

**Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística / Pós-Graduação**

**Análise da série histórica da Mortalidade por Diabetes no Brasil, 1996-2011:
comparação dos modelos de Espaço de Estados e de Suavização Exponencial**

Lenice Harumi Ishitani

Orientadora: Profa. Glaura da
Conceição Franco

**Belo Horizonte
Março / 2016**

Lenice Harumi Ishitani

**Análise da série histórica da Mortalidade por Diabetes no Brasil, 1996-2011:
comparação dos modelos de Espaço de Estados e de Suavização Exponencial**

Trabalho de conclusão de curso
apresentada ao Departamento de
Estatística do Instituto de Ciências
Exatas da UFMG como requisito
parcial para obtenção do título de
Especialista em Estatística

Orientadora: Profa. Glaura da
Conceição Franco

**Belo Horizonte
Março / 2016**

Agradecimentos

À Professora Glaura da Conceição Franco, pela paciência, compreensão e que nunca mediu esforços para me atender nas minhas dúvidas mais simples; não somente pelo TCC, mas pelas outras múltiplas demandas que trago para encontrar solução às desafiantes séries temporais da saúde. Foi com sua indicação que busquei e frequentei a Especialização.

Aos Professores do Curso, Gregório, Edna, Ilka, Marcelo, Sueli, Ela Mercedes, que me ensinaram a admirar ainda mais a Estatística, frente às possibilidades que esta ciência permite.

Às secretárias do curso, sempre disponíveis em atender às minhas solicitações.

Aos colegas de turma, companheiros de aprendizado e, por que não de angústias também, pela convivência nos 18 meses.

À Professora Elisabeth França que, gentilmente, cedeu os dados corrigidos que possibilitou esta análise.

Ao Gustavo Lana pelo auxílio nos comandos do R.

À minha Família, meus amados pais, esposo, filhas, irmãs, cunhados, sobrinhos, meus eternos amores, a quem devo todas as minhas alegrias e conquistas.

A Deus, que me possibilita conhecer e conviver com todas essas pessoas e vislumbrar as várias ciências e conhecimentos.

Resumo

O Diabetes Mellitus (DM) é uma doença metabólica, de etiologia diversa, considerado um grave problema de saúde pública, tendo em vista sua alta prevalência e mortalidade precoces. Apesar de sua importância, não existem séries históricas de incidência e/ou prevalência da doença em nosso país. Métodos estatísticos de previsão de séries temporais baseiam-se na ideia de que observações passadas da série contêm informações sobre o seu padrão de comportamento no futuro. Esses métodos consistem em identificar o padrão da série, separando-o do ruído (erro) contido nas observações individuais, para utilizá-lo na previsão de valores futuros da série. Uma grande classe de métodos de previsão, que tenta tratar as causas de flutuações em séries de tempo, é a de Suavização Exponencial. Nessa técnica é realizada ponderação dos valores da série, geralmente dando peso maior às observações mais recentes. Muito utilizado é o Modelo de Espaço de Estados, que estabelece a relação do futuro do processo em relação ao passado, dado os estados presente e passado. Considerando a importância do diabetes no país e a inexistência de modelagem e previsão para séries temporais desta doença metabólica, este trabalho se propõe a avaliar a tendência da mortalidade por diabetes no Brasil, por meio dos modelos de Alisamento Exponencial de Holt e Espaço de Estados, realizando uma comparação entre estes modelos com relação a previsão de valores futuros. As séries históricas de taxas de mortalidade padronizadas por diabetes analisadas foram a de adultos de 30 a 69 anos, no período de 1996 a 2011, no país e por sexo. Para validação dos modelos foram considerados os três últimos anos da série (2009 a 2011). As séries das taxas de mortalidade por diabetes mostraram decréscimo. No alisamento exponencial (AE) a tendência de decréscimo permaneceu constante ao longo do período para todas as séries. Não houve bom ajuste do AE para o sexo masculino. No modelo espaço de estados (MEE), a tendência de decréscimo permaneceu praticamente constante no período avaliado para as séries do Brasil e sexo masculino. A análise de resíduos para os modelos ajustados às três séries mostrou que esses atendem às suposições de variância constante, independência e normalidade. Não houve grande variabilidade e não houve autocorrelação de primeiro grau em ambos os modelos. O teste de Shapiro-Wilk demonstrou a normalidade da distribuição dos resíduos nos modelos ajustados às três séries, em ambos modelos (AE e MEE). No ajuste das previsões, observou-se que valores reais do período 2009-2011 estão contidos no intervalo de confiança de 95% das previsões tanto do alisamento exponencial (AE) quanto do modelo espaço de estados (MEE) nas séries do país e sexo feminino. No sexo masculino, os valores reais estão contidos no intervalo de confiança das previsões do modelo espaço de estados (MEE); no entanto, no AE, não houve ajuste adequado das previsões. A comparação entre os dois métodos mostrou que, para as séries de taxas de mortalidade por diabetes no Brasil, no período de 1996 a 2008, o Modelo Espaço de Estados apresentou melhor previsão que o Alisamento Exponencial. O declínio das taxas de morte por diabetes e a previsão de continuidade dessa redução são um achado positivo. No entanto, diante da epidemia de excesso de peso e obesidade, consumo não saudável de alimento e inatividade física, verifica-se um aumento da prevalência do diabetes. Tal quadro poderá reverter o declínio da mortalidade por diabetes

se as ações para o enfrentamento dessa situação não forem efetivamente implantadas por meio de ações de promoção de saúde.

Palavras-chave: tendência temporal, suavização exponencial, modelo espaço de estados, epidemiologia, diabetes, mortalidade

Sumário

1 – Introdução	8
2 - Metodologias	10
3 - Análise das séries de mortalidade	16
4 - Conclusão	28
5 - Referências Bibliográficas	30
6 - Apêndice	32

Lista de figuras

Figura 1 Taxas corrigidas de mortalidade por diabetes, em adultos de 30 a 69 anos, por sexo. Brasil, 2006-2011	18
Figura 2 Taxa corrigida de mortalidade por diabetes e valores previstos, segundo Alisamento Exponencial. Brasil, 1996-2008	21
Figura 3 Taxa corrigida de mortalidade por diabetes e valores previstos, segundo Modelo Espaço de Estados. Brasil, 1996-2008	22
Figura 4 Taxa corrigida de mortalidade por diabetes e valores previstos, no sexo masculino, segundo Alisamento Exponencial. Brasil, 1996-2008	23
Figura 5 Taxa corrigida de mortalidade por diabetes e valores previstos, no sexo masculino, segundo Modelo Espaço de Estados. Brasil, 1996-2008	24
Figura 6 Taxa corrigida de mortalidade por diabetes e valores previstos, no sexo feminino, segundo Alisamento Exponencial. Brasil, 1996-2008	25
Figura 7 Taxa corrigida de mortalidade por diabetes e valores previstos, no sexo feminino, segundo Modelo Espaço de Estados. Brasil, 1996-2008	26

Lista de tabelas

Tabela 1 Ajuste dos modelos de Alisamento Exponencial e Espaço de Estados para a série de mortalidade por diabetes, no Brasil e por sexo	19
Tabela 2 Teste de Shapiro-Wilk de resíduos dos modelos ajustados para as séries de diabetes, por sexo. Brasil, 1996-2008	19
Tabela 3 Evolução da tendência no Modelo de Espaço de Estados	20
Tabela 4 Valores reais e previsões pontuais e intervalares para o modelo Alisamento Exponencial e Espaço de Estados. Taxas de mortalidade, Brasil, 2009-2011	27
Tabela 5 Estatística do erro quadrático médio da previsão um passo à frente dos modelos de previsão	28

1 – Introdução

O Diabetes Mellitus (DM) é uma doença metabólica, de etiologia diversa, considerado um dos graves problemas de saúde pública, tendo em vista sua alta prevalência e mortalidade precoces. Além das complicações que geram custos no controle e tratamento, está associado ao maior risco de vida, principalmente, por alterações cardiovasculares e renais.

Lozano et al. (2012) verificaram um aumento de 19,7%, no período de 2000 a 2010, na taxa de mortalidade mundial padronizada. Nas projeções realizadas por Mathers & Loncar (2006), o DM pode se tornar a 7^a. causa de óbito no mundo, em 2030.

O aumento tem sido atribuído a fatores como modificações do padrão alimentar, aumento do sedentarismo e obesidade, bem como à urbanização e ao envelhecimento da população (WHO, 2014).

Apesar de sua importância, não existem séries históricas de incidência e/ou prevalência da doença em nosso país. Dessa forma, as informações de mortalidade passam a ser um importante indicador para o seu monitoramento (Mattos et al., 2012). Cesse et al. (2009) e Mattos et al. (2012) verificaram um aumento na mortalidade nas capitais brasileiras, mas não descartam a possibilidade de esse incremento ser devido à melhoria da qualidade da informação ao longo do tempo. Barreto et al. (2007) também verificaram aumento nas capitais das regiões sul e sudeste.

Esses estudos ressaltam a importância das análises de séries temporais. Segundo Latorre & Cardoso (2001) uma série temporal ou série histórica pode ser definida como uma sequência de dados obtidos em intervalos de tempo regulares durante um período específico, como por exemplo, a taxa de mortalidade por diabetes em adultos.

Métodos estatísticos de previsão de séries temporais baseiam-se na ideia de que observações passadas da série contêm informações sobre o seu padrão de comportamento no futuro. Esses métodos consistem em identificar o padrão da série, separando-o do ruído (erro) contido nas observações individuais, para utilizá-lo na previsão de valores futuros da série (Morettin & Toloí, 2006).

Uma grande classe de métodos de previsão, que tenta tratar as causas de flutuações em séries de tempo, é a da suavização exponencial. Nessa técnica é realizada ponderação dos valores da série, geralmente dando peso maior às observações mais recentes (Holt, 1957). A grande vantagem desse método está relacionada à sua simplicidade, eficiência

computacional e à sua razoável precisão, além de ser aplicável a séries com pequeno número de observações (Morettin & Toloi, 2006).

Também muito utilizado (Hemming & Shaw (2002), Chiogna & Gaetan (2002), Franco et al. (2008), Franco et al. (2009) e Alves et al. (2012)), o modelo de espaço de estados (Harvey, 1989) estabelece a relação do futuro do processo em relação ao passado, dado os estados latentes, presente e passado. O estado do processo condensa todas as informações do passado necessárias para a predição do futuro.

Considerando a importância do diabetes no país e a inexistência de modelagem e previsão para séries temporais desta doença metabólica, propõe-se avaliar a tendência da mortalidade por diabetes no Brasil, no período de 1996-2011.

Pretende-se, dessa forma:

- a) Avaliar o comportamento da série histórica para previsão, por meio do modelo de Alisamento Exponencial de Holt;
- b) Analisar o comportamento da série histórica para previsão, por meio do modelo de Espaço de Estados;
- c) Comparar os modelos para previsão.

Na Seção 2 são apresentados os modelos estatísticos utilizados neste trabalho e na Seção 3 é apresentada a aplicação a dados reais.

2. Metodologias

Seja Y_t , $t=1, \dots, N$ uma série temporal. Se Y_t apresenta apenas componente de tendência (μ_t), a série temporal pode ser modelada por $Y_t = \mu_t + \varepsilon_t$, onde ε_t é uma série ruído branco, que possui média e variância constantes.

Descreveremos abaixo os dois modelos que serão utilizados neste trabalho para fazer previsões para valores futuros da série, a saber Suavização Exponencial de Holt e o modelo de Espaço de Estados.

