

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

Modelos Dinâmicos
Para Dados Agregados

por

Leandro Tavares Correia

Orientador: Prof.^a Dr.^a Cibele Queiroz da Silva

Coorientador: Prof. Dr. Helio dos Santos Migon

Fevereiro de 2010

Leandro Tavares Correia

Modelos Dinâmicos Para Dados Agregados

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Universidade de Brasília
Brasília, Fevereiro de 2010

Agradecimentos

- Agradeço primeiramente aos professores Cibele Q. da Silva e Helio S. Migon pela oportunidade, paciência e aprendizado que foi me passado durante o desenvolvimento deste trabalho.
- Aos professores Donald Pianto, Antônio Eduardo Gomes, Alexandra Schmidt e Francisco Cribari-Neto pela troca de informações, orientações e sugestões relevantes para o trabalho.
- E, finalmente, a Capes e ao Departamento de Estatística da UnB pelo suporte e estrutura fornecida durante este período.

Sumário

Lista de Figuras	5
Lista de Tabelas	6
Resumo	7
Abstract	8
Introdução	9
1 Conceitos Básicos em Modelos Dinâmicos	11
1.1 Fundamentação Teórica	11
1.1.1 Séries Temporais	11
1.1.2 Abordagem Bayesiana em Modelos Dinâmicos	12
1.1.3 Modelo Linear Dinâmico Normal Univariado	14
1.1.4 Modelo Linear Dinâmico Normal Multivariado	18
1.1.5 Fator de Desconto	20
2 Modelagem e Previsão Multivariada	21
2.1 Modelo Linear Dinâmico Normal Matriz-Variado	21
2.1.1 Estrutura do Modelo Matricial	22
2.1.2 A Distribuição do vetor ϵ_t :	23
2.1.3 A Distribuição de Ω_t :	23
2.1.4 A distribuição de Σ :	25
2.1.5 A Distribuição Conjunta de (Θ_t, Σ) :	25
2.2 Processo de Estimação	26
2.3 Aplicação a Dados Composicionais	27

2.4	Transformação Razão Log e Distribuição Logística-Normal	27
2.4.1	Escala Razão-Log	27
2.4.2	Distribuição Logística-Normal	29
3	Modelo Linear Dinâmico Generalizado	31
3.1	Exemplo: Modelo Dinâmico Binomial	33
4	Modelo Dinâmico Beta com priori conjugada e ϕ conhecido	37
4.1	Nova parametrização	37
4.2	Contextualizando para o modelo dinâmico	39
5	Modelo Dinâmico Beta com priori Beta	42
5.1	Modelo Dinâmico Beta com ϕ conhecido	42
5.1.1	Inferência para o Modelo Dinâmico Beta	43
5.2	Modelo Dinâmico Beta com ϕ desconhecido	53
5.2.1	Procedimento Ad-Hoc para estimar ϕ em dois estágios	54
6	Simulação e Aplicação	56
6.1	A razão sinal-ruído no Modelo Dinâmico Beta	56
6.2	Lidando com observações $Y_t = 0$ ou $Y_t = 1$	58
6.3	Analisando dados simulados - ϕ conhecido	58
6.4	Comparações Entre Métodos de Estimação	67
6.5	Analisando dados simulados - ϕ desconhecido	72
6.6	Modelo Dinâmico Beta versus Modelo Linear Dinâmico Normal Matriz-Variado	76
6.7	Aplicação do Modelo Dinâmico Beta Em Dados Reais	77
6.7.1	Taxa Mensal de Desemprego no Brasil	77
6.7.2	Proporção de argila em diferentes profundidades de água em um lago ártico	85
6.7.3	Proporção de machos em uma população de gambás	88
7	Conclusões e Trabalhos Futuros	90
7.1	Conclusões	90
7.2	Sugestões de Trabalhos Futuros	92

7.2.1	Ajuste com ϕ integrado	92
7.2.2	Modelo Dinâmico Beta com μ_t e ϕ Desconhecidos	94
7.2.3	Modelo Dinâmico Dirichlet	98

Referências Bibliográficas	108
-----------------------------------	------------

Lista de Figuras

1.1	Série estimada a partir do modelo linear dinâmico de 1ª ordem, com intervalo de credibilidade de 90%.	18
5.1	Função de log-verossimilhança gerada de dados com $\phi = 50$ e $\mu = 0,4$	54
6.1	Caso 4 - Série com $W = 0.01$ e $\phi = 100$	61
6.2	Caso 4 - Gráfico de Dispersão da série com $W = 0.01$ e $\phi = 100$	62
6.3	Caso 9 - Série com $W = 0.15$ e $\phi = 25$	63
6.4	Caso 9 - Gráfico de dispersão da série com $W = 0.15$ e $\phi = 25$	64
6.5	Caso 10 - Série com $W = 0.20$ e $\phi = 15$	65
6.6	Caso 10 - Gráfico de dispersão da série com $W = 0.2$ e $\phi = 15$	66
6.7	Caso 4 - Série com $W = 0.01$ e $\phi = 100$: a) Série estimada via aproximação de Laplace e expansão de 2ª ordem; b) Dispersão da série estimada via aproximação de Laplace e expansão de 2ª ordem; c) Série estimada via moda a posteriori e expansão de 1ª ordem; d) Dispersão da série estimada via moda a posteriori e expansão de 1ª ordem.	69
6.8	Caso 9 - Série com $W = 0.15$ e $\phi = 25$: a) Série estimada via aproximação de Laplace e expansão de 2ª ordem; b) Dispersão da série estimada via aproximação de Laplace e expansão de 2ª ordem; c) Série estimada via moda a posteriori e expansão de 1ª ordem; d) Dispersão da série estimada via moda a posteriori e expansão de 1ª ordem.	70

6.9	Caso 10 - Série com $W = 0.2$ e $\phi = 15$: a) Série estimada via aproximação de Laplace e expansão de 2ª ordem; b) Dispersão da série estimada via aproximação de Laplace e expansão de 2ª ordem; c) Série estimada via moda a posteriori e expansão de 1ª ordem; d) Dispersão da série estimada via moda a posteriori e expansão de 1ª ordem. . . .	71
6.10	Caso 10 - Gráficos de dispersão dos níveis reais e estimados para a séries Beta geradas quando ϕ é estimado utilizando $\hat{\phi}_{min}$, $\hat{\phi}_{média}$, $\hat{\phi}_{mediana}$ e $\hat{\phi}_{máx}$	73
6.11	Caso 4 - Série com $W = 0.01$ e $\phi = 100$: a) Série estimada via MLDNM; b) Dispersão da série estimada via MLDNM; c) Série estimada via MDB; d) Dispersão da série estimada via MDB.	78
6.12	Caso 9 - Série com $W = 0.15$ e $\phi = 25$: a) Série estimada via MLDNM; b) Dispersão da série estimada via MLDNM; c) Série estimada via MDB; d) Dispersão da série estimada via MDB.	79
6.13	Caso 10 - Série com $W = 0.2$ e $\phi = 15$: a) Série estimada via MLDNM; b) Dispersão da série estimada via MLDNM; c) Série estimada via MDB; d) Dispersão da série estimada via MDB.	80
6.14	Taxa de desemprego mensal brasileira baseada nos dados da PME do período entre março de 2002 a julho de 2009. Fonte: PME/IBGE. . .	83
6.15	Gráfico dos efeitos de nível, tendência e sazonalidade estimados para a taxa de desemprego mensal brasileira baseada nos dados da PME do período entre março de 2002 a julho de 2009. Fonte: PME/IBGE. . .	84
6.16	Proporção de argila em diferentes níveis de profundidade de água em um lago artigo estimado via Modelo Dinâmico Beta.	86
6.17	Gráfico dos efeitos de nível e tendência estimados para a proporção de argila em diferentes níveis de profundidade de água em um lago artigo. . .	87
6.18	Proporção de gambás (<i>Gracilinamus microtarsus</i>) machos capturados nas amostras observadas. Em a) a série foi ajustada considerando $\hat{W} = 0.25$ e $\hat{\phi} = 3$. Em b) a série foi ajustada considerando o fator de desconto $\delta = 0.6$ e $\hat{\phi} = 3$	89

Lista de Tabelas

6.1	Combinações de séries avalidas	59
6.2	Erro quadrático médio das séries simuladas.	60
6.3	Algumas estatísticas para os erros $e_t = [\mu_t - \tilde{\mu}_t]$ e erro quadrático médio (<i>EQM</i>) entre os níveis reais e obtidos pelo Modelo Dinâmico Beta estimados usando tanto a moda a posteriori e aproximações de Taylor de primeira ordem ou aproximação de Laplace e aproximações de Taylor segunda ordem para os momentos $(\mu_t D_t, \phi)$	68
6.4	Comparação dos resíduos com ϕ conhecido e ϕ estimado nos Modelos Dinâmicos Beta	74
6.5	Comparação das <i>EQM</i> 's e valores estimados de ϕ para os diferentes casos e características de W_t e ϕ nos Modelos Dinâmicos Beta	75
6.6	Algumas Estatísticas dos erro $e_t = \mu_t - \tilde{\mu}_t$ e o erro quadrático médio (<i>EQM</i>) para comparar os ajustes das séries beta analisadas pelo Modelo Dinâmico Beta ou pelo Modelo Linear Dinâmico Normal Matriz-Variado.	77

Resumo

Com base na abordagem bayesiana de *Modelos Dinâmicos*, séries de tempo de dados composicionais são modeladas para análises de previsão e de comportamento.

