-se à série de valores obtidos para o período de janeiro de 1994 a dezembro de 2002, totalizando 108 observações, pesquisados de acordo com a metodologia da PME anterior à sua reformulação.

3. O vício de grupos de rotação na PME/IBGE

A ocorrência de vício de grupos de rotação da PME/IBGE foi primeiramente identificada por Pfeffermann, Silva e Freitas (2000) que analisaram dados da PME para a Região Metropolitana de São Paulo no período de maio de 1992 a maio de 1998. Especificamente, os autores trabalharam apenas com os meses que possuíam composição com visitas 1 a 4 e 5 a 8 em qualquer ordem (foram excluídos os meses de outubro, novembro e dezembro). Para este período, estimativas das médias das taxas de desemprego por visita na Região Metropolitana de São Paulo foram calculadas e são apresentadas na Tabela 2, na qual observa-se uma queda sistemática nos valores das médias por visita. Este tipo de comportamento é característico da existência de um efeito de vício relativo ao grupo de rotação (visita).

Tabela 2 - Médias da taxa de desemprego aberto por visita (%) - SP Maio 1992 a Maio 1998

Visitas	1	2	3	4	5	6	7	8
Estimativa média	7,28	6,73	6,62	6,09	5,95	5,73	5,43	5,38

Fonte: Pfeffermann, Silva e Freitas 2000.

O efeito de declínio da taxa de desemprego que é observado na tabela anterior é também observado em diversas pesquisas sobre força de trabalho que são realizadas em outros países, sempre quando as pesquisas são feitas utilizando-se painéis rotativos. Em trabalhos de autores como Pfeffermann (1991) e Kumar e Lee (1983), considera-se que a soma dos efeitos de vício de grupos de rotação é nula, ou seja, uma vez que existem painéis em todas as visitas a cada instante, os efeitos associados a cada visita em particular se cancelam, fazendo com que a estimativa conjunta – que é a taxa de desemprego publicada – não seja tendenciosa.

No caso da PME, como já mencionado, não se tem a mesma composição de visitas a cada instante. Conseqüentemente, levando-se em consideração a hipótese de que a soma dos efeitos de visitas é nula, a estimativa a cada instante é viciada. A partir desta suposição tentou-se verificar a partir dos dados se havia diferença entre as taxas médias por composição da amostra a cada instante. A Tabela 3 apresenta os resultados obtidos.

Tabela 3 - Média da taxa de desemprego aberto por composição de visitas - PME/IBGE São Paulo 1994/2001

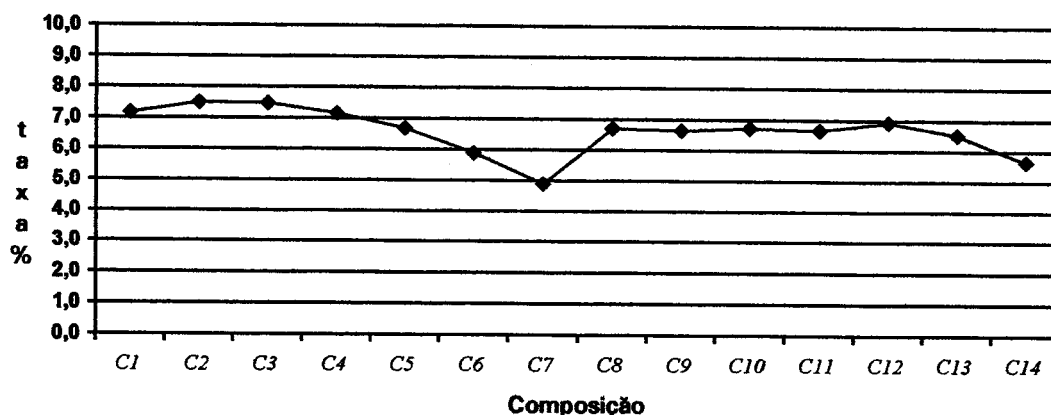
Composição	Visitas	Média da taxa de desemprego	Ano
C ¹	1-2-3-4	7,17	Par
C ²	5-6-7-8	6,71	Ímpar
C ²	2-3-4-1	7,51	Par
C ³	6-7-8-5	6,66	Ímpar
C ³	3-4-1-2	7,47	Par
C ¹⁰	7-8-5-6	6,74	Ímpar
C ⁴	4-1-2-3	7,15	Par
C ¹¹	8-5-6-7	6,68	Ímpar
C ⁵	2-3-4-5	6,69	Par
C ¹²	6-7-8-1	6,94	Ímpar
C ⁶	3-4-5-6	5,91	Par
C ¹³	7-8-1-2	6,53	Ímpar
C ⁷	4-5-6-7	4,89	Par
C ¹⁴	8-1-2-3	5,66	Ímpar

Fonte: PME/IBGE. Elaboração própria.

A partir da tabela, nota-se que as composições, quando comparadas duas a duas segundo os meses (ou mês) em que ocorrem, diferem. Sendo que as composições que possuem painéis na primeira visita têm médias superiores às que não possuem. Isto indica que existe um efeito nas médias das taxas por composição relacionado ao conjunto de visitas que compõem a taxa de desemprego mensal e não relacionado ao fenômeno desemprego, ou seja, há evidência da ocorrência de um efeito de vício de grupos de rotação para a composição de visitas mês a mês.

O gráfico a seguir apresenta as médias por composição nos meses de anos pares e ímpares (dados da Tabela 3).

Gráfico 1 - Média da taxa de desemprego aberto por composição
- Região Metropolitana de São Paulo



As composições C1, C2, C3, C4, C5, C6 e C7 ocorrem em anos pares e as composições C8, C9, C10, C11, C12, C13 e C14 ocorrem em anos ímpares. É fácil perceber que nos anos ímpares as taxas de desemprego são em média menores que as dos anos pares nos meses que vão de janeiro a setembro. Nos meses de outubro, novembro e dezembro, as taxas dos anos ímpares são em média maiores que as observadas nos anos pares. Estes últimos meses correspondem aos meses em que ocorre a transição entre anos pares e ímpares.

A análise mostrou que existem indícios da ocorrência de um efeito que não está associado ao fenômeno e sim às visitas (individualmente ou em composição). Estes efeitos serão incluídos no modelo de séries temporais.

4. Abordagem metodológica

Para a estimação das taxas de desemprego aberto da Região Metropolitana de São Paulo será utilizado um modelo de séries temporais de espaços de estados. Este tipo de representação permite a conjugação de modelos de decomposição e auto-regressivos, simultaneamente.

Modelos para representar dados provenientes de pesquisas por amostragem devem considerar explicitamente a presença do efeito associado ao método usado para obter os dados, ou seja, o erro amostral. Assim tem-se:

$$y_t = \theta_t + e_t + K_t \quad (4.1)$$

onde

y_t é a estimativa não-viciada, ou assintoticamente não-viciada, (usualmente obtida a partir de um estimador clássico) de uma característica da pesquisa no instante t ; no caso da PME, um exemplo é a média das estimativas das taxas de desemprego dos painéis,

θ_t é o valor do parâmetro que se deseja estimar;

e_t é o erro amostral associado a esta estimativa; e

K_t é o erro não-amostral.

O erro amostral não representa a estrutura inerente ao processo que se deseja estimar, no caso a taxa de desemprego aberto, e sim a variabilidade amostral, decorrente do fato de se utilizar uma amostra para obter estimativas. Em pesquisas repetidas de painéis rotativos, ou painéis fixos, a estrutura da amostra induz uma correlação entre as estimativas que não está relacionada à característica que se deseja medir (θ_t). Diversos autores, entre eles Pfeffermann (1991) e Bell e Carolan (1998), ao utilizar modelos de espaços de estado para trabalhar com dados de pesquisas repetidas, introduzem nos modelos estruturais propostos para estimação de características de interesse (como taxas de desemprego) uma componente no vetor de estado relativa ao erro amostral. Dessa forma, a equação, possivelmente estocástica, que representa a evolução do erro amostral ao longo do tempo passa a ser contemplada pelo modelo de espaços de estado.