2.1 Suavização Exponencial de Holt (duplo Holt)

Em séries temporais anuais como a da mortalidade por diabetes, que apresentam tendência, mas não sazonalidade, é adequada a utilização da Suavização Exponencial de Holt (Holt, 1957). Nesse caso, além da suavização do nível, uma nova constante é utilizada para modelar a tendência da série.

Os valores do nível, Z_t , e da inclinação da tendência da série, T_t , no instante t , serão estimados por:

$$\begin{aligned}\bar{Z}_t &= AY_t + (1 - A)(\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1 \quad e \quad t = 2, \dots, N, \\ \hat{T}_t &= C(\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C)\hat{T}_{t-1}, \quad 0 < C < 1 \quad e \quad t = 2, \dots, N.\end{aligned}$$

O cálculo das constantes A e C é feito pela minimização da soma dos erros quadráticos médios de previsão.

A previsão é calculada adicionando-se ao valor básico (\bar{Z}_t) a tendência multiplicada pelo número de passos à frente que se deseja prever (h),

$$\hat{Y}_t(h) = \bar{Z}_t + h\hat{T}_t, \quad \forall h \geq 1.$$

Supondo $\varepsilon_t \sim N(0, \sigma_\alpha^2)$, podemos construir o intervalo de previsão de $100(1-\gamma)\%$ para o valor futuro, Y_{T+h} , como:

$$(\hat{Y}_T(h) - z_{\kappa/2} d_h \Delta(T)) \quad ; \quad (\hat{Y}_T(h) + z_{\kappa/2} d_h \Delta(T))$$

onde $z_{K/2}$ é o percentil $K/2$ da distribuição Gaussiana, conforme Bowerman & O'Connell (1993), e é dado por:

$$d_h = 1,25 \left[\frac{1 + \frac{\theta}{(1+\nu)^3} \left[(1 + 4\nu + 5\nu^2) + 2\theta(1 + 3\nu)h + 2\theta^2 h^2 \right]}{1 + \frac{\theta}{(1+\nu)^3} \left[(1 + 4\nu + 5\nu^2) + 2\theta(1 + 3\nu) + 2\theta^2 \right]} \right]^{1/2};$$

θ é igual ao máximo entre A e C (as constantes de suavização);

$$\nu = 1 - \theta;$$

e

$$\Delta(T) = \frac{\sum_{t=1}^N |Y_t - [\bar{z}_{t-1} + \hat{T}_{t-1}]|}{N}.$$

2.2 Modelo de Espaço de Estados

Em modelos paramétricos, a análise é feita no domínio do tempo. Uma maneira de se modelar séries temporais é através da decomposição em seus componentes não-observáveis, por meio dos modelos estruturais (Harvey, 1989). Estes modelos consideram que uma série temporal é formada por componentes não observáveis, como por exemplo nível, tendência, sazonalidade, etc. No entanto, esta técnica foi inicialmente menosprezada devido à dificuldade computacional envolvida na implementação do filtro de Kalman.

A natureza determinística das componentes de nível, tendência e sazonalidade é bastante infrequente do ponto de vista prático. Uma maneira natural de contornar esse problema é permitir uma variabilidade nessas componentes, considerando-se os modelos estruturais ou modelos de espaço de estados, introduzidos por Harvey (1989).

Modelos estruturais consideram as observações de uma série temporal como sendo uma combinação linear de um nível e uma componente irregular, que pode representar tendências fixas ou aleatórias, além de periodicidades.

Quando a série tem um movimento crescente ou decrescente e também tem um comportamento de passeio aleatório, atende pelo nome de modelo de tendência linear local (MTL). No modelo de tendência local, a especificação da tendência é baseada em informações a priori da série e/ou no gráfico das observações. É descrito pelas equações a seguir, sabendo-se que $t = 1, 2, \dots, N$:

$$Y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2),$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad \eta_t \stackrel{iid}{\sim} N(0, \sigma_\eta^2)$$

$$\beta_t = \beta_{t-1} + \xi_t, \quad \xi_t \stackrel{iid}{\sim} N(0, \sigma_\xi^2),$$

sendo ε_t , η_t e ξ_t distúrbios tipo ruído branco mutuamente não-correlacionados.

Componentes não observáveis do modelo estrutural, para uma série temporal univariada de tendência local (MTL) são definidos por meio das duas últimas equações, acima descritos.

2.2.1 Forma de Espaço de Estados

A estimação dos modelos estruturais ocorre através da estimação das variâncias dos ruídos relacionados aos componentes não-observáveis, que são conhecidas como hiperparâmetros. O processo de estimação dos hiperparâmetros torna-se um pouco mais simples, reduzindo-se o número de equações, se a forma de espaço de estados é utilizada. Dessa forma, é possível representar a modelagem estrutural através de duas equações:

a equação das observações (ou de medida):

$$Y_t = \mathbf{z}_t' \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, h_t),$$

e a equação do estado (ou transição),

$$\alpha_t = \mathbf{T}_t \alpha_{t-1} + \mathbf{R}_t \eta_t, \quad \eta_t \sim N(0, \mathbf{Q}_t),$$

sendo $t = 1, 2, \dots, N$, ε_t , ruídos não-correlacionados, η_t vetor de ruídos serialmente não-correlacionados cuja matriz de covariância é dada por \mathbf{Q}_t e α_t é o vetor de estados. Sabe-se também que ε_t e η_t são independentes entre si, $\mathbf{z}_t, \mathbf{T}_t$ e \mathbf{R}_t são chamados matrizes do sistema e são conhecidas após a definição da modelagem a ser aplicada à série.

2.2.2 Filtro de Kalman

Para fazer previsões em um modelo na forma de espaço de estados a partir de observações passadas até o tempo t , é necessário que o vetor de estado α_t seja atualizado. Como o vetor de estado α_t é não observável, é necessário um algoritmo recursivo para estimá-lo. O filtro de Kalman (FK) (Kalman, 1960) é um algoritmo recursivo que determina a estimativa do vetor de estados no tempo t dada toda a informação disponível até o instante $t-1$.

O FK decompõe a série através de equações recursivas que atualizam sequencialmente o vetor de estado. Suas características básicas são atualização, previsão e suavização.

Sabendo-se que Y_{t-1} é o vetor das observações até o instante $t-1$, $E(\alpha_0) = \mathbf{a}_0$ e $Cov(\alpha_0, \alpha_0') = \mathbf{P}_0$, tem-se que:

$$(i) (\alpha_t | Y_{t-1}) \sim N(\mathbf{a}_{t|t-1}, \mathbf{P}_{t|t-1}), \text{ sendo } \mathbf{a}_{t|t-1} = E(\alpha_t | Y_{t-1}) = \mathbf{T}_t \mathbf{a}_{t-1},$$

$$(ii) \mathbf{P}_{t|t-1} = Var(\alpha_t | Y_{t-1}) = \mathbf{T}_t \mathbf{P}_{t-1} \mathbf{T}_t' + \mathbf{R}_t \mathbf{Q}_t \mathbf{R}_t',$$

$$(iii) (y_t | Y_{t-1}) \sim N(\tilde{y}_{t|t-1}, F_t), \text{ sendo } \tilde{y}_{t|t-1} = E(y_t | Y_{t-1}) = \mathbf{z}_t' \mathbf{a}_{t|t-1},$$

$$(iv) F_t = Var(y_t | Y_{t-1}) = \mathbf{z}_t' \mathbf{P}_{t|t-1} \mathbf{z}_t + h_t.$$

Então, pode-se notar que a média e a matriz de covariância do vetor de estado no tempo t , dado toda a informação até o tempo $t-1$, são respectivamente iguais a $\mathbf{a}_{t|t-1}$ e $\mathbf{P}_{t|t-1}$. Através de recursividade e de propriedades da distribuição Normal multivariada, encontram-se os seguintes resultados importantes:

$$\mathbf{a}_t = \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{z}_t' F_t^{-1} (y_t - \tilde{y}_{t|t-1}),$$

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \cdot \mathbf{z}_t \cdot F_t^{-1} \cdot \mathbf{z}_t' \cdot \mathbf{P}_{t|t-1}.$$

As duas equações acima são chamadas de equações de atualização, pois através delas é possível atualizar os estimadores $\mathbf{a}_{t|t-1}$ e $\mathbf{P}_{t|t-1}$ do vetor de estado para o instante t . Para facilitar o manuseio das fórmulas, chama-se de $v_t = y_t - \tilde{y}_{t|t-1}$ o erro de previsão ou inovações, sendo que $E(v_t) = 0$ e $Var(v_t) = F_t$, e chama-se de \mathbf{K}_t a matriz de ganho de Kalman. Assim, tem-se que:

$$v_t = y_t - \tilde{y}_{t|t-1} \Rightarrow v_t = y_t - \mathbf{z}_t' \mathbf{a}_{t|t-1}, \quad \mathbf{K}_t = \mathbf{T}_{t+1} \cdot \mathbf{P}_{t|t-1} \cdot \mathbf{z}_t \cdot F_t^{-1}$$

Para a inicialização do filtro de Kalman toma-se $\mathbf{a}_0 = 0$ ou $\mathbf{a}_0 = E(Y_t)$ e \mathbf{P}_0 suficientemente grande (Harvey, 1990).

2.2.3 Estimação por máxima verossimilhança

O método de máxima verossimilhança pode ser usado para se estimar os hiperparâmetros de um modelo através da maximização da função densidade conjunta $f(y_1, \dots, y_N; \psi)$ em relação a $\psi = (\sigma_\varepsilon^2, \sigma_N^2, \sigma_\xi^2)$ que é o vetor de hiperparâmetros.

Usando-se o modelo na forma de espaço de estados, a função de verossimilhança é obtida através do erro de previsão v_t e sua variância F_t , obtidos do FK, supondo que $(y_t | Y_{t-1}) \sim N(\tilde{y}_{t|t-1}, F_t)$. Então, a função de verossimilhança será descrita por:

$$L(\psi) = \prod_{t=1}^N (2\pi)^{-1/2} |F_t|^{-1/2} \exp\left\{-\frac{1}{2} v_t' F_t^{-1} v_t\right\}.$$

Aplicando-se o logaritmo natural para simplificar os cálculos:

$$\log L(\psi) = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^N \log |F_i| - \frac{1}{2} \sum_{i=1}^N v_i' F_i^{-1} v_i.$$

Para estimar o vetor de hiperparâmetros, deve-se então encontrar um valor para ψ tal que maximize a verossimilhança. Pelo fato de a função de verossimilhança ser uma função complicada do vetor de hiperparâmetros, esta estimação não é realizada analiticamente, mas via métodos numéricos.

2.2.4 Previsão

A previsão em modelos de espaço de estados é obtida pela seguinte equação:

$$\hat{Y}_t(h) = z_{t+h}' a_{t+h}, \text{ onde } a_{t+h}, \text{ é obtido do filtro de Kalman.}$$

Assumindo que os erros de previsão têm distribuição gaussiana, o intervalo de previsão de $100(1-K)\%$ para o valor futuro, Y_{t+h} , é dado por:

$$[\hat{Y}_t(h) - z_{1-K/2} \sqrt{F_{t+h}}];$$

Onde $z_{1-K/2}$ é o $(1 - (K/2))$ - percentil da distribuição normal e F_{t+h} é obtido do filtro de Kalman (Ribeiro (2006), Rodriguez & Ruiz (2009)).

3 – Análise das séries de mortalidade

As séries históricas anuais de Taxas de mortalidade por diabetes analisadas foram a de adultos de 30 a 69 anos, no país e por sexo, no período de 1996 a 2011.