Por se tratar de dados convertidos para a escala de proporções relativas a uma série agregada, os modelos são construídos utilizando-se de transformações *razão-log* e distribuição *Logística-Normal*, nos casos em que se assume a normalidade dos dados. Para casos mais gerais, os modelos baseiam-se na classe dos Modelo Lineares Dinâmicos Generalizados (MLDG), e em dados com distribuição *Beta*, e tais desenvolvimentos consistem em contribuições inéditas na área de modelos dinâmicos.

Palavras Chave: *modelos dinâmicos, abordagem bayesiana, séries de tempo, dados composicionais, distribuição Beta.*

Abstract

Using a bayesian approach for *Dynamic Models*, compositional time series data are modeled for forecasting and analysing data behavior.

Since the original data is converted to the scale of relative proportions of the aggregated series, the models are constructed using *log-ratio* transformation and *Logistic-Normal* distribution for the cases where the restriction of normally distributed data is assumed. For more general situations, we develop methodology for models that are related to the class of Dynamic Generalized Linear Models (DGLM), more specifically for the *Beta* distribution. Such developments represent new contributions in the area of dynamic models.

key words: *dynamic models, bayesian approach, time series, compositional data, Beta distribution.*

Introdução

Modelos de séries temporais têm mostrado sua utilidade em diversas áreas, tanto com a finalidade de se fazer previsões futuras, como para a análise de comportamento de eventos passados. Em muitos casos, as séries podem ser desagregadas em componentes, de modo a possibilitar a previsão e avaliação do comportamento de cada uma destas componentes. Por exemplo, a receita de uma indústria desagregada em seus itens de venda. Este estudo concentra-se nesta vertente do problema e será tratado utilizando a abordagem bayesiana de *modelos dinâmicos* para séries de tendência ao longo do tempo.

Os modelos dinâmicos têm se destacado como uma importante área da estatística nos últimos anos. Partindo da estrutura do modelo de regressão, esta abordagem se estende através da implementação de uma equação de evolução, capaz de governar a evolução temporal dos regressores. Esta classe é capaz de abranger muitos dos principais problemas envolvendo variações no tempo e espaço, inclusive os modelos de séries temporais.

Este trabalho está relacionado com a classe dos denominados *Modelos Dinâmicos Lineares Generalizados* (MLDG). Introduzido no trabalho de Migon (1984) e seguido por Migon e Harrison (1985), West, Harisson e Migon (1985), Gamerman e West (1987), Gamerman (1991,1998), Lindsey e Lambert (1995) e Godolphin e Triantafyllopoulos (2006), os Modelos Lineares Dinâmicos (MLD) foram estendidos e generalizados em diversos contextos, sem a necessidade de impor a suposição de dados normalmente distribuídos.

Neste trabalho são descritos modelos dinâmicos para a distribuição *Beta*. Alguns modelos frequentistas que tratam da regressão beta (regressão beta estática) foram propostos em Paolino (2001), Kieschnick e McCullough (2003) e Ferrari e Cribari-Neto

(2004), para modelar taxas e proporções. Suposição básica é que a variável resposta segue a lei *beta* e é ligada a um preditor linear por meio de uma função de ligação. Uma versão bayesiana de modelo de regressão estático foi proposta por Branscum *et. al.* (2007) e também utilizada em Albi *et. al.* (2009) via INLA (vide Rue *et. al.*, 2006). O Modelo Dinâmico Beta descrito neste trabalho consiste em uma extensão do modelo de regressão Beta estático, permitindo que os parâmetros do modelo evoluam com o tempo. Tal modelagem representa uma contribuição inédita.

Como subproduto desta dissertação destacamos o artigo intitulado *Bayesian Beta Dynamic Model and Applications*, que está em fase de elaboração, e uma submissão para apresentação oral no EBEB 2010 a ser realizado no período de 21 a 24 de março de 2010.

Esta dissertação está organizada como a seguir: No Capítulo 1 descreve-se conceitos básicos de modelos dinâmicos lineares, considerando-se o caso de modelos dinâmicos gaussianos uniparamétricos e multiparamétricos. No Capítulo 2 aborda-se o modelo dinâmico linear gaussiano matriz-variado e apresenta-se uma descrição adequada para dados composicionais. No Capítulo 3 apresenta-se a classe dos modelos dinâmicos lineares generalizados. No Capítulo 4 introduz-se o modelo dinâmico beta considerando-se priori conjugada e as dificuldades nesta abordagem. No Capítulo 5 descreve-se priori não-conjugada. Para o parâmetro de escala ϕ são tratados os casos (a) ϕ conhecido e (b) ϕ desconhecido. No capítulo 6 os métodos desenvolvidos no Capítulo 5 são aplicados a dados simulados e também a dados reais. No Capítulo 7 apresenta-se as conclusões finais e trabalhos futuros.

Capítulo 1

Conceitos Básicos em Modelos Dinâmicos

1.1 Fundamentação Teórica

Neste capítulo fazemos uma revisão dos Modelos Dinâmicos Lineares Gaussianos para o caso uniparamétrico e multiparamétrico.

1.1.1 Séries Temporais

No caso em que o conjunto de observações constitui-se de componentes ordenadas no tempo, o problema é caracterizado como uma série temporal. Este tipo de estrutura é muito comum em áreas como Medicina, Economia, Epidemiologia e Meteorologia. Grande parte da dificuldade ao se trabalhar com séries temporais está em tratar a dependência entre observações vizinhas. Em geral, os objetivos de uma análise de série temporal são:

- Compreender os mecanismos geradores da série:
 - ⇒ descrever eficientemente o comportamento da série;
 - ⇒ encontrar periodicidades na série;
 - ⇒ identificar os causadores de tais comportamentos.
- Predizer o comportamento futuro da série:

- ⇒ construir planos a médio, curto e longo prazo;
- ⇒ obter direcionamento para tomada de decisões.

As séries de tempo são formadas pelo conjunto de informações $\{Y_t, t \in T\}$, onde Y é a variável de interesse e T o conjunto de índices. As séries podem ser classificadas como:

- Discreta: $T = \{t_1, t_2, \dots, t_n\}$;
- Contínua: $T = \{t, t_1 < t < t_2\}$;
- Multivariada: Observações são $\{Y_{1t}, Y_{2t}, \dots, Y_{kt}, t \in T\}$.

O foco de uma análise de séries temporais está na construção de um modelo. O modelo deve ser capaz de organizar as informações de forma a propiciar aprendizagem e previsão. Devido à incerteza presente, o modelo é probabilístico. Este deve ser, também, parsimonioso, de modo que sua descrição seja relativamente simples para poder se adaptar ao futuro e facilitar o processamento da informação. A estrutura dinâmica das séries faz com que os modelos tenham que ter adaptabilidade no tempo, ou seja, deve ser parametrizado de forma a permitir mudanças locais em sua estrutura.

1.1.2 Abordagem Bayesiana em Modelos Dinâmicos

A abordagem bayesiana dos modelos dinâmicos possui a interessante capacidade de incorporar, no modelo, todas as informações relevantes disponíveis: desde dados históricos, experiências concretas ou subjetivas, assim como conhecimento de fenômenos futuros. Além de previsões rotineiras, excessões podem também ser implementadas por antecipações ou em bases retrospectivas. Algumas vantagens, em relação à análise frequentista, são obtidas naturalmente, como intervenção, função de transferência e análise retrospectiva (alisamento), entre outras.

O modelo completo pode ser decomposto em componentes lineares dinâmicas independentes (MLD) onde cada uma descreve aspectos particulares do estudo. Esta estrutura de modelo permite que tanto análises prospectivas como retrospectivas sejam acomodadas facilmente.

A inferência é feita de maneira sequencial, ou seja, é refeita a cada tempo t para o conjunto de informações Y_t . O processo inicia-se baseado no conjunto de informações disponível antes de se observar os dados, ou seja, no tempo $t = 0$. Este conjunto de dados (subjativos ou não) é denotado por D_0 . Toda a informação sobre o futuro ($t > 0$) será condicional a D_0 . Ao chegar no tempo t , a informação disponível está concentrada em $D_t = \{y_t, D_{t-1}\}$. Assim, quando o tempo evolui de t para $t + 1$, y_{t+1} é inserido no conjunto de informações, portanto $D_{t+1} = \{y_{t+1}, D_t\}$. Há, também, a possibilidade de inserir informações além daquela obtida com as observações, de modo que D_{t+1} pode ser generalizado como $D_{t+1} = \{I_{t+1}, D_t\}$, onde I_{t+1} corresponde a toda informação adicional obtida. No caso em que toda informação adicional obtida em cada tempo t é a própria observação y_t , então $D_t = \{D_0, y_1, \dots, y_n\}$, e é denominado um sistema *fechado*. Sistemas que admitem a entrada de informações além das observações da série é chamado de sistema *aberto*. Para estudos de previsão, o interesse está em obter as distribuições preditivas $(y_{t+h}|D_t)$, $h > 0$.