O erro não-amostral representado na equação (4.1) corresponde a erros que ocorrem durante o processo de coleta de dados³ e constituem uma possível fonte para o efeito tendencial denominado efeito de vício de grupos de rotação. No modelo proposto, para estimação do efeito de vícios de grupos de rotação (erro não-amostral), será definida uma componente específica para representar este fenômeno. O efeito conjunto dos erros amostral e não-amostral será denominado erro de pesquisa. O modelo será

³ Não-resposta, erro de classificação.

avaliado através de gráficos de resíduos e dos testes Jarq-Bera de normalidade dos resíduos, Box-Ljung de autocorrelação de resíduos e o teste H de homocedasticidade dos resíduos (Harvey 1989).

4.1. Representação do erro amostral

Pfeffermann, Feder e Signorelli (1998) apresentam uma maneira, em pesquisas repetidas com utilização de painéis rotativos, de estimar a função de autocorrelação e autocorrelação parcial associada à evolução do erro amostral a partir das estimativas de cada um dos painéis e a estimativa composta por todos os painéis. A função de autocorrelação e autocorrelação parcial é determinada a partir de uma quantidade denominada pseudo-erro. Os pseudo-erros são definidos como:

$$e_{pt} = y_t^j - y_t \tag{4.2}$$

sendo

y_t^j a taxa estimada com as observações do painel j no instante t ;

$y_t = \sum_{j=1}^J \frac{y_t^j}{J}$ a taxa no instante t calculada como a média das taxas dos painéis; e

J o número de painéis a cada instante na pesquisa.

Se o estimador da taxa por painel for um estimador não-viciado para a taxa populacional θ_t , então:

$$\begin{aligned} e_{pt} &= y_t^j - y_t = y_t^j - \frac{1}{J} \sum_{j=1}^J y_t^j \\ &= (y_t^j - \theta_t) - \left(\frac{1}{J} \sum_{j=1}^J y_t^j - \theta_t \right) \\ &= e_t^j - \frac{1}{J} \sum_{j=1}^J e_t^j = e_t^j - e_t \end{aligned} \tag{4.3}$$

Nota-se que, apesar de os erros amostrais não serem observados, o pseudo-erro pode ser calculado a partir das estimativas das taxas dos painéis e da taxa conjunta. Com base nestes pseudo-erros, Pfeffermann, Feder e Signorelli (1998) desenvolveram um método para estimar a função de autocorrelação dos erros amostrais. O método proposto pressupõe duas condições:

1. $Cov(e_{t-h}^j, e_t^k) = 0$ se $k \neq j, \forall t, h$, ou seja, os erros amostrais das estimativas das taxas de painéis que não fazem parte do mesmo grupo de rotação são não correlacionados.

2. $Cov(e_{t-h}^j, e_t^j) = \gamma_h^j \forall t, h$ e $j = 1, 2, \dots, J$, isto é, a autocovariância dos erros amostrais das estimativas das taxas de painéis de um mesmo grupo de rotação é função apenas da defasagem de tempo e do grupo de rotação a que se refere.

Estas duas condições determinam que cada parte da amostra – cada um dos quatro conjuntos distintos em que é dividida a amostra inicial – é homogênea no que diz respeito ao erro amostral, ou seja, quando um painel é substituído por outro do mesmo conjunto considera-se que a estrutura de autocovariância do erro amostral permanece a mesma (Pfeffermann, Feder e Signorelli, 1998, encontraram resultados empíricos que indicam que tal hipótese é válida). A partir destas suposições e de (4.3) os autores estabelecem que a autocorrelação de defasagem h do erro amostral é dada por:

$$\rho_h = \frac{\sum_{j=1}^J C_h^j}{\sum_{j=1}^J C_0^j} = \frac{\gamma_h}{\gamma_0}, \quad (4.4)$$

onde C_h^j é a autocovariância de defasagem h dos pseudo-erros associados às estimativas da subamostra referente ao painel j .

Este resultado é válido quando a estimativa é influenciada apenas pelo erro amostral. Entretanto, no caso da PME/IBGE (e de outras pesquisas que utilizam-se de painéis rotativos), obtém-se a cada instante uma taxa que incorpora, além do erro amostral, o efeito do vício de grupos de rotação, conforme evidências empíricas apresentadas anteriormente. Neste caso a taxa observada é decomposta na forma:

$$y_t^c = \theta_t + I_t + e_t + K_t^c, \quad (4.5)$$

sendo

y_t^c a estimativa da taxa de desemprego no mês t com a composição de visitas c ;

θ_t taxa no instante t ;

e_t erro amostral no instante t ;

I_t ruído associado à observação; e

K_t^c efeito da composição de visitas no instante t . A composição de visitas é determinada a partir da Figura 1.

Adicionalmente a estimativa da taxa de desemprego por painel é descrita por:

$$y_t^j = \theta_t + I_t^j + e_t^j + K_t^{vj}, \quad (4.6)$$

sendo

y_t^j a estimativa da taxa de desemprego para o painel j no instante t ;

θ_t a taxa no instante t ;

e_t^j o erro amostral do painel j no instante t ;

I_t^j o ruído associado à observação do painel j ; e

K_t^{vj} o efeito da visita v no painel j no instante t .

A composição a cada instante é formada pelas visitas dos quatro painéis observados. De acordo com esta representação, o pseudo-erro seria:

$$e_p^j = y_t^j - y_t^c = (\theta_t + e_t^j + K_t^{vj}) - (\theta_t + e_t + K_t^c) = (e_t^j - e_t) - (I_t - I_t^j) - (K_t^c - K_t^{vj}) \quad (4.7)$$

Para se utilizar o resultado apresentado por Pfeffermann, Feder e Signorelli (1998) e obter estimativas para as funções de autocorrelação e autocorrelação parcial do erro

amostral, é necessário subtrair dos pseudo-erros as parcelas que não correspondem ao fenômeno associado ao erro amostral (relativas ao efeito de vício de grupos de rotação e a componente de ruído I). Considerando-se que as componentes de ruído (I_t e I_t^j) – de média nula e variância constante – não apresentam correlação com o erro amostral ou autocorrelação serial, seria necessário estimar na expressão (4.7) apenas a parcela relativa ao efeito de vício de grupos de rotação. Optou-se por estimar $K_t^c - K_t^{vj}$ de acordo com a seguinte expressão:

$$K_t^c - K_t^{vj} = \frac{\sum_{t=1}^T y_t^c}{N^c} - \frac{\sum_{j=1}^4 \sum_{t=1}^T y_t^{vj}}{M^v} \quad (4.8)$$

sendo

N^c o número de vezes em que ocorre a composição de tipo c ;

M^v o número de vezes que ocorre a visita v considerando todos os 4 painéis;

y_t^c a taxa para a composição do tipo c de visitas; e

y_t^{vj} a taxa para a visita v do painel j .

Ou seja, estima-se o efeito da diferença entre o vício de grupos de rotação da composição (K_t^c) e da visita do painel j (K_t^{vj}), no instante t , pela diferença entre as médias das taxas com a referida composição e a média das taxas dos painéis com visita v_t . O estimador proposto foi construído de maneira similar a um apresentado por Pfeffermann, Feder e Signorelli (1998) na aplicação de seus resultados à pesquisa sobre força de trabalho realizada mensalmente na Austrália.

Estimando o efeito de vício de grupos de rotação pelo estimador proposto e aplicando-se a metodologia apresentada por Pfeffermann, Feder e Signorelli (1998) ao

resultado obtido para os pseudo-erros descontado o efeito de vício de grupos de rotação, representou-se a evolução do erro amostral no modelo de espaços de estados por um modelo auto-regressivo de ordem 1 (AR(1)).