Dados de mortalidade foram obtidos do Sistema de Informação sobre Mortalidade (SIM)/DATASUS disponíveis em http://tabnet.datasus.gov.br/cgi/sim/dados/cid10_indice.htm

Dados populacionais, para o cálculo das taxas, foram obtidos de <http://www2.datasus.gov.br/DATASUS/index.php?area=0206&VObj=http://tabnet.datasus.gov.br/cgi/deftohtm.exe?ibge/cnv/pop>

As taxas foram padronizadas, por idade e sexo, após correção de subregistro e de subnotificação de causas.

Para a modelagem foi utilizada a série de 1996 a 2008. As observações de 2009 a 2011 foram retiradas do ajuste com o propósito de comparar as previsões para estes anos. Foram utilizados os modelos de Suavização Exponencial de Holt e Espaço de Estados do tipo tendência local. Toda a análise foi realizada em linguagem R. No modelo de Suavização Exponencial foi utilizado o pacote HoltWinters. Para o modelo de Espaço de Estados, o Filtro de Kalman foi utilizado para construção da função de verossimilhança e a estimação dos parâmetros do modelo foi feita através do pacote StructTS.

Para validação dos modelos foram considerados os três últimos anos da série (2009 a 2011).

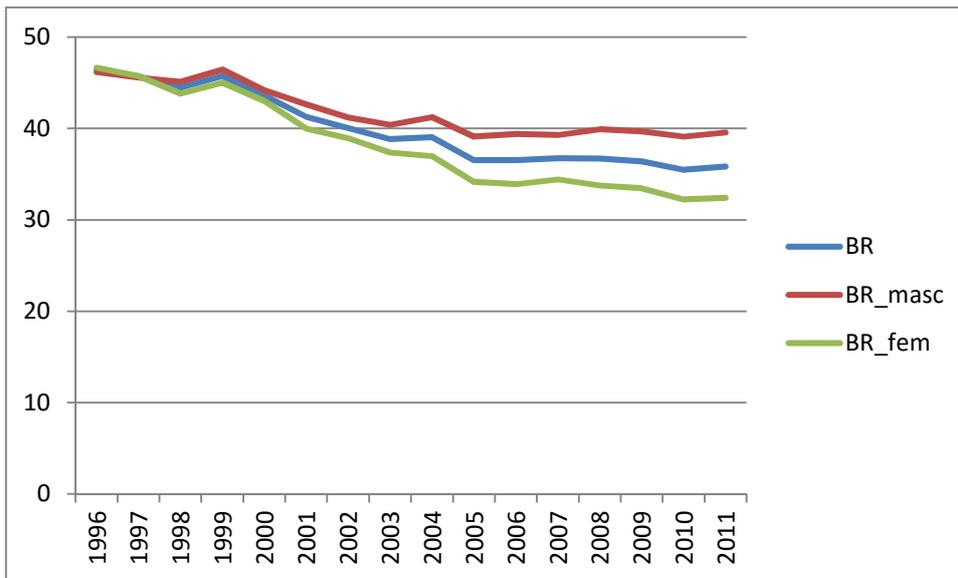
Os modelos foram comparados por meio da estatística do erro quadrático médio da previsão um passo à frente (EQMP):

$$EQMP = \sum_{t=N+1}^{N+3} (Y_t - \hat{Y}_t(1))^2.$$

3.1 – Ajuste dos modelos

As séries das taxas de mortalidade por diabetes mostraram um decréscimo, como pode ser visto na Figura 1.

Figura 1 - Taxas corrigidas de mortalidade por diabetes, em adultos de 30 a 69 anos, por sexo. Brasil, 2006-2011.



Fonte: SIM-MS

A seguir são apresentados resultados dos ajustes realizados para a série mortalidade por diabetes no Brasil, de 1996 a 2008, total e por sexo (ver Tabela 1).

O Alisamento Exponencial de Holt (AE) para a série do país e para o sexo feminino mostrou que pontos mais recentes tiveram maior peso no ajuste ($A = 0,75$ e $0,76$, respectivamente). Não houve bom ajuste do AE para o sexo masculino ($A=0,00$). Para todas as séries, a constante C foi igual a zero, indicando que a tendência de decréscimo permaneceu constante ao longo do período.

No modelo espaço de estados (MEE), a variância da componente de inclinação (σ_{ξ}^2) foi igual a zero para as séries do Brasil e Masculino, indicando também que a tendência de decréscimo permaneceu praticamente constante no período avaliado e, portanto, deve ser determinístico, o que não ocorreu no sexo feminino.

Tabela 1: Ajuste dos modelos de Alisamento Exponencial e Espaço de Estados para a série de mortalidade por diabetes, no Brasil e por sexo.

	Alisamento Exponencial		Espaço de Estados		
	A	C	σ_{η}^2	σ_{ε}^2	σ_{ε}^2
Brasil	0,74945	0,00000	1,2593	0,0000	0,0787
Masculino	0,00000	0,00000	0,9609	0,0000	0,1826
Feminino	0,76460	0,00000	0,0000	0,2646	0,6098

Fonte: SIM-MS

A análise de resíduos para os modelos ajustados às três séries (ver Apêndice) mostrou que os resíduos atendem às suposições de variância constante, independência e normalidade. Não houve grande variabilidade e não houve auto-correlação de primeiro grau em ambos os modelos.

O teste de Shapiro-Wilk demonstrou a normalidade da distribuição dos resíduos nos modelos ajustados às três séries, em ambos modelos (AE e MEE).

Tabela 2 - Teste de Shapiro-Wilk de resíduos dos modelos ajustados para as séries de diabetes, por sexo. Brasil, 1996-2008.

	AE		MEE	
	S-W	valor-p	S-W	valor-p
Brasil	0,9481	0,6202	0,9285	0,3258
Masculino	0,9504	0,6487	0,9422	0,4854
Feminino	0,9655	0,8376	0,9763	0,9563

Fonte: SIM-MS

No Modelo Espaço de Estados foi possível avaliar a evolução da tendência no tempo segundo o estimador da inclinação da curva em cada ano, em relação ao ano anterior (Tabela 3). Para o Brasil, a evolução da tendência no tempo mostrou um decréscimo em relação ao início da série. Ao final da série, a partir de 2006, observou-se uma redução do decréscimo.

No sexo masculino, verificou-se redução no início da série com um incremento atípico em 1999, mas depois a tendência de decréscimo aumentou até 2003. A partir de 2006, verificou-se que houve uma diminuição desse decréscimo.

No sexo feminino, houve uma importante diminuição em 1998. A partir de 1999, ocorreu aumento desse decréscimo e, a partir de 2006, houve uma diminuição dessa tendência.

Tabela 3: Evolução da tendência ($\hat{\beta}_t$) no Modelo Espaço de Estados

	Estimativas da inclinação		
	Brasil	Masculino	Feminino
1996	0,00000	0,00000	0,00000
1997	-0,15455	-0,12596	-0,29260
1998	-0,33205	-0,18750	-0,92178
1999	-0,12008	0,00439	-0,37788
2000	-0,35304	-0,23130	-0,79571
2001	-0,57272	-0,38468	-1,74619
2002	-0,64721	-0,49743	-1,78168
2003	-0,69892	-0,53443	-1,71135
2004	-0,63053	-0,43583	-1,21569
2005	-0,76487	-0,53975	-1,65254
2006	-0,71809	-0,50054	-1,26203
2007	-0,65670	-0,47363	-0,48917
2008	-0,61783	-0,40894	-0,31679

Fonte: SIM-MS

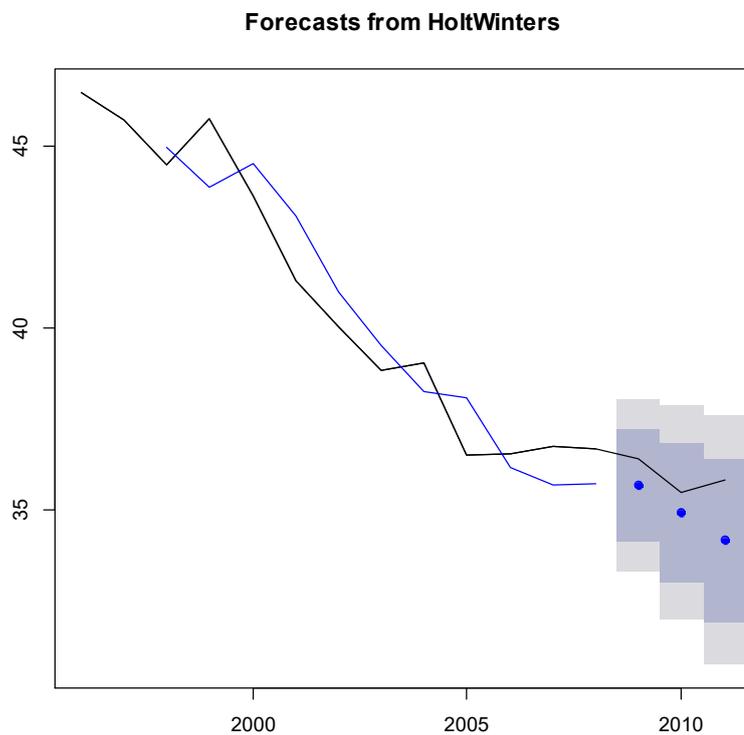
3.2 Previsões para o período 2009-2011

A seguir são apresentados os resultados de previsão um passo-à-frente para os três últimos anos da série, de 2009 a 2011.

3.2.1 Resultados da previsão para ambos os sexos (Brasil – total)

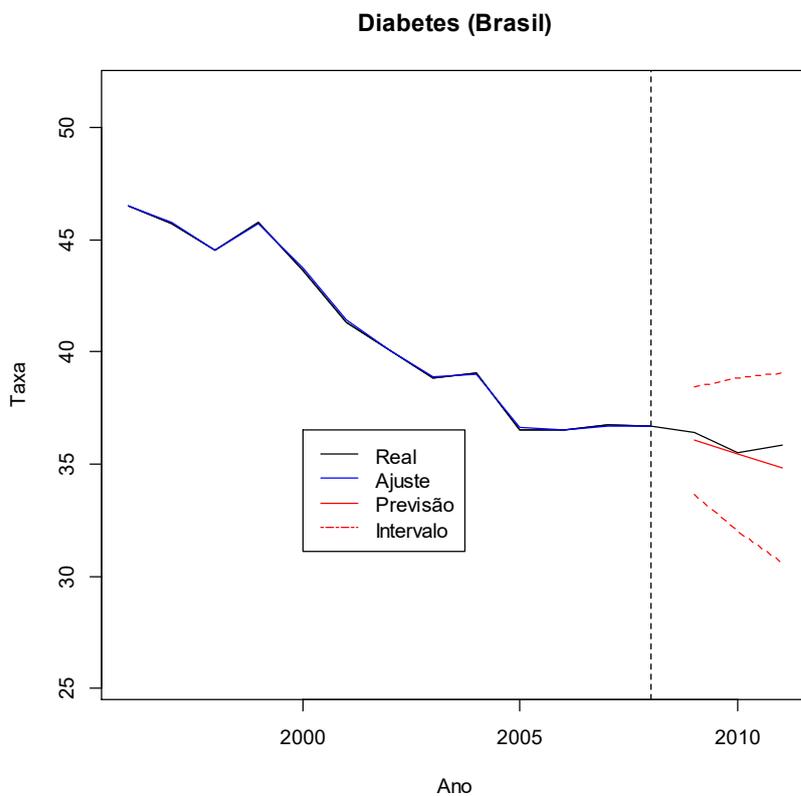
No ajuste das previsões, observou-se que os valores reais do período 2009-2011 estão contidos no intervalo de confiança de 95% das previsões tanto do alisamento exponencial (AE) quanto do modelo espaço de estados (MEE) (Figuras 2 e 3 e Tabela 4).