A modelagem dinâmica parte da idéia de que as observações flutuam em torno de uma média., denominada nível. Esta média, no entanto, está sujeita a variações ao longo do tempo. Estas variações com o tempo são, essencialmente, estocásticas e vão depender de erros, ou perturbações, da evolução. Sendo assim, a evolução é modelada através de um passeio aleatório e a novidade desta abordagem está na evolução paramétrica caracterizada pela equação que relaciona sucessivos valores dos parâmetros, denominada *Equação de evolução*. A forma geral dos modelos dinâmicos lineares apresentam a seguinte estrutura (West e Harrison, 1997, pg. 102):

- *Equação de observação:*

$$y_t = F_t' \theta_t + \epsilon_t$$

Em geral, assume-se que $\epsilon_t \sim N(0, V_t)$.

- *Equação de evolução ou Sistema de Equações:*

$$\theta_t = G_t \theta_{t-1} + w_t$$

Em geral, assume-se que $w_t \sim N(0, W_t)$.

Esta estrutura parte da idéia de que a série de observações y_t são independentes condicionalmente a θ_t e V_t , onde $F_t'\theta_t = E(y_t|\theta_t)$ e $V_t = V(y_t|\theta_t)$. F_t é um vetor de constantes conhecidas, podendo ser um conjunto de variáveis explicativas. θ_t corresponde a um vetor de parâmetros. Na abordagem bayesiana, θ_t é interpretado como um vetor de variáveis aleatórias onde $E(\theta_t|D_t) = m_t$ e $V(\theta_t|D_t) = W_t$. G_t corresponde à matriz de termos conhecidos que define a evolução sistemática dos parâmetros. Em geral, assume-se que ϵ_t e ω_t são mutuamente independentes. O sistema de equações é iniciado assumindo que θ_0 possui uma distribuição a priori conhecida.

O modelo linear dinâmico é caracterizado pelo conjunto

$$(F_t, G_t, V_t, W_t)$$

Partindo desta estrutura, pode-se caracterizar os seguintes casos:

- Modelos clássicos de séries temporais com: $F_t = F$ e $G_t = G, \forall t$.
- Modelos clássicos de regressão linear com: $G_t = I_p$ e $V_t = 0, \forall t$.

Nas próximas seções serão introduzidos os modelos dinâmicos normal univariado e normal multivariado.

1.1.3 Modelo Linear Dinâmico Normal Univariado

O Modelo Linear Dinâmico Normal Univariado (Vide West e Harrison, 1997, pg. 32) é a forma mais simples de modelo dinâmico linear. Porém, incorpora muito dos principais conceitos de modelagem dinâmica. O modelo é denotado de 1ª ordem e como já foi mencionado anteriormente, este é formalizado a partir das duas equações:

- *Equação de observação:*

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim N(0, V_t).$$

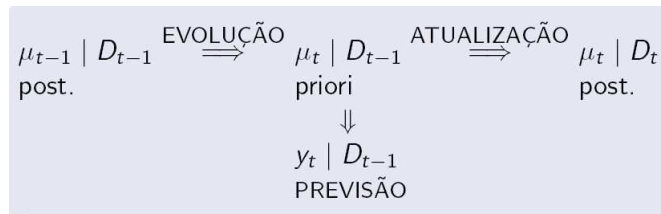
- *Equação de evolução:*

$$\mu_t = \mu_{t-1} + w_t, \quad w_t \sim N(0, W_t),$$

onde a priori inicial é determinada como $(\mu_0|D_0) \sim N(m_0, C_0)$.

Vale lembrar que ϵ_t e ω_t são assumidos todos independentes entre si e de $(\mu_0|D_0)$. Os erros ω_t controlam a evolução através da variância W_t , ou seja, quanto menor o seu valor, mais suave será a evolução. A sua média igual a zero garante uma certa localidade fixa. O tipo de trajetória descrita está relacionada com a razão W_t/V_t . Quando esta é relativamente pequena, a maior parte do movimento se dá devido às observações. No caso em que a razão apresenta um valor alto, os movimentos ocorrem devido às observações, mas também devido às variações de μ_t .

O processo inferencial é composto, a cada passo no tempo, pelos processos de *evolução*, *previsão* e *atualização*, organizados da seguinte forma:



A *evolução* é obtida a partir da equação do sistema. A *atualização* é feita incorporando a observação do tempo “presente”, y_t , através do *teorema de Bayes* (vide apêndice A). A *previsão* segue da distribuição marginal de $(y_t|D_{t-1})$, ou seja, a informação obtida antes de observar y_t .

O processo de estimação dos parâmetros segue da seguinte forma:

1. Posteriori em $t - 1$: $(\mu_{t-1}|D_{t-1}) \sim N(m_{t-1}, C_{t-1})$,

para μ_{t-1} e C_{t-1} já conhecidos.

2. Priori em t : $(\mu_t|D_{t-1}) \sim N(a_t, R_t)$

onde:

$$a_t = m_{t-1} \quad \text{e} \quad R_t = C_{t-1} + W_t.$$

3. Previsão 1 passo a frente: $Y_t|D_{t-1} \sim N(f_t, Q_t)$

onde:

$$f_t = m_{t-1} \quad \text{e} \quad Q_t = R_t + V_t.$$

4. Posteriori em t : $(\mu_t|D_t) \sim N(m_t, C_t)$

onde:

$$m_t = m_{t-1} + A_t e_t; \quad A_t = R_t/Q_t;$$
$$e_t = (Y_t - f_t) \quad \text{e} \quad C_t = R_t - A_t^2 Q_t.$$

O algoritmo acima pode ser provado a partir da aplicação do *teorema de Bayes* (vide apêndice A) e das propriedades da distribuição normal (vide apêndice B). A prova parte do princípio de indução, ou seja, assumindo que $(\mu_{t-1}|D_{t-1}) \sim N(m_{t-1}, C_{t-1})$. A seguir, apresentamos um esboço da prova. Sabe-se que:

$$E(\mu_{t-1}|D_{t-1}) = m_{t-1} \quad \text{e} \quad V(\mu_{t-1}|D_{t-1}) = C_{t-1}.$$

Como $\mu_t = \mu_{t-1} + \omega_t$, devido a independência entre μ_{t-1} e ω_t , segue que:

$$E(\mu_t|D_{t-1}) = E(\mu_{t-1} + \omega_t|D_{t-1}) = m_{t-1},$$

$$V(\mu_t|D_{t-1}) = V(\mu_{t-1} + \omega_t|D_{t-1}) = C_{t-1} + W_t.$$

Como a combinação linear de variáveis aleatórias normalmente distribuídas mantém a mesma distribuição, segue que:

$$(\mu_t|D_{t-1}) \sim N(m_t, C_t).$$

Da mesma forma, sabe-se que:

$$Y_t = \mu_t + \epsilon_t.$$

Portanto,

$$E(Y_t|D_{t-1}) = E(\mu_t + \epsilon_t|D_{t-1}) = m_{t-1} \quad \text{e}$$

$$V(Y_t|D_{t-1}) = V(\mu_t + \epsilon_t|D_{t-1}) = C_{t-1} + V_t,$$

e, devido à normalidade,

$$(Y_t|D_{t-1}) \sim N(f_t, Q_t).$$

Pelo *teorema de bayes*, tem-se:

$$P(\mu_t|D_t) = P(\mu_t|Y_t, D_{t-1}) \propto P(Y_t|\mu_t, D_{t-1})P(\mu_t|D_{t-1}).$$

Como $(\mu_t|D_t)$ possui, necessariamente, distribuição normal, as equações dos momentos podem ser obtidas através das propriedades das médias e variâncias condicionais da distribuição normal. Lembrando que qualquer função linear de Y_t e μ_t é uma combinação de quantidades normais independentes ϵ_t , ω_t e μ_{t-1} . Portanto, condicionalmente a D_{t-1} , esta é normalmente distribuída.