Quanto a representação no modelo do efeito de vício de grupos de rotação, uma análise da Figura 1, apresentada anteriormente, mostra que as composições repetem-se, na mesma ordem, a cada 24 meses, como um ciclo. Pfeffermann (1991) representou o efeito de vício de grupos de rotação como variáveis auxiliares indicadoras, com a restrição de que, a cada instante, a soma dos efeitos de visita seja nula. A suposição de soma nula para os efeitos de visita foi inicialmente proposta por Kumar e Lee (1983). Em ambos os casos as pesquisas que foram consideradas possuíam planos amostrais com painéis rotativos em que, a cada instante, todas as visitas estavam presentes.

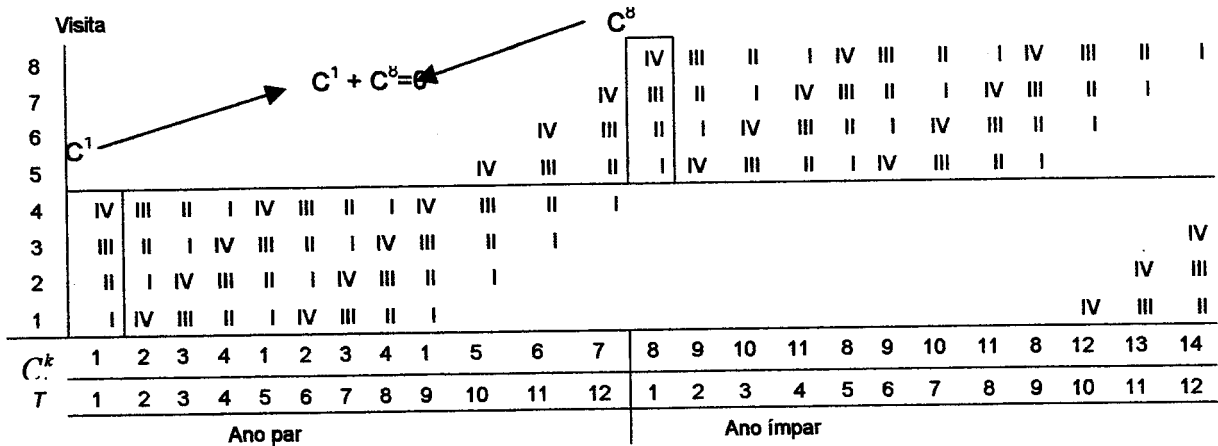
No caso da PME, como já mencionado, não se tem a mesma composição de visitas a cada instante. Conseqüentemente, levando-se em consideração a hipótese de que a soma dos efeitos de visitas é nula, a estimativa a cada instante é viciada. Propõe-se, então, um modelo para os efeitos de composição de visitas a cada instante do tempo. Estes efeitos foram incluídos no modelo através de uma representação diferenciada da equação de observação. Nesta representação, o modelo de espaços de estados deixa de ser invariante no tempo: a variável que corresponde ao efeito de composição de visitas não evolui estocasticamente de forma Markoviana, mas evolui segundo um mecanismo determinístico – o plano de rotação.

Para incorporação dos efeitos de vícios de grupos de rotação, o modelo proposto considera fixos os efeitos, podendo os mesmos ser distintos nos diferentes painéis. A representação destes efeitos é incorporada na equação de observação do modelo por meio de vetores Z_t que variam de acordo com a composição observada a cada instante da pesquisa. Assim, para cada observação, seleciona-se o efeito da composição que está presente, fazendo com que Z_t varie de forma determinística ao longo do tempo.

Para que seja possível estimar os parâmetros do modelo, faz-se necessária a utilização de uma restrição de maneira análoga à definida para os fatores sazonais. Observando-se o conjunto de composições distintas possíveis, considerada a ordem dos painéis, o esquema de rotação e a suposição de que os efeitos de visitas individuais têm

soma nula², optou-se pela representação do efeito de forma que duas composições separadas por 12 meses tenham soma nula. A figura a seguir apresenta as composições possíveis durante o período de um ciclo de rotação.

Figura 2 - Ciclo de composições em um período de 24 meses



Na Figura 2, o eixo horizontal apresenta as C_t^k , $k=1,2,\dots,14$ possíveis composições no período de dois anos consecutivos (1994-1995, por exemplo). Ao todo temos 14 composições, sendo que, segundo a restrição proposta, duas a duas elas se cancelam, ou seja:

$$C^1 + C^8 = 0; C^2 + C^9 = 0; C^3 + C^{10} = 0; C^4 + C^{11} = 0; C^5 + C^{12} = 0; C^6 + C^{13} = 0 \text{ e } C^7 + C^{14} = 0$$

Sob esta restrição, a soma das composições é nula a cada período completo de 24 meses (dado que todas as visitas ocorrem o mesmo número de vezes) e nula para a soma das composições distintas separadas por 12 meses.

O modelo proposto para a estimativa da taxa de desemprego é dado por:

$$y_t = \theta_t + I_t + e_t + C_t^k, \tag{4.9}$$

² Soma dos efeitos das visitas 1 a 8 é nula.

sendo que

y_t é a estimativa não-viciada da taxa de desemprego no instante t ;

I_t é o ruído associado à observação;

e_t é o erro amostral associado a esta estimativa;

C_t^k é o efeito da composição de visitas do tipo k no instante t ; e

θ_t é o valor do parâmetro que se deseja estimar.

Supõe-se ainda que:

$e_t \sim N(0, \sigma_e^2)$, $I_t \sim N(0, \sigma_I^2)$ e o efeito das composições fixo.

A partir deste modelo proposto, e das considerações anteriormente feitas para a representação do efeito de vício de grupos de rotação, as equações estruturais do modelo são:

$$\begin{aligned} y_t &= \theta_t + I_t + e_t + C_t^k & L_t &= L_{t-1} + R_{t-1} + \eta_t^L & S_t &= -\sum_{j=1}^{11} S_{t-j} + \eta_t^S & (4.10) \\ \theta_t &= L_t + S_t & R_t &= R_{t-1} + \eta_t^R \end{aligned}$$

$$e_t = \phi e_{t-1} + \eta_t^e$$

$$C_t^k = -C_{t+12}^k$$

com

L_t correspondendo a componente de tendência;

R_t correspondendo a componente de variação de tendência; e

S_t correspondendo a componente sazonal.

$$\eta_t^L \sim N(0, \sigma_L^2)$$

$$\eta_t^R \sim N(0, \sigma_R^2)$$

$$\eta_t^S \sim N(0, \sigma_S^2)$$

$$\eta_t^e \sim N(0, \sigma_e^2)$$

$$I_t \sim N(0, \sigma_I^2)$$

e $\eta_t^L, \eta_t^R, \eta_t^S, \eta_t^e, I_t$ não-correlacionadas.

A representação em modelos de espaço de estados é dada pelas equações de observação e transição:

$$\begin{aligned} y_t &= Z_t' \alpha_t + I_t \\ \alpha_t &= T \alpha_{t-1} + G \eta_t \end{aligned} \quad (4.11)$$

Os valores iniciais para a estimação dos hiperparâmetros foram obtidos a partir de dois outros modelos ajustados anteriormente: um modelo estrutural básico ajustado por meio da utilização do programa STAMP e um modelo que considerou apenas o erro amostral em sua representação.