Figura 2 - Taxa corrigida de mortalidade por diabetes e valores previstos, segundo Alisamento Exponencial. Brasil, 1996-2008.



Fonte: SIM-MS

Figura 3 - Taxa corrigida de mortalidade por diabetes e valores previstos, segundo Modelo Espaço de Estados. Brasil, 1996-2008.

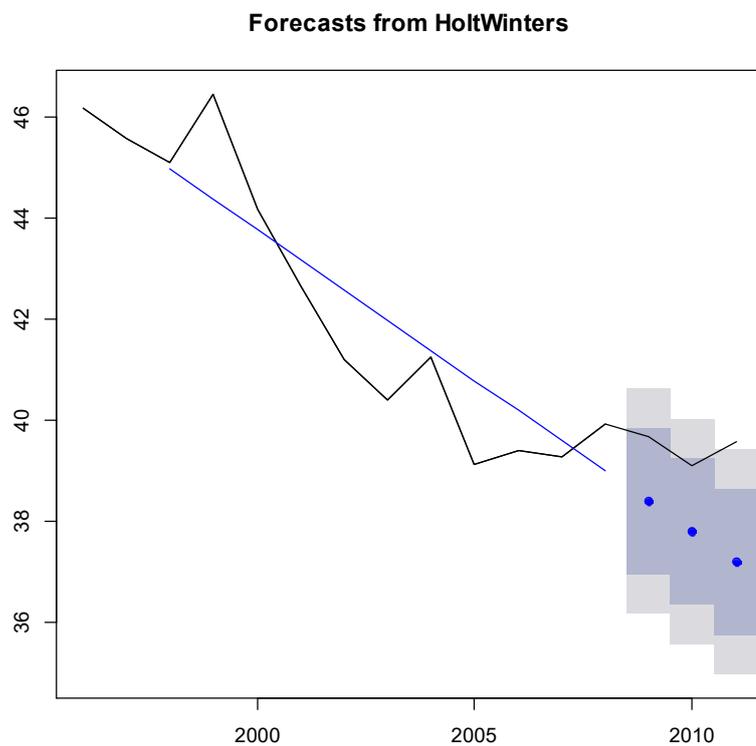


Fonte: SIM-MS

3.2.2 Resultados da previsão para o sexo masculino

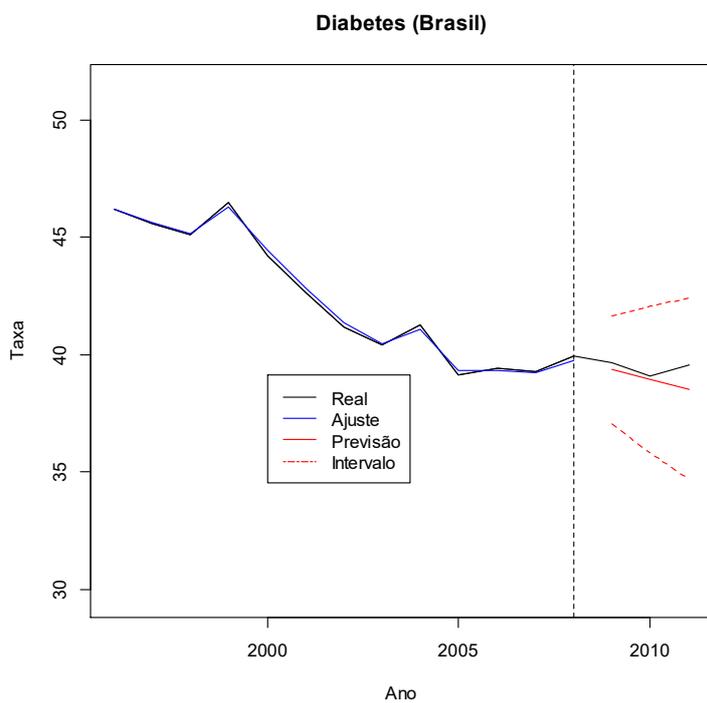
Observou-se que os valores reais do período 2009 a 2011, no sexo masculino, estão contidos no intervalo de confiança das previsões do modelo espaço de estados (MEE). No entanto, no AE, não houve ajuste adequado das previsões: observou-se que o ponto referente a 2011, encontra-se fora do intervalo previsto (Figuras 4 e 5 e Tabela 4).

Figura 4 - Taxa corrigida de mortalidade por diabetes e valores previstos, no sexo masculino, segundo Alisamento Exponencial. Brasil, 1996-2008.



Fonte: SIM-MS

Figura 5 - Taxa corrigida de mortalidade por diabetes e valores previstos, no sexo masculino, segundo Modelo Espaço de Estados. Brasil, 1996-2008.

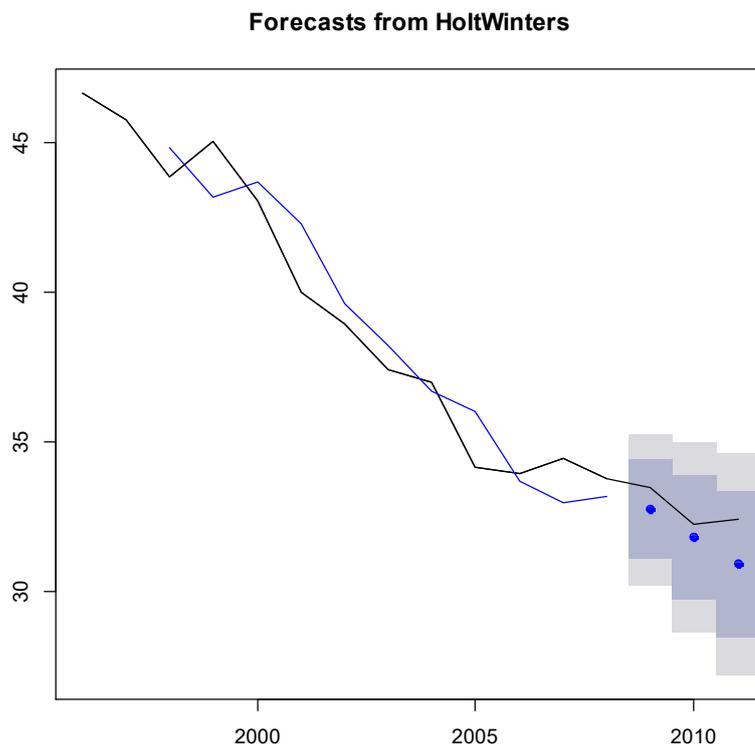


Fonte: SIM-MS

3.2.3 Resultados da previsão para o sexo feminino

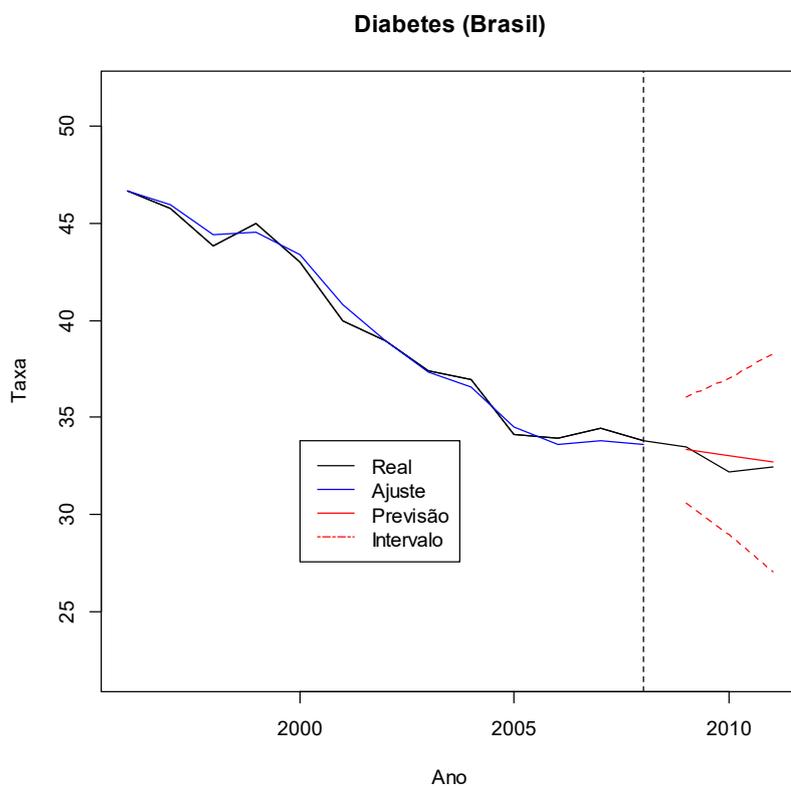
Observou-se que os valores reais estão contidos no intervalo de confiança das previsões tanto do alisamento exponencial (AE) quanto do modelo espaço de estados (MEE) (Figuras 6 e 7 e Tabela 4).

Figura 6 - Taxa corrigida de mortalidade por diabetes e valores previstos, no sexo feminino, segundo Alisamento Exponencial. Brasil, 1996-2008.



Fonte: SIM-MS

Figura 7 - Taxa corrigida de mortalidade por diabetes e valores previstos, no sexo feminino, segundo Modelo Espaço de Estados. Brasil, 1996-2008.



Fonte: SIM-MS

Apesar das séries mostrarem variações no declínio, as previsões mostraram tendência de decréscimo.

3.3 Comparação dos modelos através das previsões para o período 2009-2011

Os valores das previsões para os três anos subsequentes (2009 a 2011) encontram-se na Tabela 4. Observa-se que o valor real de 2011, para o sexo masculino, não está contido no Intervalo de Previsão de 95% do modelo de Alisamento Exponencial.

Tabela 4: Valores reais e previsões pontuais e intervalares para o modelo Alisamento Exponencial e Espaço de Estados. Taxas de mortalidade, Brasil, 2009-2011.

Ano	Valor Real	Previsão - AEH	Previsão - MEE
Total			
2009	36,400	35,702	36,049
IC 95%		[33,353 - 38,050]	[33,648-38,451]
2010	35,483	34,945	35,432
IC 95%		[32,010 - 37,880]	[32,032-38,831]
2011	35,816	34,189	34,814
IC 95%		[30,767 - 37,611]	[30,577-39,051]
Masculino			
2009	39,669	38,391	39,366
IC 95%		[36,164 - 40,619]	[37,061-41,671]
2010	39,108	37,792	38,957
IC 95%		[35,565 - 40,019]	[35,828-42,085]
2011	39,580	37,193	38,548
IC 95%		[34,965 - 39,420]	[34,710-42,386]
Feminino			
2009	33,457	32,728	33,315
IC 95%		[30,205 - 35,250]	[30,559-36,072]
2010	32,219	31,817	32,999
IC 95%		[28,641 -34,993]	[28,987-37,011]
2011	32,424	30,907	32,682
IC 95%		[27,191 - 34,622]	[27,071-38,293]

Fonte: SIM-MS

A Tabela 5 mostra o EQMP para os dois modelos. Verifica-se que o Modelo Espaço de Estados apresentou melhores previsões um passo-à-frente para o período 2009-2011 para as séries históricas do país e na análise por sexo.

Tabela 5 - Estatística do erro quadrático médio da previsão um passo à frente dos modelos de previsão.