Levando em conta o raciocínio acima, basta identificar o vetor de médias e a matriz de covariância da distribuição conjunta $(Y_t, \mu_t|D_{t-1})$. As médias já são conhecidas: $E(Y_t|D_{t-1}) = E(\mu_t|D_{t-1}) = m_{t-1}$. As variâncias também são conhecidas pelas distribuições marginais: $V(Y_t|D_{t-1}) = Q_t$ e $V(\mu_t|D_{t-1}) = R_t$. Logo, resta determinar a covariância entre Y_t e μ_t :

$$Cov(Y_t, \mu_t|D_{t-1}) = Cov(\mu_t + \epsilon_t, \mu_t|D_{t-1}) = Cov(\mu_t, \mu_t|D_{t-1}) + Cov(\epsilon_t, \mu_t|D_{t-1}).$$

Como μ_t e ϵ_t são independentes,

$$Cov(Y_t, \mu_t|D_{t-1}) = V(\mu_t|D_{t-1}) = R_t.$$

Logo,

$$\begin{pmatrix} Y_t \\ \mu_t \end{pmatrix} \Big|_{D_{t-1}} \sim N_2 \left[\begin{pmatrix} m_{t-1} \\ m_{t-1} \end{pmatrix}, \begin{pmatrix} Q_t & R_t \\ R_t & R_t \end{pmatrix} \right].$$

Seguindo propriedades da distribuição normal bivariada, pode-se obter a distribuição condicional:

$$(\mu_t|Y_t, D_{t-1}) \sim N(m_t, C_t).$$

É interessante observar que e_t (vide Passo 4) pode ser interpretado como o *erro de previsão* e A_t como o *peso adaptativo* dada a observação mais recente y_t , pois m_t pode ser reescrito como:

$$m_t = m_{t-1} + A_t e_t = \frac{R_t V_t}{Q_t} \left(\frac{m_{t-1}}{R_t} + \frac{y_t}{V_t} \right) = A_t y_t + (1 - A_t) m_{t-1}.$$

e

$$C_t = (1 - A_t) R_t = \frac{1}{R_t^{-1} + V_t^{-1}} = \frac{R_t V_t}{Q_t} = A_t V_t.$$

A_t pode ser interpretado como o quadrado da correlação entre y_t e μ_t , ou como o coeficiente de regressão linear entre μ_t e y_t .

Avaliando o aspecto preditivo do modelo, previsões considerando h passos a frente são baseadas em:

$$\begin{aligned}
 E(y_{t+h}|\mu_t) &= E(\mu_{t+h} + \epsilon_{t+h}) = E(\mu_{t+h}|D_t) \\
 &= E(\mu_{t+h-1} + \omega_{t+h}|D_t) = \dots \\
 &= E(\mu_t + \omega_{t+1} + \dots + \omega_{t+h}|D_t) = m_t.
 \end{aligned} \tag{1.1}$$

Portanto, $f_t(h) = E(y_{t+h}|\mu_t) = E(\mu_{t+h}|D_t) = E(\mu_t|D_t) = m_t$.

As distribuições preditivas considerando h passos a frente são dadas por:

$$(y_{t+h}|D_t) \sim N(f_t(h), Q_t(h)),$$

onde $Q_t(h) = V(\mu_t + \omega_{t+1} + \dots + \omega_{t+h}|D_t) = C_t + \sum_{j=1}^h W_{t+j} + V_{t+j}$.

A Figura (1.1) ilustra o exemplo de uma série de tempo construída com as variâncias constantes $W_t = W = 0,5$ e, $V_t = V = 1$. A série é estimada pelo método descrito:

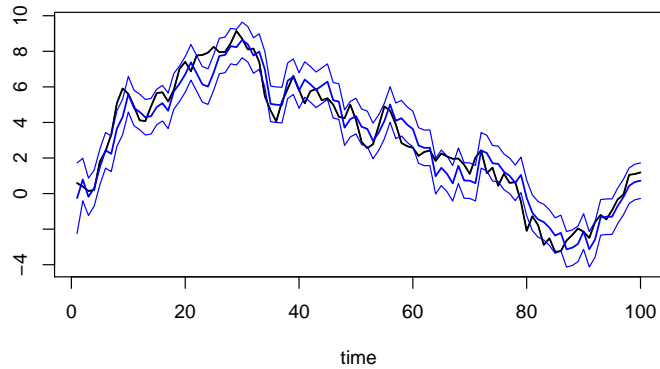


Figura 1.1: Série estimada a partir do modelo linear dinâmico de 1ª ordem, com intervalo de credibilidade de 90%.

1.1.4 Modelo Linear Dinâmico Normal Multivariado

No modelo multivariado, μ_t é substituído por um vetor de parâmetros desconhecidos, θ_t . Seguem abaixo as equações:

- *Equação de observação:*

$$y_t = F_t' \theta_t + \epsilon_t, \quad \epsilon_t \sim N(0, V_t).$$

- *Equação de evolução:*

$$\theta_t = G_t \theta_{t-1} + \omega_t, \quad \omega_t \sim N(0, W_t).$$

- F_t e θ_t são vetores de dimensão p , com F_t conhecido.
- G_t é uma matriz $p \times p$ que define a evolução de θ_t .
- $(\theta_0 | D_0) \sim N_p(m_0, C_0)$.

O processo de estimação do modelo multivariado segue o mesmo procedimento do modelo de 1ª ordem.

1. Posteriori em $t - 1$: $(\theta_{t-1} | D_{t-1}) \sim N(m_{t-1}, C_{t-1})$,

para θ_{t-1} e C_{t-1} já conhecidos.

2. Priori em t : $(\theta_t | D_{t-1}) \sim N(a_t, R_t)$

onde:

$$a_t = G_t m_{t-1} \quad \text{e} \quad R_t = G_t C_{t-1} G_t' + W_t$$

3. Previsão 1 passo a frente: $Y_t | D_{t-1} \sim N(f_t, Q_t)$

onde:

$$f_t = F_t' a_t \quad \text{e} \quad Q_t = F_t' R_t F_t + V_t$$

4. Posteriori em t : $(\theta_t | D_t) \sim N(m_t, C_t)$

onde:

$$m_t = a_t + A_t e_t; \quad A_t = R_t F_t Q_t^{-1}; \quad e_t = (Y_t - f_t)$$

$$\text{e} \quad C_t = R_t - A_t A_t' Q_t.$$

As demonstrações seguem do mesmo raciocínio descrito no caso univariado, levando em conta que agora trata-se de vetores e, portanto, distribuições multivariadas. Para maiores detalhes vide West e Harrison (1997) pg. 97.

1.1.5 Fator de Desconto

Até o momento não foi abordado o processamento da incerteza relativa à variância dos erros de evolução W_t . Este tratamento não pode ser feito de forma analítica e, como alternativa, é usado o *Fator de Desconto*.

Sabe-se que o valor da informação decresce ao longo do tempo e esta queda é controlada pela evolução do sistema através do aumento de sua incerteza. De acordo com a equação:

$$\theta_t = G_t \theta_{t-1} + \omega_t, \quad \omega_t \sim N(0, W_t).$$

Como:

$$V(\theta_{t-1}|D_{t-1}) = C_{t-1} \quad \text{e} \quad V(G_t \theta_{t-1}|D_{t-1}) = G_t C_{t-1} G_t' = P_t,$$

segue que:

$$R_t = V(\theta_t|D_{t-1}) = P_t + W_t \quad \text{e} \quad W_t = R_t - P_t.$$

Logo, se for definido um δ tal que $R_t = P_t/\delta$, então δ pode ser interpretado como a proporção da informação que se mantém entre os períodos $t-1$ e t . Assim, W_t pode ser definido por:

$$W_t = R_t - P_t = P_t/\delta - P_t = P_t(\delta^{-1} - 1).$$

Em geral, δ costuma assumir valores acima de 0,9 para sistemas sem variações muito bruscas. Valores abaixo de 0,8 tendem a introduzir muita incerteza, produzindo intervalos de predição muito largos. Valores muito elevados tornam o sistema com mudanças muito suaves. Esta idéia de desconto pode ser aplicada a modelos mais gerais com diferentes fatores de desconto em diferentes partes do modelo.

Capítulo 2

Modelagem e Previsão

Multivariada

Este capítulo é baseado no capítulo 16 de West e Harrison (1997), sendo útil no estudo de dados composicionais.

2.1 Modelo Linear Dinâmico Normal Matriz-Variado

No caso do Modelo Linear Dinâmico Normal Matriz-Variado, deseja-se trabalhar com um conjunto de séries temporais e, ao mesmo tempo, estabelecer uma estrutura de covariância entre elas. Para compreender melhor, suponha que se deseja investigar, conjuntamente, q séries temporais univariadas. Esta formulação é útil no estudo de dados composicionais (Quintana e West, 1988).

Considere que as q séries univariadas Y_{tj} , $j = 1, \dots, q$, são tais que Y_{tj} segue um *modelo dinâmico* definido por:

$$\{F_t, G_t, V_t\sigma_j^2, W_t\sigma_j^2\}, \quad j = 1, 2, \dots, q$$

assumindo todos estes elementos conhecidos, exceto σ_j^2 , $j = 1, \dots, q$.

As séries univariadas são descritas da forma já conhecida:

- *Equação de observação:*

$$Y_{tj} = F_t'\theta_{tj} + \epsilon_{tj}, \quad \epsilon_{tj} \sim N(0, V_t\sigma_j^2).$$

- *Equação de evolução:*

$$\theta_{tj} = G_t \theta_{t-1,j} + \omega_{tj}, \quad \omega_{tj} \sim N(0, W_t \sigma_j^2).$$

- F_t e θ_t são vetores de dimensão $n \times 1$ (quando $n = 1$, trata-se de um modelo de 1ª ordem);
- G_t é uma matriz $n \times n$ que define a evolução de θ_t .
- $\theta_0 | D_0 \sim N_n(m_0, C_0)$;

Note que, para as q séries de tempo, F_t, G_t, V_t e W_t permanecem os mesmos, ou seja, as q séries apresentam a mesma estrutura de evolução ao longo do tempo, alterando apenas os seus parâmetros (os valores de estado θ_{tj}) e as escalas σ_j^2 .