5. Aplicação

O modelo proposto foi ajustado ao conjunto de dados da taxa de desemprego aberto para a Região Metropolitana de São Paulo no período de janeiro de 1994 a dezembro de 2002. Os resultados obtidos foram satisfatórios no que diz respeito às hipóteses de normalidade, homocedasticidade e autocorrelação nula dos resíduos encontrados. Os gráficos e estatísticas de teste utilizadas para avaliar tais hipóteses

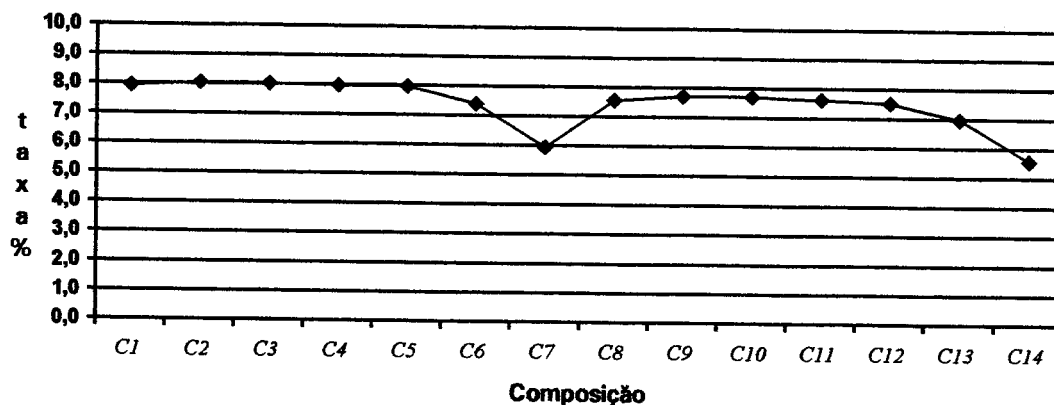
encontram-se em anexo. As estimativas finais dos hiperparâmetros do modelo são:

Tabela 4 - Estimativas dos hiperparâmetros

Hiperparâmetro	Estimativa	Variância
σ_L^2	5,603E-6	4,627E-14
σ_S^2	4,691E-7	7,184E-16
σ_I^2	2,794E-7	2,184E-16

A partir do sinal estimado por este modelo tem-se o gráfico das médias das taxas por composição.

Gráfico 2 - Média da taxa de desemprego por composição



Percebe-se, a partir do Gráfico 2, que ainda existe uma pequena diferença entre as médias das taxas por composição para anos pares (C1 a C7) e ímpares (C8 a C14). No entanto, a diferença notada no comportamento é bem menor do que a obtida para a taxa publicada. A Tabela 5 apresenta as estimativas para as médias das taxas por composição.

Tabela 5 - Média das taxas de desemprego por composição

Composição	Média da taxa de desemprego	Diferença	Modelo	Diferença
C^7	7,17	0,46	7,96	0,40
C^8	6,71		7,56	
C^2	7,51	0,85	8,05	0,34
C^9	6,66		7,71	
C^3	7,47	0,73	8,03	0,32
C^{10}	6,74		7,71	
C^4	7,15	0,47	7,99	0,36
C^{11}	6,68		7,63	
C^5	6,69	-0,25	7,97	0,45
C^{12}	6,94		7,52	
C^6	5,91	-0,62	7,39	0,42
C^{13}	6,53		6,97	
C^7	4,89	-0,77	5,94	0,37
C^{14}	5,66		5,57	

Fonte: Elaboração própria.

As estimativas dos efeitos de cada composição são apresentadas na Tabela 8 a seguir:

Tabela 6 - Estimativas para o efeito de cada composição

Composição	Estimativa	Variância
$C^7 = - C^8$	0,003761	1,260E-3
$C^2 = - C^9$	0,00495	1,499E-3
$C^3 = - C^{10}$	0,00454	1,547E-3
$C^4 = - C^{11}$	0,00418	1,498E-3
$C^5 = - C^{12}$	0,00229	1,688E-3
$C^6 = - C^{13}$	0,000226	1,757E-3
$C^7 = - C^{14}$	-0,00434	3,017E-3

Fonte: Elaboração própria.

A partir destes resultados é elaborada, a seguir, uma comparação entre a taxa de desemprego estimada por meio do sinal – tendência e sazonalidade - do modelo e a taxa de desemprego divulgada mensalmente pelo IBGE.

5.1. Análise comparativa dos resultados obtidos pelo modelo e as taxas publicadas

Ao utilizar um modelo estrutural, considera-se como estimativa da taxa de desemprego a cada mês (série) a soma das estimativas das componentes de tendência e sazonalidade (estimativas suavizadas do sinal). Este é o valor estimado pelo modelo para a taxa de desemprego aberto. Os gráficos das estimativas de tendência e sazonalidade para a taxa de desemprego são apresentados a seguir:

Gráfico 3 - Tendência da taxa de desemprego aberto

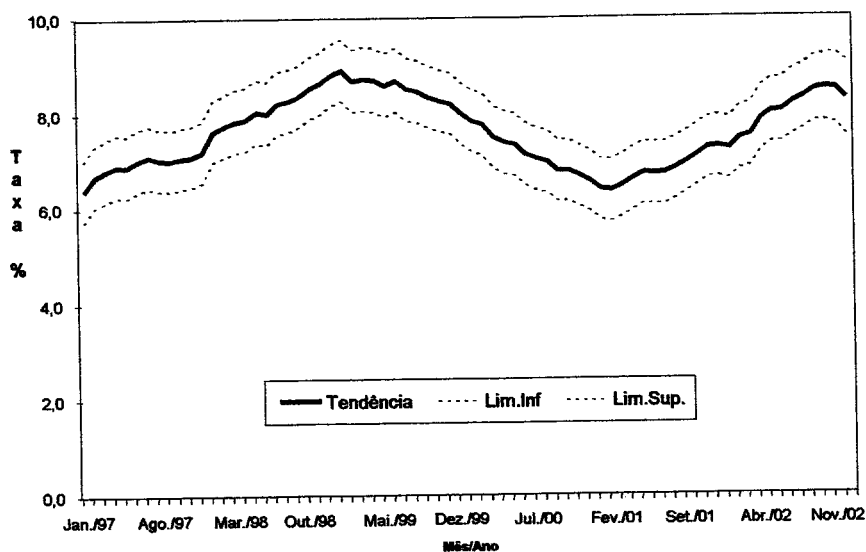
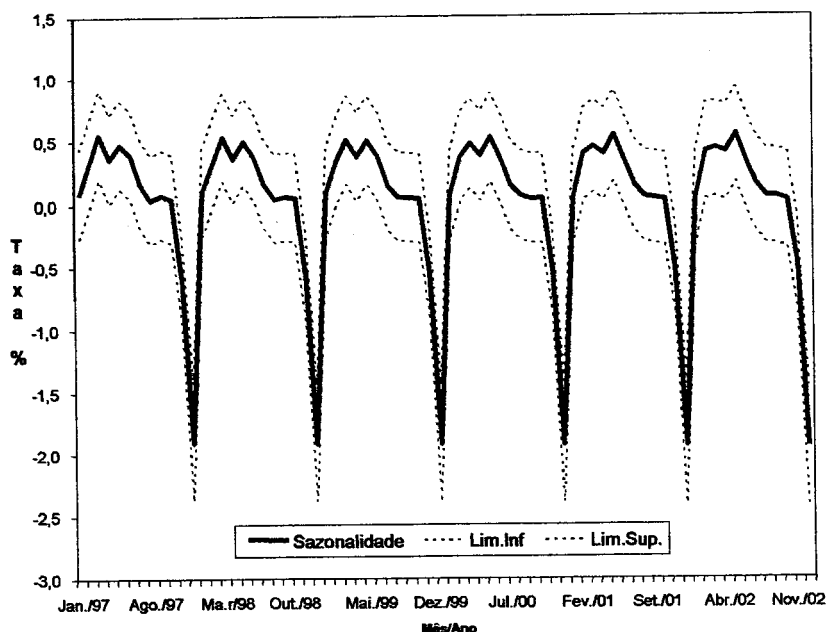