	Alisamento exponencial	Espaço de Estados
Brasil	2,7314	0,6568
Masculino	8,2690	0,7358
Feminino	2,3389	0,2677

Fonte: SIM-MS

4 – Conclusão

O diabetes é, atualmente, importante problema de saúde pública devido às suas complicações e por ser fator de risco para as doenças crônicas não transmissíveis, principal causa de mortalidade no Brasil e no mundo (World Health Organization, 2014). Avaliar o perfil de mortalidade dessa causa, portanto, torna-se subsídio na definição de ações para o enfrentamento dessa situação epidemiológica.

Estudos referentes às taxas de mortalidade no Brasil encontraram resultados controversos. Entre os que encontraram aumento da mortalidade, foram trabalhados dados brutos sem correção (Cesse et al. (2009), Mattos et al. (2012)).

Neste estudo, foram analisados dados corrigidos de subregistro e causas de morte, para assegurar melhor qualidade do dado. Foram utilizados dois modelos para ajuste de previsão, Alisamento Exponencial e Modelo Espaço de Estados, a fim de se comparar os dois métodos.

No Alisamento Exponencial, verificou-se que todas as séries apresentaram decréscimo constante ao longo do período, apesar de não apresentar um bom ajuste para o sexo masculino. No modelo espaço de estados (MEE), a variância da componente de inclinação (σ_{ξ}^2) foi igual a zero para as séries Brasil e Masculino, indicando também que a tendência de decréscimo permaneceu praticamente constante no período avaliado. A tendência de queda encontrada corrobora estudos que também analisaram dados corrigidos (Alves et al. (2015), Malta et al. (2014)).

Observou-se que valores reais do período 2009-2011 estão contidos no intervalo de confiança de 95% das previsões tanto do alisamento exponencial (AE) quanto do modelo espaço de estados (MEE), nas séries Brasil e sexo feminino. No entanto, no AE da série masculino, não houve ajuste adequado das previsões: observou-se que o ponto referente a 2011, encontra-se fora do intervalo previsto. (Figuras 4 e 5 e Tabela 4).

A comparação entre os dois métodos mostrou que, para as séries de taxas de mortalidade por diabetes no Brasil, no período de 1996 a 2008, o Modelo Espaço de Estados apresentou melhor previsão que o Alisamento Exponencial.

O declínio das taxas de morte por diabetes e a previsão de continuidade dessa redução são um achado positivo, pois podem ser decorrentes de ações de saúde. No entanto, diante da epidemia de excesso de peso e obesidade, consumo não saudável de alimento e

inatividade física, que vem se observando no país e no mundo, verifica-se também um aumento da incidência do diabetes e, conseqüentemente, de sua prevalência (WHO, 2014). Tal quadro poderá reverter o declínio da mortalidade por diabetes se as ações para o enfrentamento dessa situação epidemiológica não forem efetivamente implantadas por meio de ações de promoção de saúde.

Referências Bibliográficas

- Alves CG, Morais Neto OL. Tendência da mortalidade prematura por doenças crônicas não transmissíveis nas unidades federadas brasileiras. *Ciência & Saúde Coletiva* 2015; 20(3):641-654.
- Alves GJ, Gomes MFM, Lima JE, Gomes MTM. Modelo de previsão de preços: um estudo para o algodão brasileiro. *Revista de Economia e Administração* 2012; 11(2): 235-249.
- Barreto SM, Passos VMA, Almeida SKF, Assis TD. The increase of diabetes mortality burden among Brazilian adults. *Rev Panam Salud Publica* 2007; 22(4): 239–45.
- Bowerman, B.L., O'Connell, R.T. (1993) *Forecasting and Time Series*. 3rd ed. Belmont: Duxbury Press.
- Cesse EAP, Carvalho EF, Souza WV, Luna CF. Tendência da mortalidade por diabetes melito no Brasil: 1950 a 2000. *Arq Bras Endocrinol Metab.* 2009; 53(6):760-6.
- Chiogna M, Gaetan C. Dynamic generalized linear models with application to environmental epidemiology. *Applied Statistics* 2002; 51: 453-468.
- Franco GC. Minicurso de séries temporais. Outro Preto: 2010.
- Franco GC, Santos TR, Ribeiro JA, Cruz FRB. Confidence intervals for hyperparameters in structural models. *Communications in Statistics: Simulation and Computation* 2008; 37(3): 486-497.
- Franco GC, Gamerman D, Santos TR. Modelos de Espaço de Estados: abordagens clássica e bayesiana. Livro de minicurso da 13a ESTE. 2009; São Carlos: ABE.
- Harvey AC. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: University Press, 1989
- Harvey AC. *The Econometric Analysis of Time Series* New Jersey: Prentice-Hall, 1990.
- Hemming K, Shaw JEH. A parametric dynamic survival model applied to breast cancer survival times. *Applied Statistics* 2002; 51: 421-435.
- Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. ONR Research Memorandum, Carnigie Institute, v. 52, 1957.
- Kalman, RE. A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng* 1960; 82: 35-45.
- Latorre MRDO, Cardoso MRA. Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. *Rev. Bras. Epidemiol.* 2001; 4(3): 145-152.
- Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* 2012; 380 (December 15/22/29): 2095-2128.
- Malta DC, Moura L, Prado RR, Escalante JC, Schmidt MI, Duncan BB. Mortalidade por doenças crônicas não transmissíveis no Brasil e suas regiões, 2000 a 2011. *Epidemiol. Serv. Saúde* 2014; 23(4):599-608, out-dez.
- Mathers CD, Loncar D. Projections of Global Mortality and Burden of Disease from 2002 to 2030. *Plos Medicine* 2006; 3(11): 2011-2029.

Mattos PE, Luz LL, Santiado LM, Mattos IE. Tendência da mortalidade por diabetes melito em capitais brasileiras, 1980-2007. *Arq Bras Endocrinol Metab.* 2012; 56(1):39-46.

Morettin PA, Toloí CMC. Análise de séries temporais. Blucher, 2006.

Ribeiro JA. Inferência sobre os hiperparâmetros dos modelos estruturais usando bootstrap. Orientadora: Glaura da Conceição Franco. 2006. Dissertação (Mestrado em Estatística) - Departamento de Estatística do Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

Rodriguez A, Ruiz E. Bootstrap prediction intervals in state-space models. *Journal of time series analysis.* 2009; 30(2): 167-178.

World Health Organization. Global status report on noncommunicable diseases 2014: Attaining the nine global noncommunicable diseases targets; a shared responsibility. World Health Organization: 2014.

APÊNDICE

Apêndice A

Análise de resíduos

Série: Brasil

Alisamento exponencial

Figura A1: Gráfico de resíduos para o modelo Alisamento exponencial - Brasil

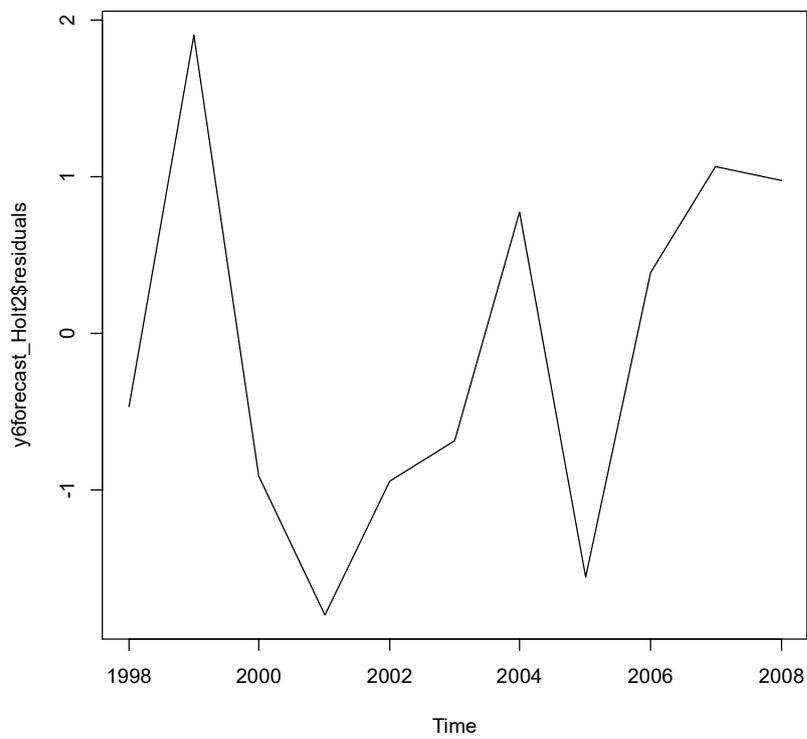


Figura A2: ACF dos resíduos para o modelo Alisamento exponencial - Brasil

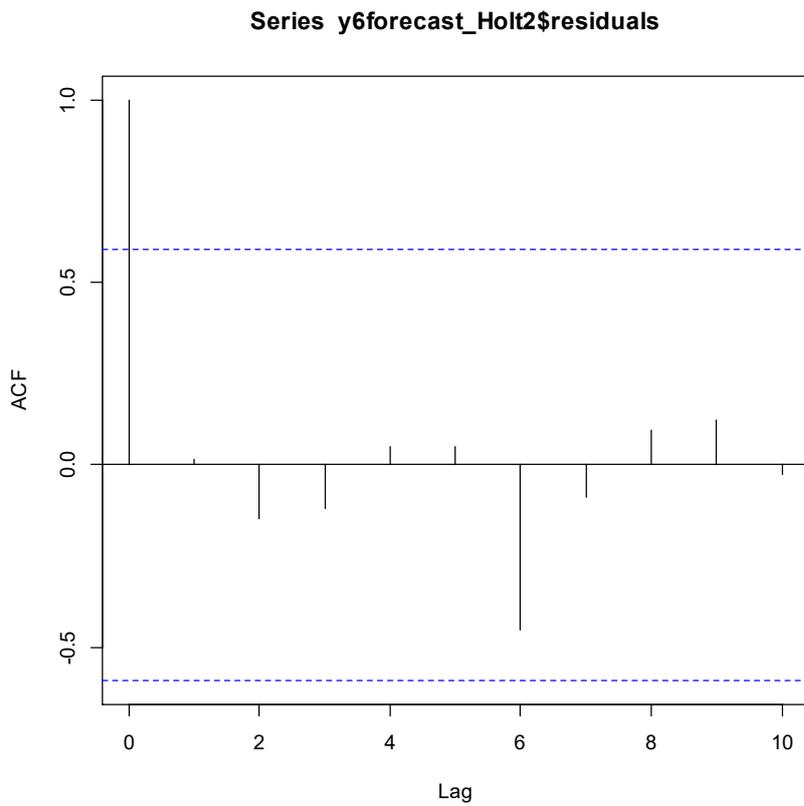
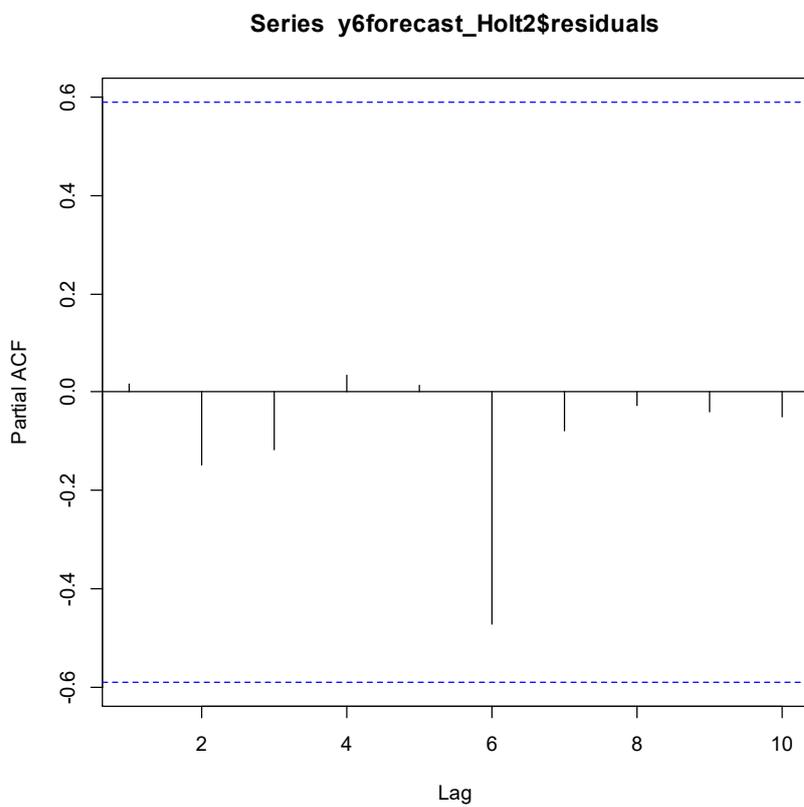