Assume-se que para todo j , os erros observacionais ϵ_{tj} são independentes no tempo, assim como os erros de evolução ω_{tj} . Para quaisquer j e t , as sequências ϵ_{tj} e ω_{tj} são, também, independentes.

A estrutura de covariância entre as séries, tanto para os erros observacionais ϵ_{tj} quanto para os erros de evolução ω_{tj} , é determinada pela matriz Σ de dimensão $q \times q$, dada por:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \cdots & \sigma_{1,q} \\ \sigma_{2,1} & \sigma_2^2 & \cdots & \sigma_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,q} & \sigma_{2,q} & \cdots & \sigma_q^2 \end{pmatrix}$$

Com a matriz Σ definida, segue que a covariância entre os erros é dada por:

$$Cov(\epsilon_{tj}, \epsilon_{ti}) = V_t \sigma_{ji}$$

$$Cov(\omega_{tj}, \omega_{ti}) = W_t \sigma_{ji}$$

para $i \neq j$ e $V_t > 0$.

2.1.1 Estrutura do Modelo Matricial

Com estas estruturas definidas, o modelo pode ser generalizado para a seguinte forma matricial:

- $Y_t = (Y_{t1}, \dots, Y_{tq})'$, com dimensão $q \times 1$.
- $\epsilon_t = (\epsilon_{t1}, \dots, \epsilon_{tq})'$, com $\epsilon_{t,(q \times 1)}$ um vetor $q \times 1$ de erros observacionais no tempo t .
- $\Theta_t = (\theta_{t1}, \dots, \theta_{tq})'$, com $\Theta_{t,(n \times q)}$ uma matrix $n \times q$ cujas colunas são os vetores de estados dos modelos individuais.
- $\Omega_t = (\omega_{t1}, \dots, \omega_{tq})'$, com $\Omega_{t,(n \times q)}$ uma matrix $n \times q$ cujas colunas são os erros de evolução dos modelos individuais.

Assim, as equações de observação e evolução são expressas por:

- $Y_t' = F_t' \Theta_t + \epsilon_t'$
- $\Theta_t = G_t \Theta_{t-1} + \Omega_t$

Naturalmente, nota-se que, no caso de $q = 1$, tem-se um modelo linear dinâmico convencional com *fator de escala observacional* desconhecido, $\Sigma = \sigma_1^2$. No caso em que $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$, então as q séries Y_{tj} não são relacionadas.

2.1.2 A Distribuição do vetor ϵ_t :

ϵ_t é um vetor aleatório formado pelos escalares ϵ_{tj} com $j = 1, \dots, q$. Sabe-se que cada componente ϵ_{tj} possui média zero e a relação de covariância entre eles é determinada pela matrix Σ . Adicionalmente, admite-se que todos os elementos de ϵ_t possuem uma escala comum V_t , portanto segue que:

$$\epsilon_{tj} \sim N_q(0, V_t \Sigma).$$

2.1.3 A Distribuição de Ω_t :

A matrix aleatória Ω_t é uma matrix composta por q colunas de dimensão $n \times 1$ com distribuição *normal multivariada* (Gupta, 1992, Khoan e Truc, 1994) com vetor de médias zero e variância estabelecida pela matrix W_t . Ao mesmo tempo, estes q vetores estão relacionados de modo que a covariância entre eles é estabelecida pela matrix Σ . Desta forma, a matrix aleatória Ω_t segue distribuição normal matrix-variada com

matriz de médias $0_{n \times n}$. A relação entre as linhas da matriz é determinada pela *matriz de covariância à esquerda*, W_t , enquanto a relação entre as colunas é determinada pela *matriz de covariância à direita* Σ . A notação é dada por:

$$\Omega_t \sim N_{n \times q}(0, W_t, \Sigma).$$

A função de densidade é dada por:

$$P(\Omega_t) = K(W_t, \Sigma) \exp \left\{ -\frac{1}{2} \text{tr}[\Omega_t' W_t^{-1} \Omega_t \Sigma^{-1}] \right\} \quad (2.1)$$

onde,

$$K(W_t, \Sigma) = (2\pi)^{-qn/2} |W_t|^{-q/2} |\Sigma|^{-n/2}.$$

Em relação à distribuição *normal matriz-variada*, vale mencionar algumas de suas propriedades:

- Todas as marginais e distribuições condicionais dos elementos de Ω_t , assim como as funções lineares destas, são distribuições normais univariadas, multivariadas ou matriz-variadas.
- A definição das distribuições permanece válida ainda que cada uma, ou mesmo ambas as matrizes W_t e Σ , sejam não negativas definidas.
- A distribuição é não singular se e somente se cada um das matrizes de covariância é positiva definida.
- Se W_t ou Σ é uma matriz nula, então $P(\Omega_t = 0) = 1$.
- Se $\Omega_t \sim N_{n \times q}(0, W_t, \Sigma)$ então, a “vetorização” de Ω_t segue uma distribuição normal multivariada dada por:

$$\text{vec}(\Omega_t) \sim N_{nq}(0, W_t \otimes \Sigma).$$

- Seja $\Omega_t^* = (w_{t1}^*, \dots, w_{tq}^*) \sim N(H_t, W_t, \Sigma)$,
em que $H_t = (h_{t1}, \dots, h_{tq})$, portanto $E(w_{tj}^*) = h_{tj}$ para $j = 1, \dots, q$. Então,

$$\Omega_t = \Omega_t^* - H_t \sim N(0, W_t, \Sigma).$$

Uma rotina em **R** para a simulação de dados de uma distribuição normal matriz-variada pode ser encontrada no Apêndice F.

2.1.4 A distribuição de Σ :

Em geral, os valores da matriz Σ não são conhecidos. Desta forma, pode-se atribuir à matriz aleatória Σ a distribuição *Wishart Invertida*.

A distribuição *Wishart Invertida* é descrita por:

$$P(\Sigma) \propto |\Sigma|^{-(q+n/2)} \exp \left\{ -\frac{1}{2} \text{tr}(nS\Sigma^{-1}) \right\} \quad (2.2)$$

com $n > 0$ graus de liberdade e $S_{q \times q}$ é uma matriz positiva definida.

Segue a notação: $\Sigma \sim W_n^{-1}[S]$ com $E(\Sigma) = \left(\frac{n}{n-2}S\right)$; $n > 2$.

Se $\Phi = \Sigma^{-1}$, então $\Phi \sim W_n$ com $E(\Phi) = S^{-1}$.

Se Σ segue uma *Wishart Invertida*, então cada elemento σ_{ij} de Σ segue uma distribuição *Gama Inversa*. A relação entre as distribuições *Gama Inversa* e *Normal* tem sido muito explorada na inferência bayesiana, devido à tratabilidade analítica que esta relação permite. Da mesma forma, ao expandir a dimensão das variáveis aleatórias, esta relação se estende entre as distribuições *Wishart Invertida* e *Normal Matriz-Variada*.

2.1.5 A Distribuição Conjunta de (Θ_t, Σ) :

Atribuindo $\Sigma \sim W_n^{-1}[S]$ e, condicionalmente a Σ conhecido, $(\Theta|\Sigma) \sim N(m, C, \Sigma)$ onde m tem dimensão $n \times q$, C e Σ com dimensão $p \times p$, segue que:

$$P(\Theta, \Sigma) \stackrel{d}{=} P(\Theta|\Sigma)P(\Sigma).$$

Esta distribuição é denominada *Normal Matriz-Variada Wishart Invertida* e é definida pela notação:

$$(\Theta, \Sigma) \sim NW_n^{-1}(m, C, S).$$

Analogamente ao caso univariado, a distribuição marginal da matriz Θ_t segue distribuição *matriz t-Student*, de modo que para

$$\Theta = (\theta_1, \dots, \theta_q) \quad \text{e} \quad m = (m_1, \dots, m_q),$$

tem-se

$$\theta_j \sim T(m_j, CS_{jj}), \quad j = 1, \dots, q.$$

Logo,

$$\Theta \sim T_n(m, C, S).$$

onde $E(\theta_j) = m_j$ para $n > 1$; $V(\theta_j) = CS_{jj}\frac{n}{n-2}$ para $n > 2$ e $Cov(\theta_i, \theta_j) = CS_{ij}\frac{n}{n-2}$.

2.2 Processo de Estimação

Seguindo a mesma lógica do caso univariado, as previsões a 1 passo são obtidas por:

1. Posteriori em $t - 1$: $(\Theta_{t-1}, \Sigma | D_{t-1}) \sim NW_{n_{t-1}}^{-1}(m_{t-1}, C_{t-1}, S_{t-1})$

onde $m_{t-1}, C_{t-1}, S_{t-1}$ e n_{t-1} já são conhecidos.