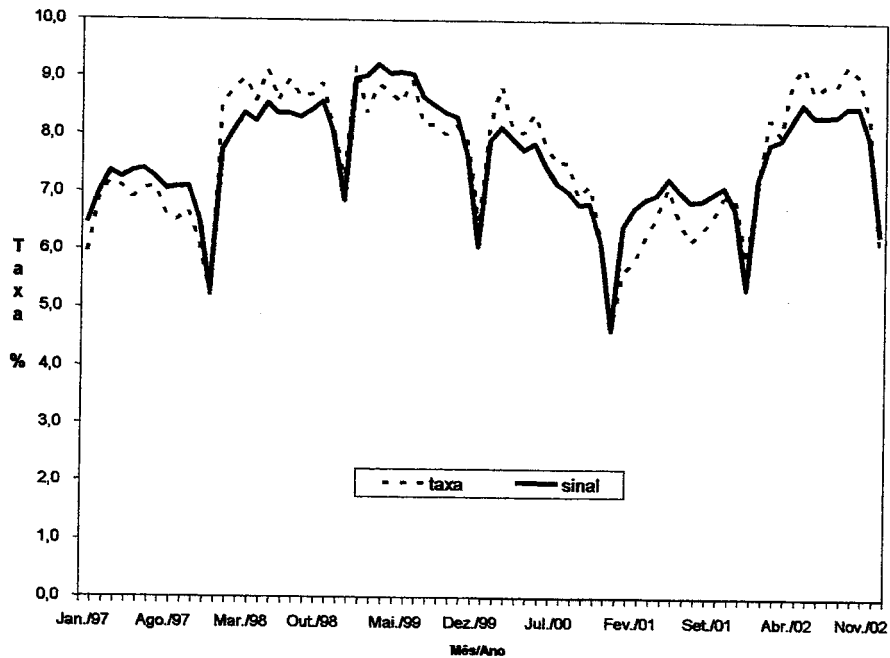
Gráfico 4 - Sazonalidade da taxa de desemprego aberto



Os Gráficos 3 e 4 indicam que a sazonalidade possui valores muito inferiores aos da tendência. A tendência varia entre 0,05 e 0,095, aproximadamente, enquanto a sazonalidade varia entre -0,03 e 0,01. O comportamento sazonal caracteriza-se por uma forte queda nos meses de outubro, novembro e dezembro – tradicionalmente meses de maior oferta de emprego.

O sinal – taxa de desemprego estimada pelo modelo – deve ser comparado à taxa publicada para permitir uma avaliação da diferença no comportamento das duas séries.

Gráfico 5 - Taxa de desemprego aberta publicada x Taxa de desemprego aberto estimada



Comparando-se os resultados das duas taxas, sob a hipótese de que o modelo adotado explica satisfatoriamente o processo (fenômeno e método de pesquisa), percebe-se que em algumas ocasiões a taxa publicada subestimaria, e em outros superestimaria, a taxa de desemprego aberto obtida pelo método de extração do sinal. Da mesma forma pode-se comparar as variações mensais e anuais da taxa de desemprego para as duas estimativas.

Gráfico 6 - Variação mensal da taxa de desemprego aberto

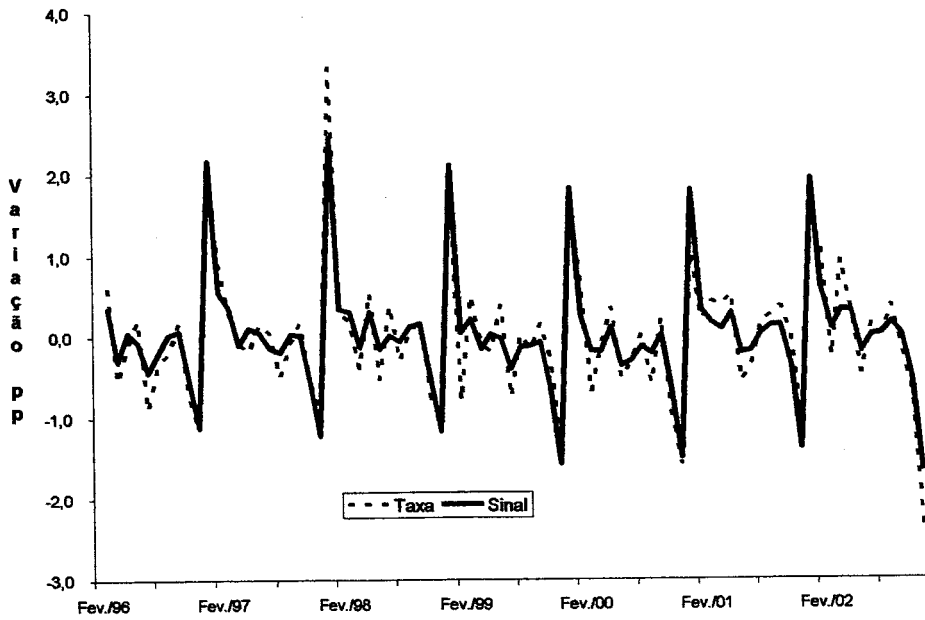
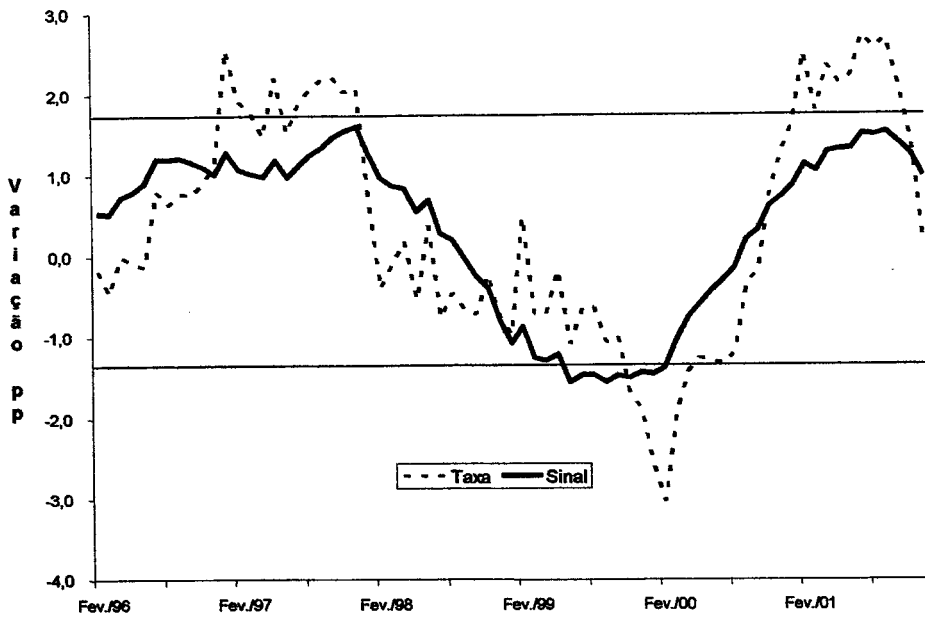


Gráfico 7 - Variação anual da taxa de desemprego aberto



Os resultados obtidos para a diferença referente à variação entre meses consecutivos para o modelo e a taxa publicada não são tão expressivos quanto aos obtidos no caso da comparação das duas metodologias na variação anual. Na variação mensal, em alguns meses a diferença é superior para o sinal e em outros é inferior, quando comparada à diferença da taxa publicada. Verifica-se, no caso da diferença de taxas entre dois meses em anos consecutivos, que a evolução do sinal é bem mais suave que a observada para a taxa publicada. Este comportamento está fortemente relacionado à restrição imposta pelo modelo para a introdução da componente de efeito de vício de grupos de rotação.

Desse modo, a adoção da taxa de desemprego estimada pelo modelo de séries temporais alteraria de forma significativa os resultados para as variações anuais e de forma mais atenuada as variações mensais. Por exemplo, a variação anual em fevereiro de 2001 passaria de -3% para -1,4%, uma alteração positiva de aproximadamente 50% entre o resultado publicado e o obtido no modelo. Por outro lado, em agosto de 2002 a variação anual passaria de 2,7% para 1,5%, uma alteração negativa na variação de aproximadamente 44%.