Figura A3: PACF dos resíduos para o modelo Alisamento exponencial - Brasil



```
> shapiro.test(y6forecast_Holt2$residuals)
```

Shapiro-Wilk normality test

data: y6forecast_Holt2\$residuals

W = 0.9481, p-value = 0.6202

Modelo Espaço de estados

Figura A4: Gráfico de resíduos para o modelo Espaço de Estados - Brasil

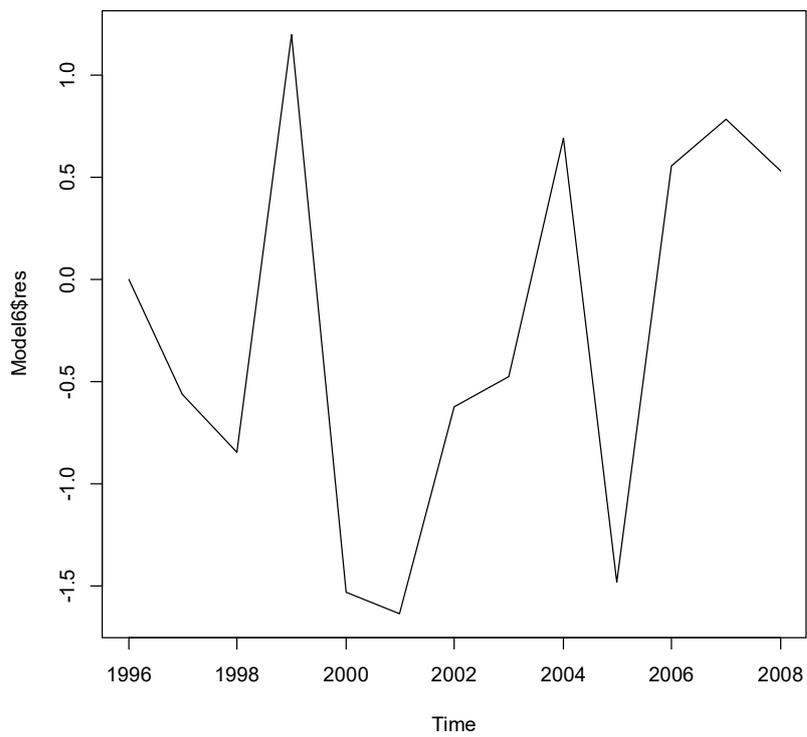


Figura A5: ACF dos resíduos para o modelo Espaço de Estados - Brasil

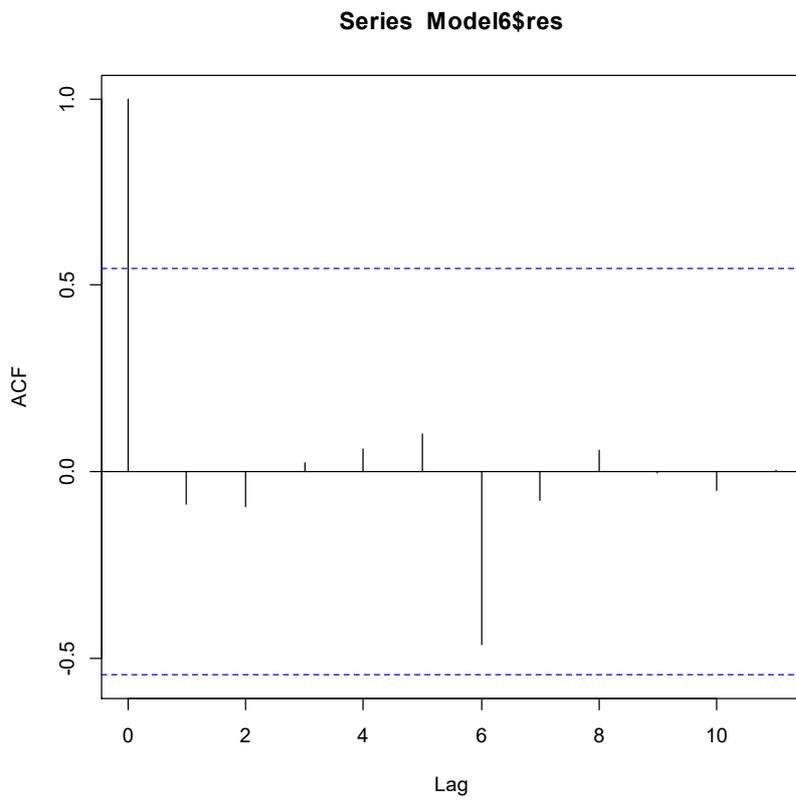
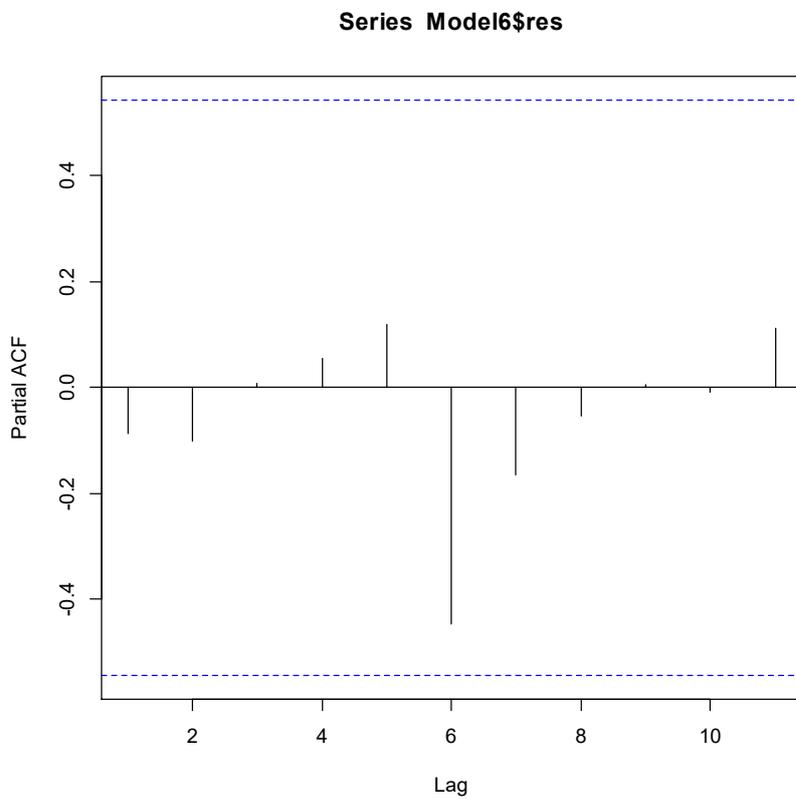


Figura A6: PACF dos resíduos para o modelo Espaço de Estados - Brasil



```
>shapiro.test(Model6$res)
Shapiro-Wilk normality test
```

```
data: Model6$res
W = 0.9285, p-value = 0.3258
```