2. Priori em t : $(\Theta_t, \Sigma | D_{t-1}) \sim N(a_t, R_t, S_{t-1})$

onde:

$$a_t = G_t m_{t-1}; \quad e \quad R_t = G_t C_{t-1} G_t' + W_t.$$

3. Previsão 1 passo a frente: $(Y_t | \Sigma, D_{t-1}) \sim N(f_t, Q_t \Sigma)$

onde:

$$f_t = F_t' a_t \quad e \quad Q_t = F_t' R_t F_t + V_t$$

com marginal $(Y_t | D_{t-1}) \sim T_{n_{t-1}}(f_t, Q_t S_{t-1})$.

4. Posteriori em t : $(\Theta_t, \Sigma | D_t) \sim NW_{n_t}^{-1}(m_t, C_t, S_t)$

onde:

$$m_t = a_t + A_t e_t'; \quad A_t = R_t F_t Q_t^{-1}; \quad e_t = (Y_t - f_t)$$

$$R_t = A_t A_t' Q_t; \quad n_t = n_{t-1} + 1 \quad e$$

$$S_t = n_t^{-1} [n_{t-1} S_{t-1} + e_t e_t' Q_t^{-1}].$$

A demonstração do algoritmo segue os mesmos passos do caso da série univariada, lembrando que as variáveis com distribuição *normal matriz-variada* podem ser vetorizadas em Normais multivariadas. A evolução de S_t segue também do *Teorema de Bayes*, de modo que:

$$P(\Phi | D_{t-1}) \propto |\Phi|^{(n_{t-1}/2 - q)} \exp \left\{ -\frac{1}{2} \text{tr}(n_{t-1} S_{t-1} \Phi) \right\} \quad (2.3)$$

$$P(Y_t|D_{t-1}, \Phi) \propto |\Phi|^{q/2} \exp \left\{ -\frac{1}{2} e_t' \Phi Q_t^{-1} e_t \right\} \quad (2.4)$$

Por definição do teorema de bayes:

$$P(\Phi|D_t) \propto P(\Phi|D_{t-1})P(Y_t|D_{t-1}, \Phi),$$

onde

$$P(\Phi|D_t) \propto |\Phi|^{[(n_{t-1}+1)/2-q]} \exp \left\{ -\frac{1}{2} \text{tr}[(n_{t-1}S_{t-1} + e_t e_t' Q_t^{-1})\Phi] \right\}. \quad (2.5)$$

Portanto, substituindo $n_t = n_{t-1} + 1$ e $S_t = n_t^{-1}[n_{t-1}S_{t-1} + e_t e_t' Q_t^{-1}]$, obtém-se:

$$P(\Phi|D_t) \propto |\Phi|^{[n_t/2-q]} \exp \left\{ -\frac{1}{2} \text{tr}[n_t S_t \Phi] \right\}. \quad (2.6)$$

2.3 Aplicação a Dados Composicionais

A análise do comportamento relativo de séries de tempo com dados composicionais pode, frequentemente, ser simplificada ao converter os dados para séries de proporções relativas.

Para compreender melhor, suponha uma série de tempo S_t seja formada por q séries que compõem o vetor $z_t = (z_{t1}, \dots, z_{tq})$, ou seja:

$$S_t = \mathbf{1}' z_t, \quad \forall t$$

de modo que as q séries z_{tj} possuem natureza similar.

A série temporal das *proporções* é obtida da forma:

$$p_t = (\mathbf{1}' z_t)^{-1} z_t = \frac{z_t}{\sum_{j=1}^q z_{tj}} \quad (2.7)$$

2.4 Transformação Razão Log e Distribuição Logística-Normal

2.4.1 Escala Razão-Log

Ao trabalhar com todas as séries q , simultaneamente, a restrição $\sum_{j=1}^q p_{tj} = 1$ deve ser considerada na análise. Em várias áreas Aitchison (1982, 1986) descreve

o uso de transformações *razão-log* para a análise de proporções. Para tratar deste problema especificamente, West e Harrison (1997) sugerem uma versão simétrica da transformação razão-log, dada por:

$$Y_{tj} = \log\left(\frac{p_{tj}}{\tilde{p}_t}\right) = \log(p_{tj}) - \log(\tilde{p}_t), \quad j = 1, \dots, q. \quad (2.8)$$

onde

$$\tilde{p}_t = \prod_{j=1}^q p_{tj}^{1/q}$$

é a média geométrica dos p_{tj} 's e $Y_{tj} \in \mathbb{R}^{q-1}$.

A transformação inversa é dada por:

$$p_{tj} = \frac{\exp(Y_{tj})}{\sum_{i=1}^q \exp(Y_{ti})}. \quad (2.9)$$

Trabalhar com esta transformação é interessante devido à relação entre a distribuição normal e a distribuição *logística-normal* (vide Seção 2.4.2), uma vez que se observa a seguinte propriedade:

$$Y_t \sim N(\mu, \Sigma) \Rightarrow p_t \sim LN(\mu, \Sigma).$$

Observa-se, também, que:

$$\sum_{j=1}^q Y_{tj} = \sum_{j=1}^q \log\left(\frac{p_{tj}}{\tilde{p}_t}\right) = \sum_{j=1}^q \log p_{tj} - q \log \tilde{p}_t = 0.$$

Portanto, modelar diretamente Y_t implica em Σ singular.

Para lidar com o problema de singularidade em Σ , West e Harrison (1997) sugerem que a restrição $\sum_{j=1}^q Y_{tj} = 0$ seja imposta, diretamente no modelo, através da transformação $H_t = KY_t$, onde:

$$K = K' = I_q - q^{-1}\mathbf{1}\mathbf{1}'.$$

Ao modelar H_t no lugar de Y_t , tal procedimento resolve o problema de singularidade e as equações passam a ser descritas por:

- *Equação de observação:*

$$H_t' = Y_t'K = F_t'\Psi_t + (K\epsilon_t)', \quad \epsilon_t \sim N(0, V_t\Delta^{-1})$$

- *Equação de evolução:*

$$\Psi_t = G_t \Psi_{t-1} + \Omega_t K, \quad \Omega_t K \sim N(0, W_t, \Delta^{-1})$$

onde $\Psi_t = \Theta_t K$ e $\Delta = K \Sigma K$.

Similarmente,

$$(\Psi_0, \Delta | D_0) \sim NW_{n_0}^{-1}(m_0 K, C_0, K S_0 K).$$

Cabe ressaltar que:

- O procedimento de estimação do modelo restrito segue os mesmos passos descritos para o modelo matricial.
- As quantidades F_t, G_t, V_t, W_t e C_t permanecem inalteradas pelas transformações lineares impostas por K . Com isso, o uso de *fatores de descontos* em W_t não é afetado pelas transformações.

2.4.2 Distribuição Logística-Normal

A distribuição *Logística Normal* foi descrita por Aitchison e Shen (1980) e destacou-se por apresentar uma relação direta com a distribuição *Normal*. Se p_t é um vetor aleatório com distribuição *Logística Normal*, $p_t \sim LN_{q-1}(\mu_t, \Sigma_t)$, sua função de densidade é definida por:

$$f(p_t | \mu_t, \Sigma_t) = \left(\frac{1}{2\pi} \right)^{\frac{q-1}{2}} |\Sigma_t|^{-\frac{1}{2}} \left(\frac{1}{\prod_{i=1}^q p_{it}} \right) \exp \left\{ -\frac{1}{2} (p_t - \mu_t)' \Sigma_t^{-1} (p_t - \mu_t) \right\}, \quad (2.10)$$

Seja v_t o vetor aleatório obtido pelo simplex em p_t da seguinte forma:

$$v_t = alr(p_t) = \log \left(\frac{p_{t,-q}}{p_{tq}} \right) = \left[\log \left(\frac{p_{t1}}{p_{tq}} \right), \dots, \log \left(\frac{p_{tq-1}}{p_{tq}} \right) \right]' \in \mathbb{R}^{q-1}, \quad (2.11)$$

com $p_t \in \nabla^{q-1}$ (espaço no simplex) e $p_{tq} = 1 - (p_{t1} + \dots + p_{tq-1})$.

A partir desta transformação, tem-se:

$$v_t \sim N_{k-1}(\mu_t, \Sigma_t).$$

Esta relação torna-se muito útil para a modelagem de dados composicionais. Cabe observar que a transformação conserva os parâmetros para as duas distribuições e a

transformação inversa é obtida facilmente da forma:

$$p_t = alr^{-1}(v_t) = \left[\frac{e^{v_t1}}{1 + \sum_{i=1}^{q-1} e^{v_{ti}}}, \dots, \frac{e^{v_{t(q-1)}}}{1 + \sum_{i=1}^{q-1} e^{v_{ti}}}, \frac{1}{1 + \sum_{i=1}^{k-1} e^{v_{ti}}} \right]. \quad (2.12)$$

Com estas transformações, é possível construir um modelo dinâmico de previsão para dados composicionais utilizando-se a teoria baseada na distribuição normal através da transformação $y_t = alr(\cdot)$. Retorna-se à escala de proporção dos valores estimados por meio da transformação inversa $alr^{-1}(\cdot)$.

Na Seção 6.6 apresenta-se uma aplicação do modelo aqui descrito para dados composicionais.