Gráfico 8 - *Box-plot* da variação mensal da taxa de desemprego aberto

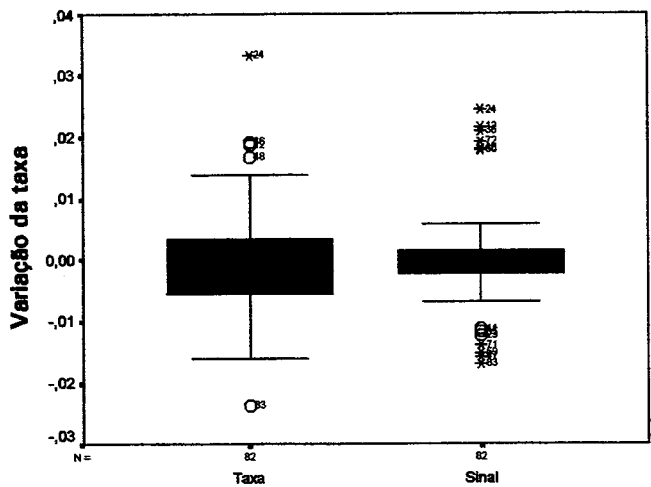
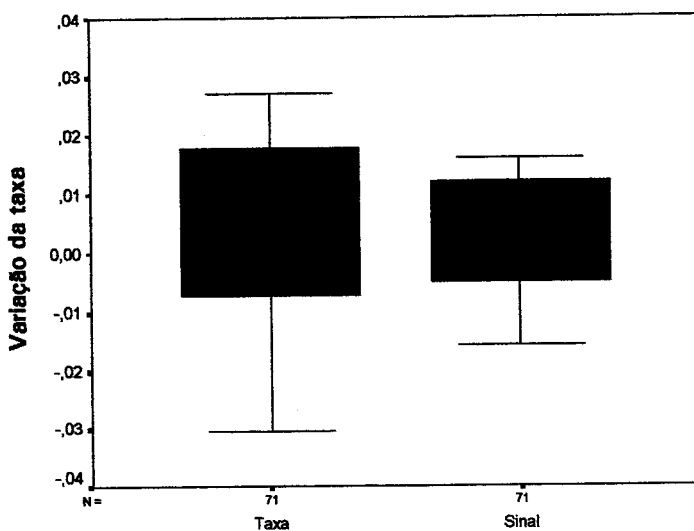


Gráfico 9 - Box-plot da variação anual da taxa de desemprego aberto



Os Gráficos 8 e 9 apresentam as distribuições das variações mensal e anual das taxas estimadas de desemprego aberto, para a taxa publicada e o sinal estimado pelo modelo. Em ambos os casos há uma redução considerável na variabilidade observada para as variações obtidas a partir de taxas estimadas pelo modelo. A mediana das variações mensais publicada e obtida pelo modelo não são diferentes, por outro lado a mediana das variações anuais obtidas pelo modelo é superior a observada para as taxas publicadas.

6. Considerações finais e trabalhos futuros

O efeito de vício de grupos de rotação em pesquisas repetidas já foi estudado por diversos autores e todos concordam que tal fenômeno ocorre sistematicamente nas pesquisas que utilizam este tipo de desenho amostral, sem no entanto apresentar uma causa específica para tal comportamento. Este trabalho procurou, por meio da utilização de metodologia de séries temporais aplicadas a pesquisas repetidas estimar o efeito resultante do desenho amostral adotado na PME antes de sua reformulação, desconsiderando as causas que podem estar associadas à ocorrência do efeito tendencial comum em tais pesquisas.

Os resultados obtidos com a utilização de metodologia para a estimação da taxa de desemprego aberto (PME/SP) apresentaram como característica marcante a suavização das diferenças entre as taxas de dois meses de anos consecutivos (variação anual), quando comparadas com as obtidas segundo a taxa publicada. Este resultado está fortemente relacionado às hipóteses restritivas que foram feitas sobre o comportamento do efeito de vício de grupos de rotação na composição, hipóteses que são necessárias para que a metodologia aplicada forneça resultados estáveis (estabilidade do filtro). Outras restrições poderiam ser feitas sobre o comportamento do efeito, mas entre todas as que foram testadas, sempre considerando o efeito como fixo (não-estocástico), apenas esta apresentou adequação no ajuste. As demais apresentaram não-estabilidade da variância, ou problemas de autocorrelação forte dos resíduos obtidos para o modelo ajustado. Entre as diversas tentativas deve-se mencionar aqui aquela na qual o modelo ajustado apresentou estabilidade na variância, mas resultados de resíduos que violaram a hipótese de não-autocorrelação. Tal restrição considerou os efeitos de vício para composições fixas, mas sendo que apenas oito composições distintas foram representadas. As oito composições são: as que possuem visitas 1 a 4, independentemente de que painel esteja em cada visita – C1;

a composição de visitas 2,3,4,5 – C2;

a composição de visitas 3,4,5,6 – C3;

a composição de visitas 4,5,6,7 – C4;

as que possuem visitas 5 a 8, independentemente de que painel esteja em cada visita – C5;

a composição de visitas 6,7,8,1 – C6;

a composição de visitas 7,8,1,2 – C7; e

a composição de visitas 8,1,2,3 – C8.

Para este conjunto de composições a restrição adotada foi de que a soma de cada par: (C1,C5); (C2,C6); (C3,C7) e (C4,C8) é nula. Esta forma de representação, que considera não-existência de diferença nos efeitos para cada uma das partes em que é dividida a amostra de setores, não apresentou bons resultados. Os outros modelos

utilizados foram omitidos por não apresentarem nenhum tipo de informação relevante sobre o problema.

Outras tentativas de aprimorar o modelo podem ser feitas. Ainda trabalhando com modelos de espaços de estados univariados, pode-se construir um modelo mais complexo, no qual os efeitos não sejam considerados fixos, podendo variar estocasticamente. Este tipo de abordagem implicaria aumentar o número de hiperparâmetros do modelo. Uma outra abordagem, a partir das estimativas dos quatro painéis, por meio de modelos de espaços de estados multivariados, permitiria estimar os efeitos de visita ao invés dos efeitos das composições de visitas. Neste tipo de representação, as componentes de tendência e sazonalidade são componentes comuns, o que implica ter que se representar o problema por meio de um modelo de análise fatorial dinâmica. Tal modelo forneceria resultados para o efeito de vício de grupos de rotação associado a cada visita, e não à composição mensal de visitas. Estas propostas de diferentes abordagens constituem linhas de trabalhos que podem ser elaborados futuramente.

Ainda existem muitas possibilidades de avanço para o tema abordado neste trabalho, algumas considerando mudanças nas pesquisas, outras baseadas na comparação dos resultados a outras pesquisas e bases de dados distintas. Este trabalho apresentou alguns resultados e desenvolvimentos com os quais se espera contribuir para o aumento do conhecimento sobre o problema colocado.

Referências bibliográficas

- ANDERSON, B. D. O., MOORE, J. B., (1979). *Optimal filtering*. Prentice-Hall.
- BAILAR, B. A., (1975). The effects of rotation group bias on estimates from panel surveys, *Journal of American Statistical Association*, 70, pp. 23-30.
- BELL, P. A., CAROLAN, A. M., (1998). Trend estimation for small areas from continuing survey with controlled sample overlap, *Working Paper N° 98/1*, Australian Bureau of Statistics.
- CRUZ, M.M.,(2001). Estimação de variâncias para séries dessazonalizadas pelo método X-12 ARIMA, considerando o desenho amostral. Rio de Janeiro: IBGE, Escola Nacional de Ciências Estatísticas, 194 p.
- DUNCAN, G. J.; KALTON, G., (1987). Issues of design and analysis of surveys across time, *International Statistical Review*, vol. 55, n. 1, 97-117.