Apêndice B

Análise de resíduos

Série: Sexo masculino

Alisamento exponencial

Figura A7: Gráfico de resíduos para o modelo Alisamento exponencial – masculino

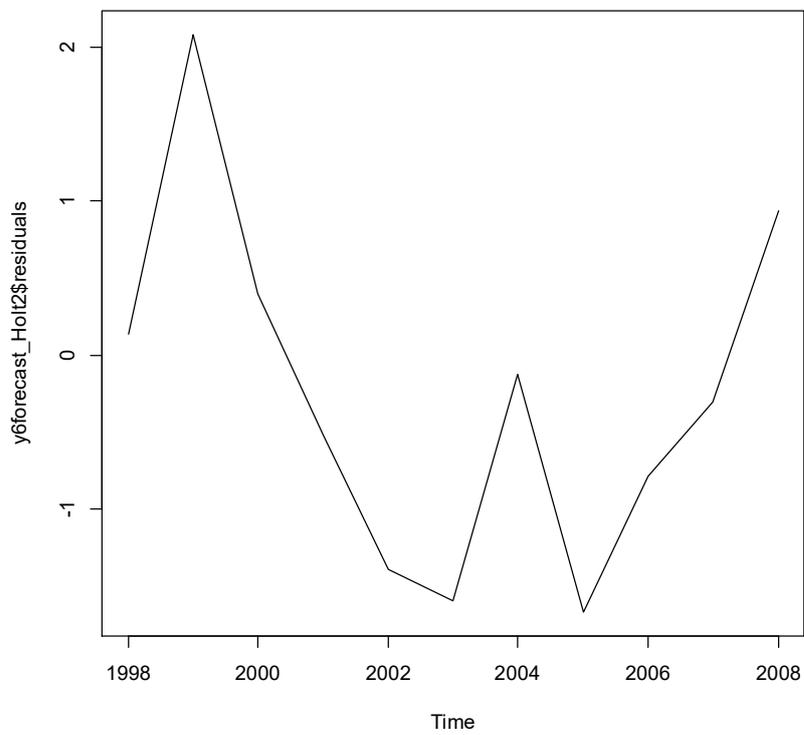


Figura A8: ACF dos resíduos para o modelo Alisamento exponencial – masculino

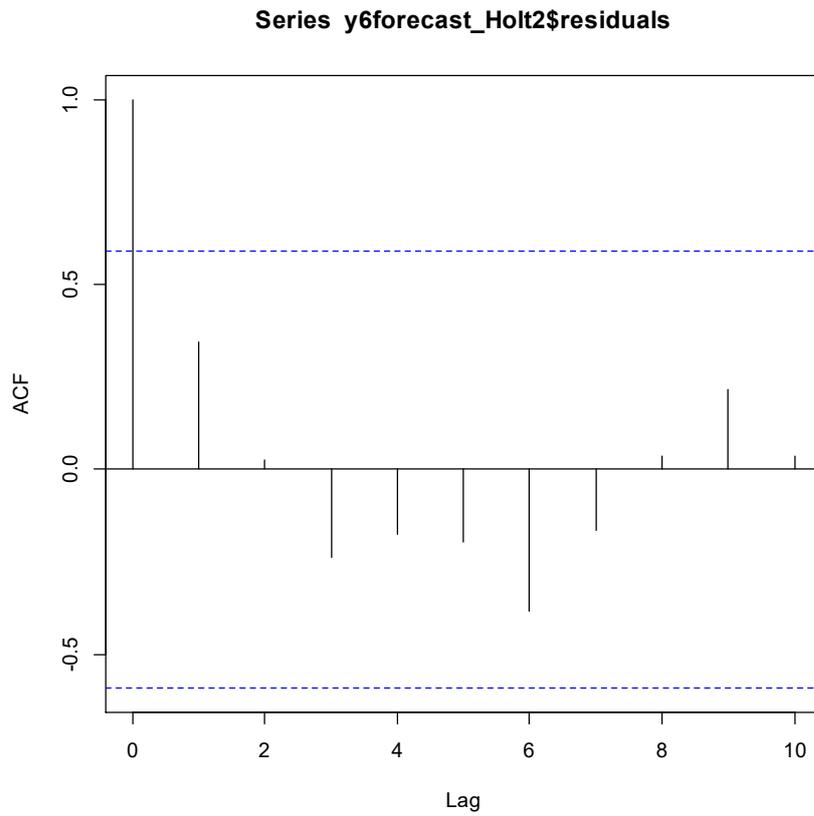
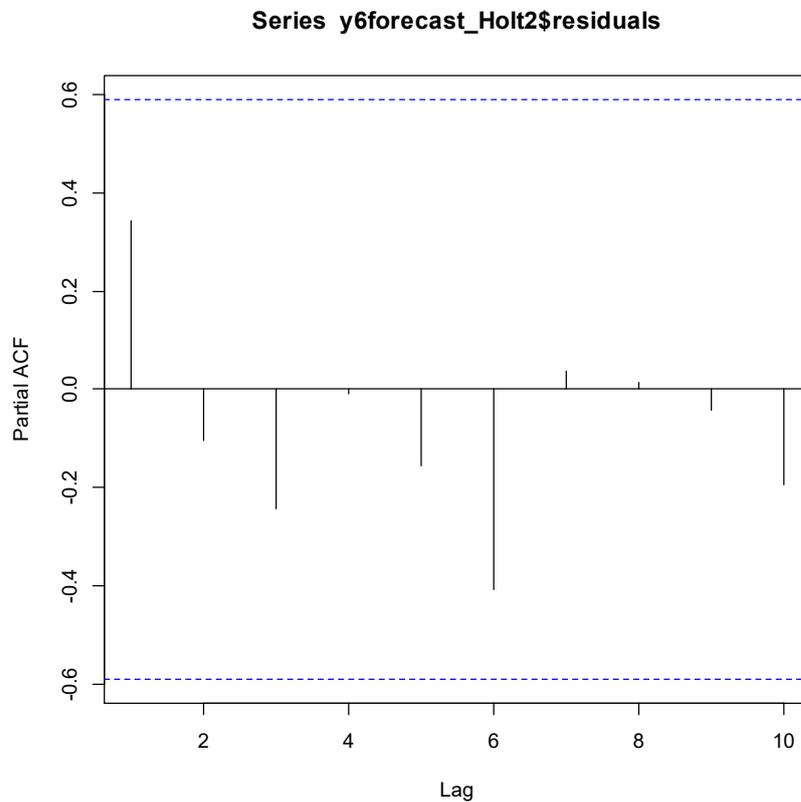


Figura A9: PACF dos resíduos para o modelo Alisamento exponencial – masculino



```
>shapiro.test(y6forecast_Holt2$residuals)
```

Shapiro-Wilk normality test

data: y6forecast_Holt2\$residuals

W = 0.9504, p-value = 0.6487

Modelo Espaço de Estados

Figura A10: Gráfico de resíduos para o modelo Espaço de Estados - masculino

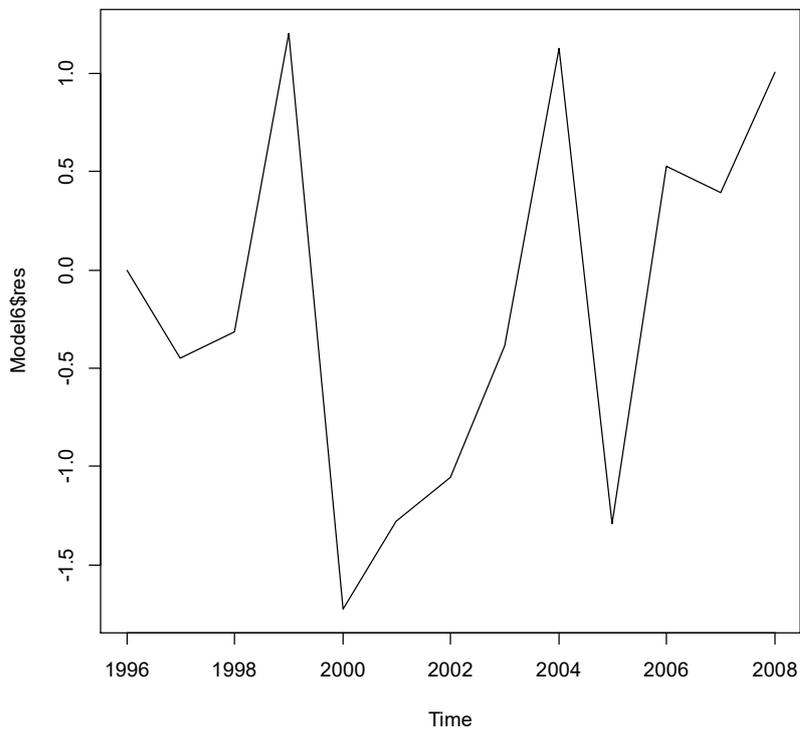


Figura A11: ACF dos resíduos para o modelo Espaço de Estados – masculino

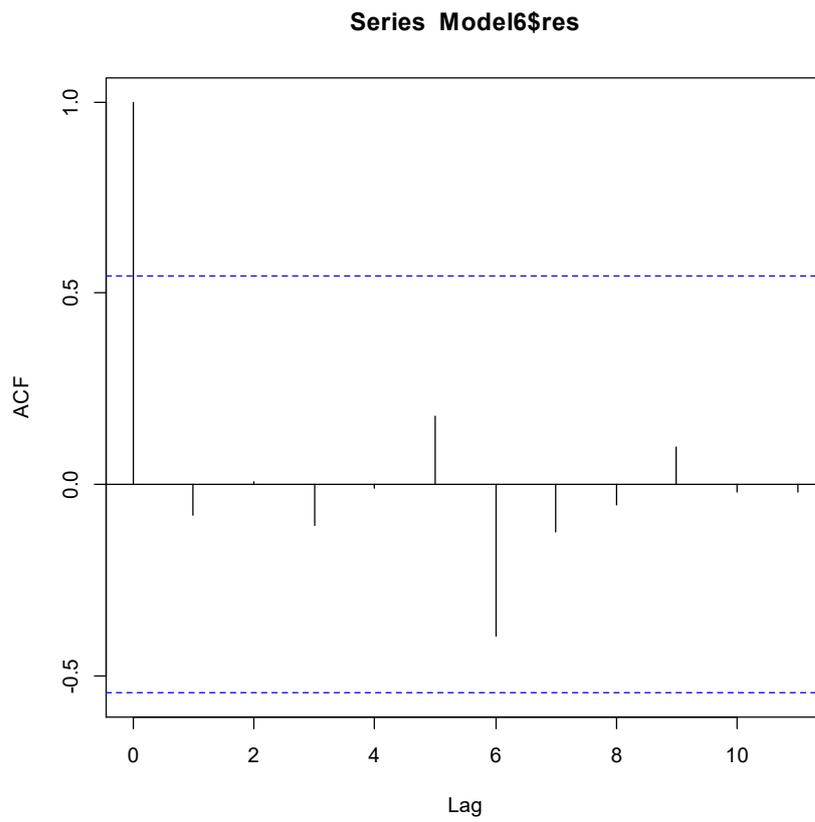
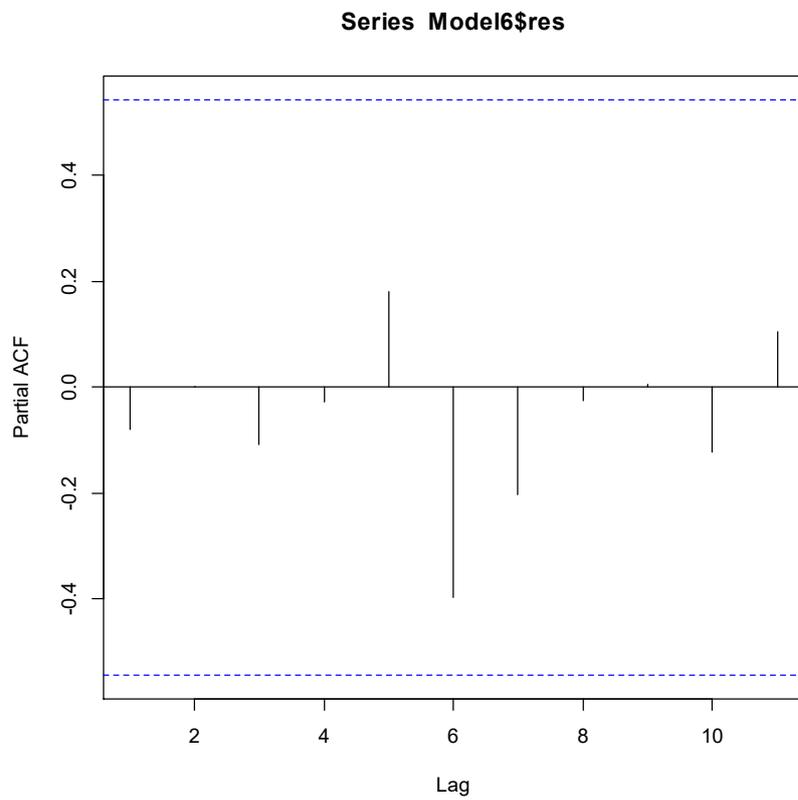


Figura A12: PACF dos resíduos para o modelo Espaço de Estados - masculino



```
>shapiro.test(Model6$res)
Shapiro-Wilk normality test
```

```
data: Model6$res
```

```
W = 0.9422, p-value = 0.4854
```