Capítulo 3

Modelo Linear Dinâmico

Generalizado

Nesta abordagem, a suposição de que as observações sejam *normalmente distribuídas* é relaxada. De forma bem mais abrangente, assume-se, simplesmente, que a distribuição das observações seja membro da *família exponencial* de Nelder e Wedderburn (1972), isto é:

$$P(Y_t|\eta_t) = \exp\{\phi_t^{-1}[y_t\eta_t - b(\eta_t)] + c(y_t, \phi_t)\}$$

onde η_t é o *parâmetro natural* e $\phi_t = V_t^{-1}$ é a precisão da distribuição. Neste caso, η_t é o parâmetro que tem uma relação linear com o vetor de estados θ_t , mas a equação de evolução permanece a mesma em relação ao modelo normal, exceto pela distribuição dos erros. A distribuição normal faz parte desta família, portanto o modelo normal é um caso particular do modelo dinâmico generalizado.

O modelo dinâmico generalizado é especificado por:

- *Função Média:* $\mu_t = E(y_t|\eta_t) = b'(\eta_t) = \frac{\partial b(\eta_t)}{\partial \eta_t}$;
- *Função de ligação:* $\eta_t = g(\mu_t) = F_t'\theta_t$;
- *Equação do sistema:* $\theta_t = G_t\theta_{t-1} + \omega_t$;
- *Erro do sistema:* $\omega_t \sim (0, W_t)$ (observa-se que somente os momentos estão definidos, mas não a distribuição de ω_t);

- *Informação a priori:* $\theta_1 \sim (a_1, R_1)$

(da mesma forma, não é imposta uma distribuição específica para θ_1).

Observa-se que, no caso da distribuição normal, o parâmetro natural η_t é a própria média μ_t , logo $g(\mu_t) = \mu_t$.

A sequência de procedimentos utilizada na estimação do modelo é resumida a seguir.

- Posteriori em $t - 1$: $(\theta_{t-1}|D_{t-1}) \sim (m_{t-1}, C_{t-1})$, e

- Priori em t : $(\theta_t|D_{t-1}) \sim (a_t, R_t)$

onde $a_t = G_t m_{t-1}$ e $R_t = G_t C_{t-1} G_t' + W_t$.

- A priori para η_t é dada por: $(\eta_t|D_{t-1}) \sim (f_t, q_t)$.

No caso em que a priori de η_t é conjugada, tem-se que priori e posteriori pertencem na mesma família e portanto $(\eta_t|D_t) \sim (f_t^*, q_t^*)$.

- A distribuição conjunta de $(\eta_t, \theta_t|D_{t-1})$, considerando apenas os momentos, é dada por:

$$(\eta_t, \theta_t|D_{t-1}) \sim \left[\begin{pmatrix} f_t \\ a_t \end{pmatrix}, \begin{pmatrix} q_t & F_t' R_t \\ R_t F_t & R_t \end{pmatrix} \right],$$

onde $f_t = F_t' a_t$ e $q_t = F_t' R_t F_t$.

- Uma vez conhecidos f_t e q_t , os parâmetros da priori conjugada de μ_t são obtidos pela relação:

$$f_t = E[g(\mu_t)|D_{t-1}] \quad \text{e} \quad q_t = V[g(\mu_t)|D_{t-1}]$$

- Os momentos associados a $(\theta_t|\eta_t, D_{t-1})$ são estimados via *Linear Bayes* (vide West e Harrison, 1997 e apêndice E):

$$\hat{E}(\theta_t|\eta_t, D_{t-1}) = a_t + R_t F_t (\eta_t - f_t) / q_t$$

$$\hat{V}(\theta_t|\eta_t, D_{t-1}) = R_t - R_t F_t F_t' R_t / q_t.$$

- Após atualizada a posteriori $(\mu_t|D_t)$ e obtidos $E[\mu_t|D_t]$ e $V[\mu_t|D_t]$, é mantida a relação:

$$f_t^* = E[g(\mu_t)|D_t] \quad \text{e} \quad q_t^* = V[g(\mu_t)|D_t],$$

com $(\eta_t|D_t) \sim (f_t^*, q_t^*)$.

- A posteriori em t , $(\theta_t|D_t)$, segue de:

$$p(\theta_t|D_t) = \int P(\theta_t|\eta_t, D_{t-1})P(\eta_t|D_t)d\eta_t$$

onde $P(\theta_t|\eta_t, D_{t-1})$ não é conhecida, mas os momentos de $(\theta_t|D_t)$ podem ser estimados por:

$$\begin{aligned} m_t &= E(\theta_t|D_t) \\ &= E[\hat{E}(\theta_t|\eta_t, D_{t-1})|D_t] \\ &= a_t + R_t F_t (f_t^* - f_t)/q_t \end{aligned}$$

e

$$\begin{aligned} C_t &= V(\theta_t|D_t) \\ &= V[\hat{E}(\theta_t|\eta_t, D_{t-1})|D_t] + E[\hat{V}(\theta_t|\eta_t, D_{t-1})|D_t] \\ &= R_t - R_t F_t F_t' R_t (1 - q_t^*/q_t)/q_t. \end{aligned}$$

3.1 Exemplo: Modelo Dinâmico Binomial

Seja $Y_t \sim Bin(n_t, \pi_t)$

$$P(Y_t|\pi_t) = \binom{n_t}{y_t} \pi_t^{y_t} (1 - \pi_t)^{n_t - y_t} \quad (3.1)$$

$$\begin{aligned} &= \binom{n_t}{y_t} \left(\frac{\pi_t}{1 - \pi_t} \right)^{y_t} (1 - \pi_t)^{n_t} \\ &= \exp \left\{ y_t \log \left(\frac{\pi_t}{1 - \pi_t} \right) + n_t \log(1 - \pi_t) \right\} \binom{n_t}{y_t} \end{aligned} \quad (3.2)$$

onde

$$\eta_t = \log \left(\frac{\pi_t}{1 - \pi_t} \right). \quad \text{Logo, } \pi_t = \frac{e^{\eta_t}}{1 + e^{\eta_t}}.$$

$P(Y_t|\eta_t)$ pode ser reescrito como:

$$\begin{aligned} P(Y_t|\eta_t) &= \binom{n_t}{y_t} \exp\{y_t \eta_t - n_t \log(1 + e^{\eta_t})\} \\ &= h(y_t) \exp\{\phi_t^{-1}[y_t \eta_t - b(\eta_t)] + c(y_t, \phi_t)\} \end{aligned}$$

com $b(\eta_t) = n_t \log(1 + e^{\eta_t})$; $\mu_t = E(Y_t) = b'(\eta_t) = n_t \pi_t$ e $\phi_t = 1$. Logo,

$$\eta_t = \log \left(\frac{\mu_t}{n_t - \mu_t} \right); \quad \mu_t = n_t \left(\frac{e^{\eta_t}}{1 + e^{\eta_t}} \right).$$

Com isso, a função $g(\eta_t)$ é tal que $g : \mathbb{R} \rightarrow \mathbb{R}$ é a própria *função identidade*, com $g(\eta_t) = \eta_t$, isto é,

$$g(\eta_t) = \eta_t = \log \left(\frac{\pi_t}{1 - \pi_t} \right) = F'_t \theta_t. \quad (3.3)$$

A única incerteza em relação à distribuição de Y_t , dado o passado D_{t-1} , é devido à incerteza com respeito a η_t .

A priori conjugada também pertence à família exponencial. Portanto, tem densidade da forma:

$$P(\eta_t | D_{t-1}) = c(r_t, s_t) \exp[r_t \eta_t - s_t b(\eta_t)].$$

Os parâmetros s_t e r_t da priori conjugada devem satisfazer as condições dos momentos de 1ª e 2ª ordem da priori ($g(\eta_t) | D_{t-1}$).

Desta forma, tem-se:

$$\begin{aligned} f_t &= E(\eta_t | D_{t-1}) = E \left[\log \left(\frac{\mu_t}{1 - \mu_t} \right) \middle| D_{t-1} \right] \\ &= E(\log(\mu_t) | D_{t-1}) - E(\log(1 - \mu_t) | D_{t-1}) \\ &= \psi(s_t) - \psi(r_t), \end{aligned}$$

e

$$\begin{aligned} q_t &= V(\eta_t | D_{t-1}) = V(\log(\mu_t) | D_{t-1}) + V(\log(1 - \mu_t) | D_{t-1}) \\ &= \psi'(s_t) + \psi'(r_t), \end{aligned} \quad (3.4)$$

onde $\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}$, é a função digama. Com isso, dados f_t e q_t , define-se os parâmetros da priori conjugada $Beta(r_t, s_t)$, a ser explicitada a seguir.