- FIGUEIRÊDO, J. S., (2003).** Não-resposta diferencial e tendenciosidade de grupos de rotação na Pesquisa Mensal de Emprego. Rio de Janeiro: IBGE, Escola Nacional de Ciências Estatísticas, 325 p.
- HARVEY, A. C., (1989).** Forecasting, structural time series models and the kalman filter, Cambridge University Press. 554 p.
- IBGE, (1983).** Série Relatórios Metodológicos, vol. 2, 82 p, Rio de Janeiro: IBGE.
- KALMAN, R. E., (1960).** A new approach to filtering and prediction problems, *Journal of Basic Engineering Transactions ASME, Series D* 82, 35-45.
- KALMAN, R. E.; BUCY, R. S., (1961).** New results in linear filtering and prediction theory, *Journal of Basic Engineering Transactions ASME, Series D* 83, 95-108.
- KOOPMAN, S. J.; HARVEY, A. C.; DOORNIK, J. A.; SHEPHARD, N., (1995).** Stamp 5.0: structural time series analyser, modeller and predictor. London: Chapman & Hall, 382 p.
- KUMAR, S.; LEE, H., (1983).** Evaluation of Composite Estimation for the Canadian Labor Force Survey, *Survey Methodology*, 9, 1-24.
- MORETTIN, P. A.; TOLOI, C. M. C., (1987).** Previsão de séries temporais, São Paulo: Atual, 436 p.
- PFEFFERMANN, D., (1991).** Estimation and seasonal adjustment of population means using data from repeated surveys, *Journal of Business & Economic Statistics*, vol. 9, n. 2, 163-175.
- PFEFFERMANN, D., FEDER, M.; SIGNORELLI, D., (1998).** Estimation of autocorrelations of survey errors with application to trend estimation in small areas, *Journal of Business & Economic Statistics*, vol. 9, n. 2, 339-348.
- PFEFFERMANN, D., SILVA, P.L.N. e FREITAS, M.P.S., (2000).** Implicações do esquema de rotação da Pesquisa Mensal de Emprego do IBGE na qualidade das estimativas publicadas, Rio de Janeiro: IBGE.
- SAS Institute Inc., (1989).** SAS/IML software: usage & reference, version 6, first edition. Cary, NC.
- SAS Institute Inc., SAS/IML software: changes and enhancements through. Release 6.11.** Cary, NC.
- SILVA, D. B. N., 1996,** Modelling compositional time series from repeated surveys, Tese (Doutorado), University of Southamoton, Faculty of Mathematical Studies, 237 p.
- SILVA, D.B.N.; CRUZ, M. M., (2002).** Séries temporais de pesquisas amostrais periódicas, 15º SINAPE, 141 p.
- SILVA, P.L.N., PESSOA, D.G.C., ALBIERI, S., FREITAS, M.P.S. e BIANCHINI, Z.M., (1998).** Aspectos sobre a estrutura longitudinal no contexto da Pesquisa sobre Mercado de Trabalho (versão preliminar), Rio de Janeiro: IBGE.
- SMITH, T. M. F., (1978).** Principles and problems in the analysis of repeated surveys, *In Survey sampling and Measurement*, 201-216 Eds. N. K. namboodini. Academic Press.
- WEI, W. W. S., (1993).** Time series analysis: univariate and multivariate methods. Redwood city: Addison-Wesley, 478 p.