Apêndice C

Análise de resíduos

Série: Sexo feminino

Alisamento exponencial

Figura A13: Gráfico de resíduos para o modelo Alisamento exponencial – feminino

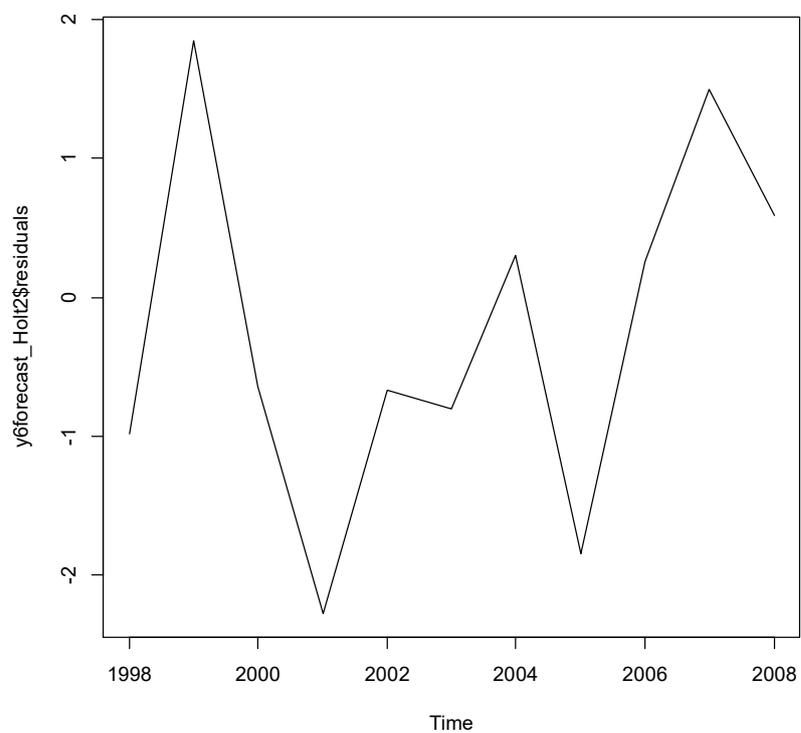


Figura A14: ACF dos resíduos para o modelo Alisamento exponencial – feminino

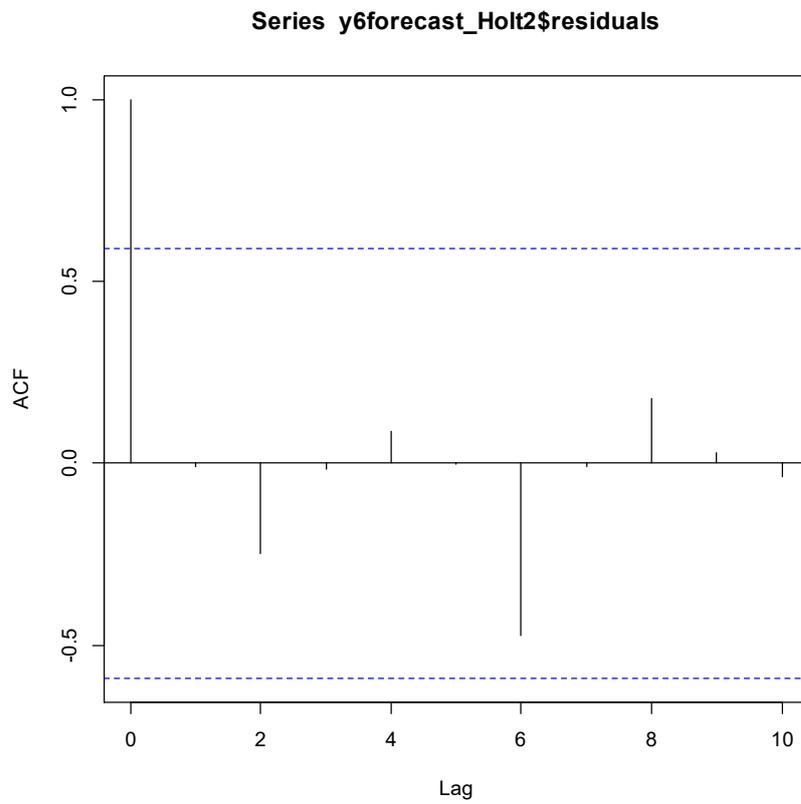
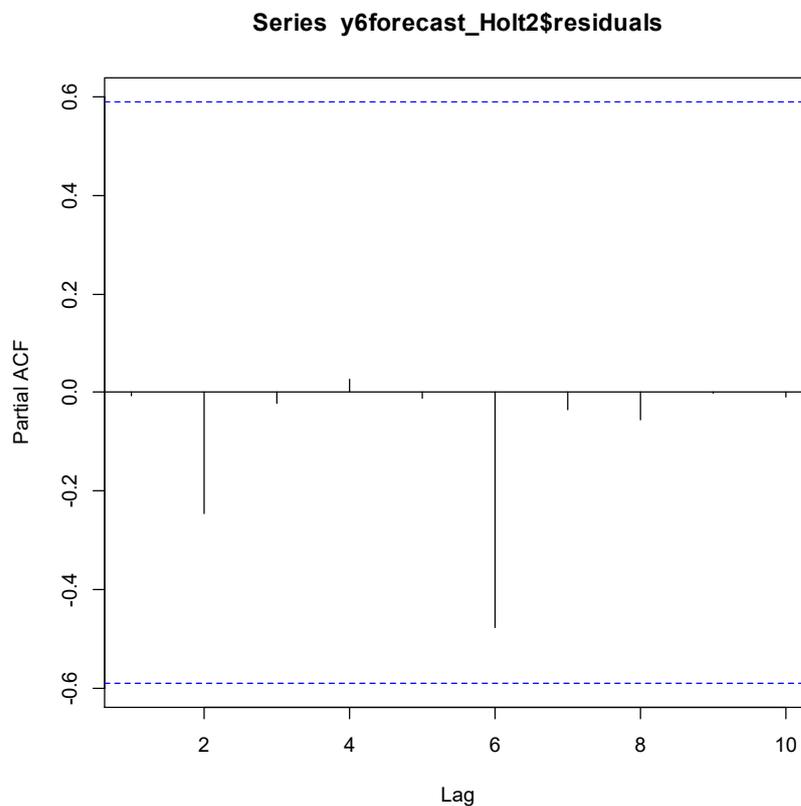


Figura A15: PACF dos resíduos para o modelo Alisamento exponencial – feminino



```
>shapiro.test(y6forecast_Holt2$residuals)
```

Shapiro-Wilk normality test

data: y6forecast_Holt2\$residuals

W = 0.9655, p-value = 0.8376

Modelo Espaço de Estados

Figura A16: Gráfico de resíduos para o modelo Espaço de Estados - feminino

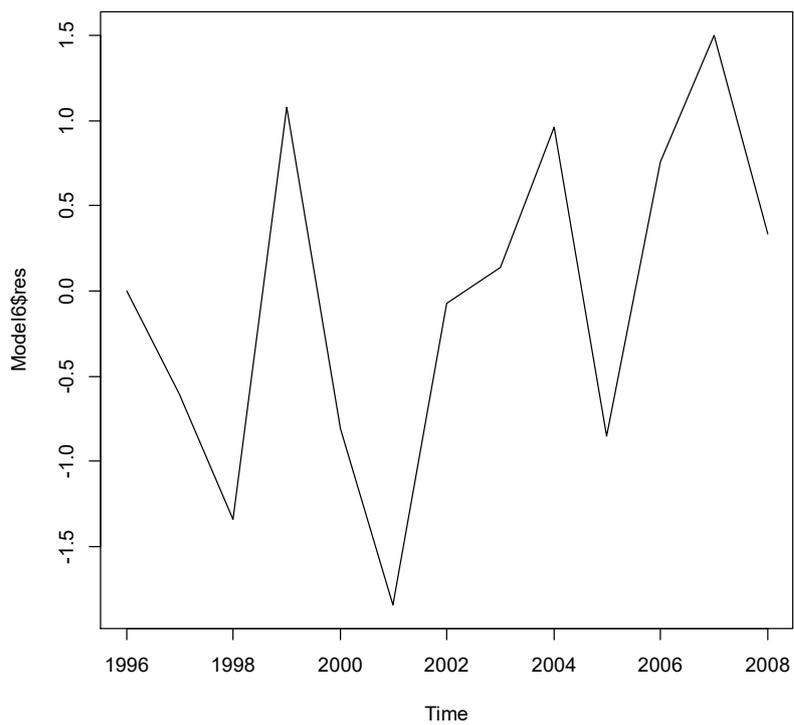


Figura A17: ACF dos resíduos para o modelo Espaço de Estados – feminino

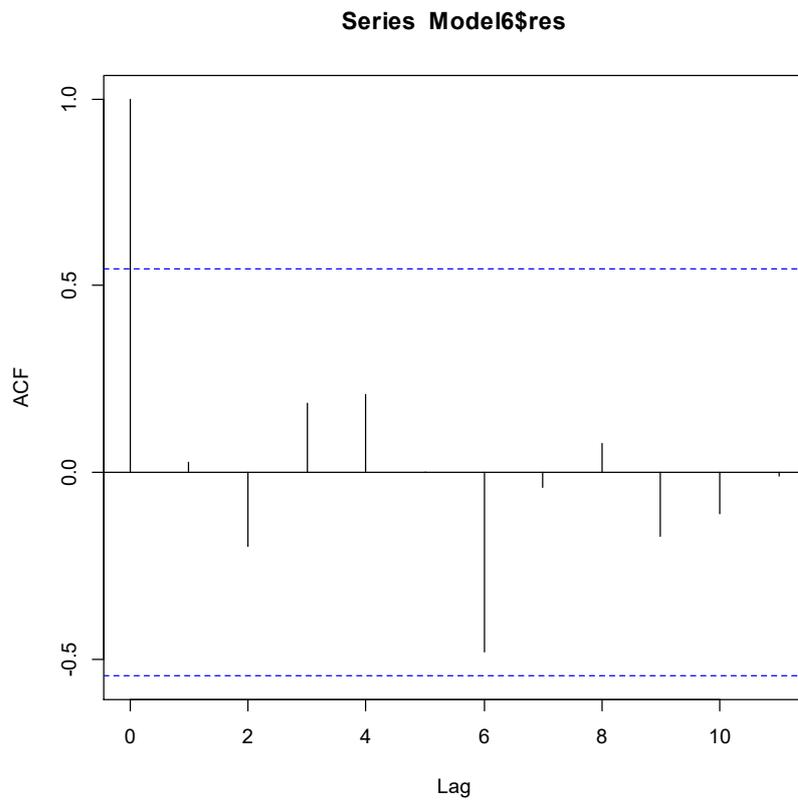
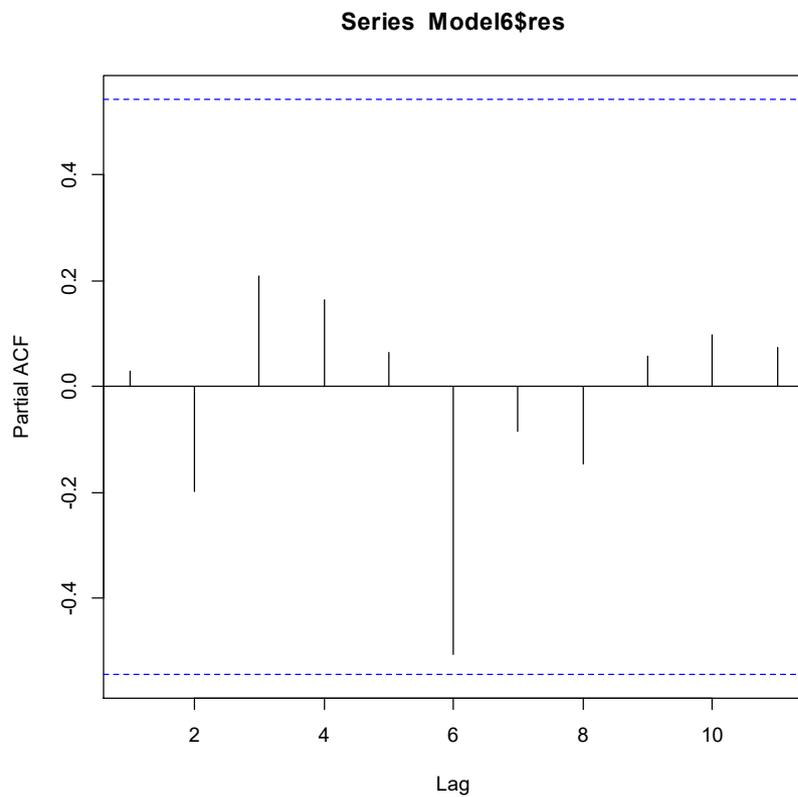


Figura A18: PACF dos resíduos para o modelo Espaço de Estados - feminino



```
>shapiro.test(Model6$res)
```

Shapiro-Wilk normality test

data: Model6\$res

W = 0.9763, p-value = 0.9563