Anexo A

Gráfico A1 - Gráfico de autocorrelação dos resíduos

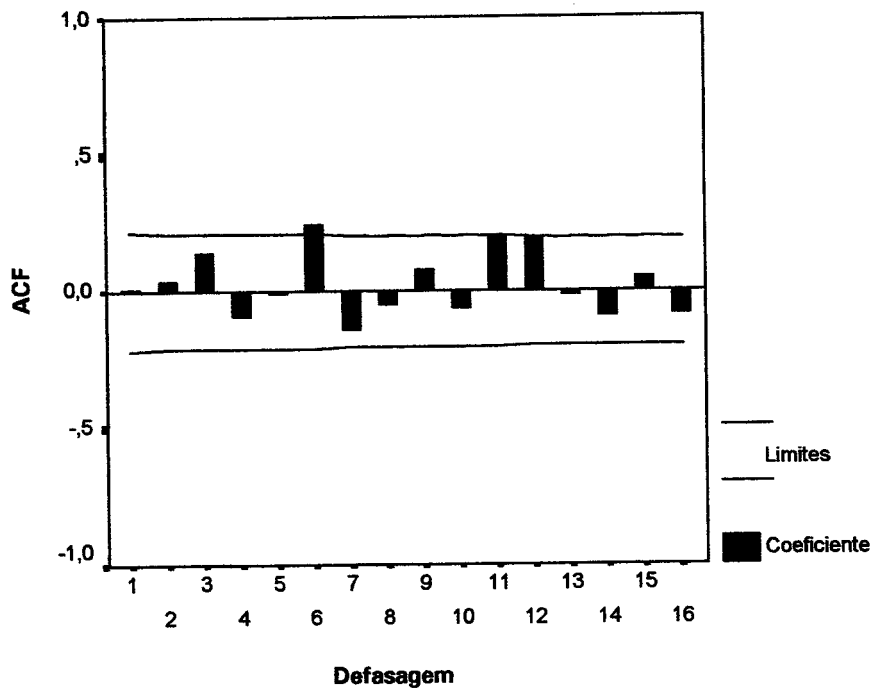


Gráfico A2 - Gráfico de autocorrelações parciais dos resíduos

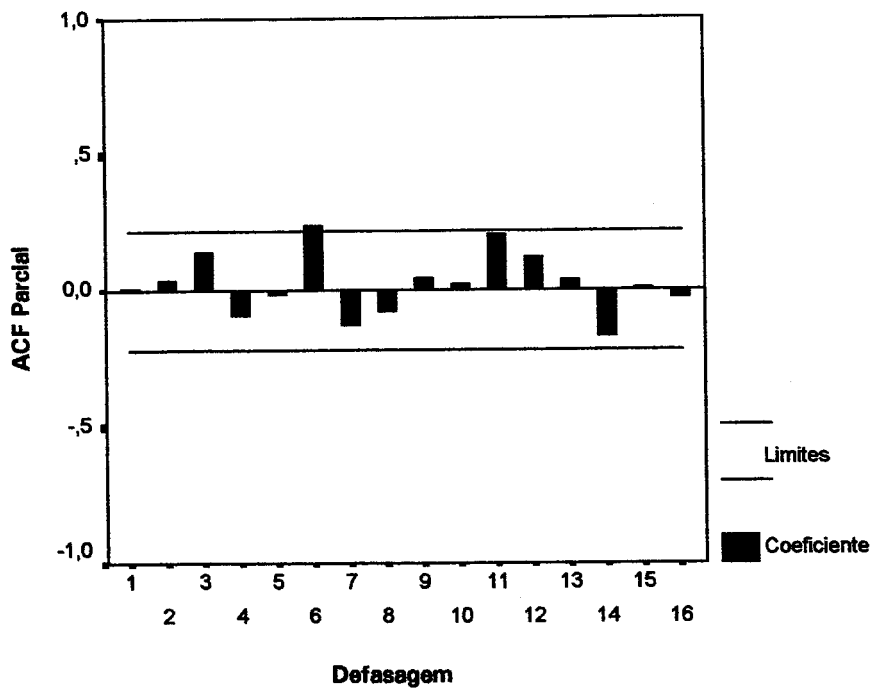


Gráfico A3 - Gráfico de normalidade dos resíduos

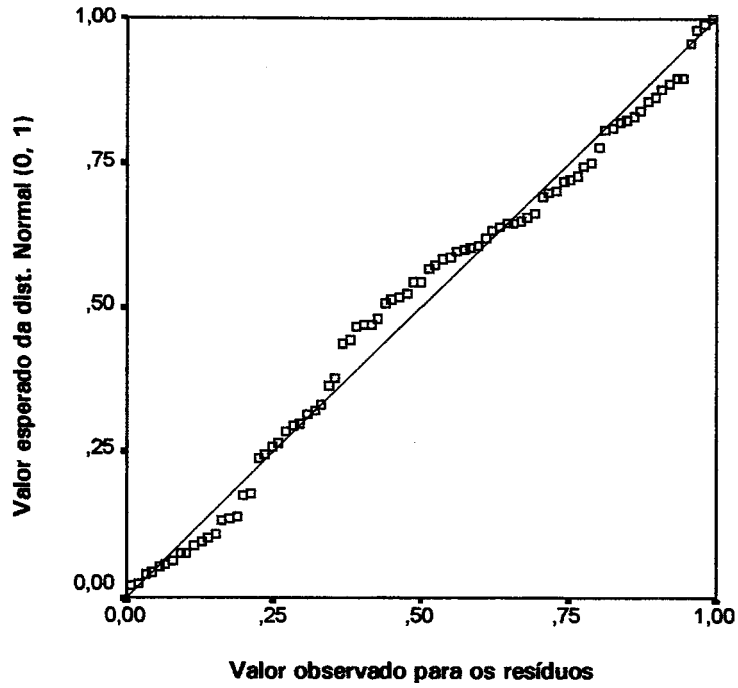


Gráfico A4 - Gráfico dos resíduos ao longo do tempo

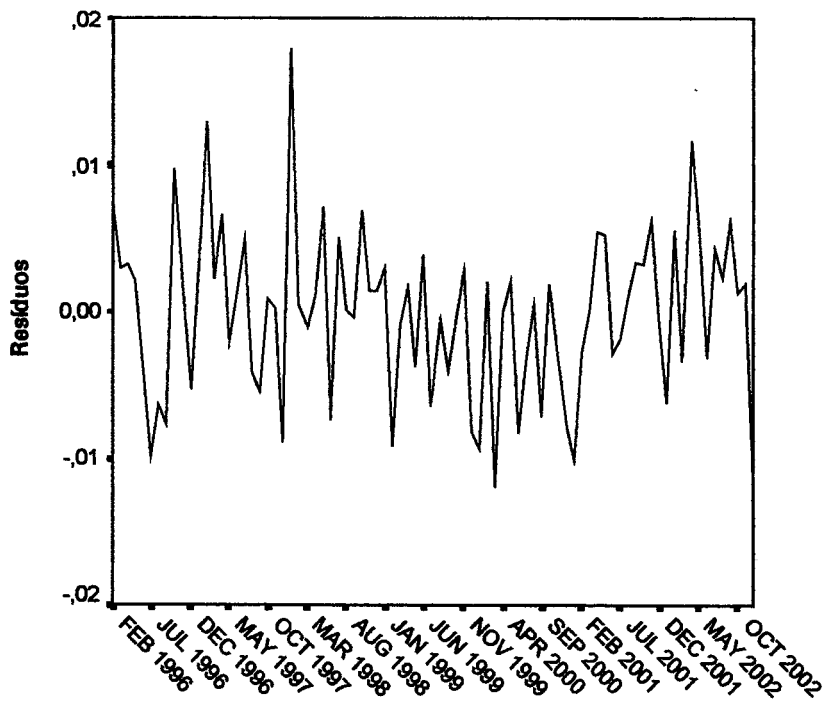


Tabela A1 - Testes dos resíduos

Teste	Estatística	p-valor
Jarq.Bera	5,786	0,055
Box-Ljung	11,276	0,257
H(27)	0,927	0,578

REVISTA BRASILEIRA DE ESTATÍSTICA - RBEs

POLÍTICA EDITORIAL

A Revista Brasileira de Estatística - RBEs publica trabalhos relevantes em Estatística Aplicada, não havendo limitação no assunto ou matéria em questão. Como exemplos de áreas de aplicação citamos as áreas de advocacia, ciências físicas e biomédicas, criminologia, demografia, economia, educação, estatísticas governamentais, finanças, indústria, medicina, meio ambiente, negócios, políticas públicas, psicologia e sociologia, entre outras. A RBEs publicará, também, artigos abordando os diversos aspectos de metodologias relevantes para usuários e produtores de estatísticas públicas, incluindo planejamento, avaliação e mensuração de erros em censos e pesquisas, novos desenvolvimentos em metodologia de pesquisa, amostragem e estimação, imputação de dados, disseminação e confiabilidade de dados, uso e combinação de fontes alternativas de informação e integração de dados, métodos e modelos demográficos e econométricos. Os artigos submetidos devem ser inéditos e não devem ter sido submetidos simultaneamente a qualquer outro periódico.

O periódico tem como objetivo a apresentação de artigos que permitam fácil assimilação por membros da comunidade em geral. Os artigos devem incluir aplicações práticas como assunto central, com análises estatísticas exaustivas e apresentadas de forma didática. Entretanto, o emprego de métodos inovadores, apesar de ser incentivado, não é essencial para a publicação.

Artigos contendo exposição metodológica são também incentivados, desde que sejam relevantes para a área de aplicação pela qual os mesmos foram motivados, auxiliem na compreensão do problema e contenham interpretação clara das expressões algébricas apresentadas.

A RBEs tem periodicidade semestral e também publica artigos convidados e resenhas de livros, bem como incentiva a submissão de artigos voltados para a educação estatística.

Artigos em espanhol ou inglês só serão publicados caso nenhum dos autores seja brasileiro e nem resida no País.

Todos os artigos submetidos são avaliados quanto à qualidade e à relevância por dois especialistas indicados pelo Comitê Editorial da RBEs.

O processo de avaliação dos artigos submetidos é do tipo 'duplo cego', isto é, os artigos são avaliados sem identificação de autoria e os comentários dos avaliadores também são repassados aos autores sem identificação.

INSTRUÇÃO PARA SUBMISSÃO DE ARTIGOS À RBEs

O processo editorial da RBEs é eletrônico. Os artigos devem ser submetidos via e-mail para: hortega@ibge.gov.br

Após a submissão o autor correspondente receberá um código para acompanhar o processo de avaliação do artigo. Caso não receba um aviso com este número no prazo de uma semana, fazer contato com a secretaria da revista no endereço:

Revista Brasileira de Estatística

IBGE – Diretoria de Pesquisas - Coordenação de Métodos e Qualidade

Av. República do Chile, nº 500, 10º andar

Centro, Rio de Janeiro – RJ

CEP: 20031-170

Tel.: 55 21 2142-0472

55 21 2142-4549

Fax: 55 21 2142-4802

INSTRUÇÕES PARA PREPARO DOS ORIGINAIS

Os originais entregues para publicação devem obedecer às normas seguintes.

1. Originais processados pelo editor de textos Word for Windows são preferidos. Entretanto, serão aceitos também, originais processados em LaTeX desde que sejam encaminhados acompanhados de versões em pdf, conforme descrito no item 3 a seguir;
2. A primeira página do original (folha de rosto) deve conter o título do artigo, seguido do(s) nome(s) completo(s) do(s) autor(es), indicando-se, para cada um, a afiliação e endereço para correspondência. Agradecimentos a colaboradores e instituições, e auxílios recebidos, também, devem figurar nesta página;
3. No caso da submissão não ser em Word for Windows, três arquivos do original devem ser enviados. O primeiro deve conter os originais no processador de texto utilizado (por exemplo, Latex). O segundo e terceiro devem ser no formato pdf, sendo um com a primeira página, como descrito no item 2, e outro contendo apenas o título, sem identificação do(s) autor(es) ou outros elementos que possam permitir a identificação da autoria;
4. A segunda página do original deve conter resumos em português e inglês (abstract), destacando os pontos relevantes do artigo. Cada resumo deve ser digitado seguindo o mesmo padrão do restante do texto, em um único parágrafo, sem fórmulas, com, no máximo, 150 palavras;
5. O artigo deve ser dividido em seções, numeradas progressivamente, com títulos concisos e apropriados. Todas as seções e subseções devem ser numeradas e receber título apropriado;

6. Tratamentos algébricos exaustivos devem ser evitados ou alocados em apêndices;
7. A citação de referências no texto e a listagem final de referências devem ser feitas de acordo com as normas da ABNT;
8. As tabelas e gráficos devem ser precedidos de títulos que permitam perfeita identificação do conteúdo. Devem ser numeradas seqüencialmente (Tabela 1, Figura 3, etc.) e referidas nos locais de inserção pelos respectivos números. Quando houver tabelas e demonstrações extensas ou outros elementos de suporte, podem ser empregados apêndices. Os apêndices devem ter título e numeração, tais como as demais seções de trabalho; e
9. Gráficos e diagramas para publicação devem ser incluídos nos arquivos com os originais do artigo. Caso tenham que ser enviados em separado, devem ter nomes que facilitem a sua identificação e posicionamento correto no artigo (ex. Gráfico 1; Figura 3; etc.). É fundamental que não existam erros, quer no desenho, quer nas legendas ou títulos.