Para valores elevados de s_t e r_t , $\psi \approx \log(z)$, enquanto $\psi'(z) \approx z^{-1}$ (West e Harrison, 1997). Desta forma,

$$f_t \approx \log \left(\frac{s_t}{r_t} \right) \quad \text{e} \quad q_t \approx \frac{1}{s_t} + \frac{1}{r_t}.$$

Então,

$$s_t \approx [1 + \exp(f_t)]/q_t \quad \text{e} \quad r_t \approx [1 + \exp(-f_t)]/q_t.$$

Reparametrizando o modelo $P(Y_t|\pi_t)$ em função da média μ_t :

$$\begin{aligned} P(Y_t|\mu_t) &\propto \mu_t^{y_t}(1 - \mu_t)^{n_t - y_t} \\ &\propto \exp[y_t \log \mu_t + (n_t - y_t) \log(1 - \mu_t)], \end{aligned}$$

a priori conjugada pode ser reescrita, também, em função de μ_t , como:

$$\begin{aligned} P(\mu_t|D_{t-1}) &\propto \exp(s_t^* \log \mu_t + r_t^* \log(1 - \mu_t)) \\ &\propto \mu_t^{s_t^*} (1 - \mu_t)^{r_t^*} \\ &\propto \mu_t^{s_t - 1} (1 - \mu_t)^{r_t - 1}. \end{aligned}$$

Logo, $(\mu_t|D_{t-1}) \sim \text{Beta}(s_t, r_t)$.

Portanto, a posteriori $P(\mu_t|D_t)$ é dada por:

$$\begin{aligned} P(\mu_t|D_t) &\propto \mu_t^{y_t}(1 - \mu_t)^{n_t - y_t} \mu_t^{s_t - 1} (1 - \mu_t)^{r_t - 1} \\ &\propto \mu_t^{y_t + s_t - 1} (1 - \mu_t)^{n_t + r_t - y_t - 1}. \end{aligned}$$

Logo,

$$(\mu_t|D_t) \sim \text{Beta}(y_t + s_t; n_t + r_t - y_t).$$

A previsão a 1 passo a frente é obtida por:

$$P(Y_t|D_{t-1}) = \int P(Y_t|\eta_t, D_{t-1})P(\eta_t|D_{t-1})d\eta_t.$$

Utilizando a parametrização em μ_t :

$$\begin{aligned} P(Y_t|D_{t-1}) &= \int P(Y_t|\mu_t, D_{t-1})P(\mu_t|D_{t-1})d\mu_t \\ &\propto \int \mu_t^{y_t + s_t - 1} (1 - \mu_t)^{n_t + r_t - y_t - 1} d\mu_t \\ &= \binom{n_t}{y_t} \frac{\Gamma(s_t + r_t)}{\Gamma(s_t)\Gamma(r_t)} \frac{\Gamma(y_t + s_t)\Gamma(n_t + r_t - y_t)}{\Gamma(n_t + r_t + s_t)}. \end{aligned} \quad (3.5)$$

Desta forma, $(Y_t|D_{t-1})$ segue distribuição *Beta-Binomial*.

Em resumo:

- Priori para μ_t : $(\mu_t|D_{t-1}) \sim \text{Beta}(s_t, r_t)$.
- Previsão a 1 passo:

$$P(Y_t|D_{t-1}) = \binom{n_t}{y_t} \frac{\Gamma(s_t + r_t)}{\Gamma(s_t)\Gamma(r_t)} \frac{\Gamma(y_t + s_t)\Gamma(n_t + r_t - y_t)}{\Gamma(n_t + r_t + s_t)}.$$

- Posteriori para μ_t : $(\mu_t|D_t) \sim \text{Beta}(y_t + s_t; n_t + r_t - y_t)$.

No próximo capítulo, discutiremos modelos dinâmicos Beta que foram desenvolvidos durante os trabalhos de pesquisa desta dissertação.

Capítulo 4

Modelo Dinâmico Beta com priori conjugada e ϕ conhecido

Na Seção 2.4 discutiu-se a abordagem dada por West e Harrison (1997) na modelagem de dados composicionais. Como descrito, o uso da transformação razão-log e o pressuposto de normalidade dos erros na escala original dos dados é imprescindível em tal abordagem. Neste Capítulo descreve-se um modelo dinâmico para séries de proporções para o qual não se faz necessário o uso de transformações.

Com base na metodologia descrita para *Modelos Lineares Dinâmicos Generalizados* (vide Capítulo 3), deseja-se, inicialmente, construir um procedimento para modelar as proporções ao longo do tempo, considerando o caso de uma série Y formada por duas componentes, de modo que $Y \sim Beta(p, q)$. Portanto, seja

$$f_Y(y|p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1}; \quad 0 < y_t < 1, \quad p, q > 0.$$

Desta forma,

$$E(Y) = \frac{p}{p+q} \quad \text{e} \quad V(Y) = \frac{pq}{(p+q)^2(p+q+1)}.$$

4.1 Nova parametrização

Utilizando a parametrização proposta por Ferrari e Cribari-Neto (2004), tem-se:

$$\mu = \frac{p}{p+q} \quad \text{e} \quad \phi = p+q \Rightarrow p = \mu\phi \quad \text{e} \quad q = \phi(1-\mu).$$

Logo,

$$f_Y(y|\mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad (4.1)$$

onde $0 < y < 1$; $0 < \mu < 1$; $\phi > 0$.

Portanto, a distribuição de y pode ser reparametrizada na forma:

$$(Y|\mu, \phi) \sim \text{Beta}(\mu\phi, (1-\mu)\phi),$$

onde $E(Y|\mu, \phi) = \mu$ e $V(Y|\mu, \phi) = \frac{\mu(1-\mu)}{1+\phi}$.

Considere, agora, $f_Y(y|\mu, \phi)$ expresso em termos da família exponencial:

$$\begin{aligned} f_Y(y|\mu, \phi) &= \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1} \\ &= \exp\{(\mu\phi-1)\log y + [(1-\mu)\phi-1]\log(1-y) \\ &\quad + \log \Gamma(\phi) - \log \Gamma(\mu\phi) - \log \Gamma((1-\mu)\phi)\} \\ &= h(y) \exp\{\mu\phi[\log y - \log(1-y)] + \phi \log(1-y) \\ &\quad + \log \Gamma(\phi) - \log \Gamma(\mu\phi) - \log \Gamma((1-\mu)\phi)\}. \end{aligned}$$

Considere ϕ conhecido, então,

$$\begin{aligned} f_Y(y|\mu, \phi) &= h^*(y) \exp\left\{\mu \left[\phi \log \left(\frac{y}{1-y}\right)\right] + \log \left[\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)}\right]\right\} \\ &= h^*(y) \exp\{t(y)\mu + a(\mu)\}, \end{aligned} \quad (4.2)$$

com

$$h^*(y) = (1-y)^{(\phi-1)} y^{-1}, \quad t(y) = \phi \log \left[\frac{y}{1-y}\right], \quad e$$

$$a(\mu) = \log \left[\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)}\right].$$

Considere o caso de uma priori conjugada de μ e sua correspondente posteriori:

- *Priori conjugada de μ :*

$$P(\mu | \phi) = c(r, s) \exp\{r\mu + sa(\mu)\}, \quad (4.3)$$

onde

$$c(r, s)^{-1} = \int_0^1 \left[\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)}\right]^s e^{r\mu} d\mu.$$

- *Posteriori de μ :*

$$\begin{aligned} P(\mu | D, \phi) &\propto f_Y(y; \mu, \phi)P(\mu) \\ &\propto \exp\{t(y)\mu + a(\mu)\} \exp\{r\mu + sa(\mu)\} \\ &\propto \exp\{\mu(r + t(y)) + a(\mu)[s + 1]\}. \end{aligned}$$

Logo,

$$P(\mu | D, \phi) = c((r + t(y)), s + 1) \left[\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1 - \mu)\phi)} \right]^{s+1} e^{\mu(r+t(y))}, \quad (4.4)$$

onde

$$c((r + t(y)), s + 1)^{-1} = \int_0^1 \left[\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1 - \mu)\phi)} \right]^{s+1} e^{\mu(r+t(y))} d\mu.$$

Vale ressaltar que $(\mu | \phi)$ e $(\mu | D, \phi)$ não têm forma conhecida.

4.2 Contextualizando para o modelo dinâmico

- *Equação das observações:*

$$f(y_t | \mu_t, \phi) = h^*(y_t) \exp\{\mu_t t(y_t) + a(\mu_t)\}$$

com ϕ conhecido, $t(y_t) = \phi \log \left[\frac{y_t}{1 - y_t} \right]$ e $a(\mu_t) = \log \left[\frac{\Gamma(\phi)}{\Gamma(\mu_t\phi)\Gamma((1 - \mu_t)\phi)} \right]$,

$$g(\mu_t) = \log \left[\frac{\mu_t}{1 - \mu_t} \right] = \lambda_t = F'_t \beta_t \quad \text{e} \quad h^*(y) = (1 - y)^{(\phi-1)} y^{-1}.$$

A função de ligação natural é a função identidade, $g(\mu_t) = \mu_t$. No entanto, $\mu_t \in (0, 1)$, sendo mais adequado o uso de uma função de ligação logito, que leva μ_t em \mathbb{R} . Partindo desta transformação, o processo de estimação pode ser realizado, via *Linear Bayes*, da seguinte forma:

- *Equação de evolução:*

$$\beta_t = G_t \beta_{t-1} + w_t \quad \text{com} \quad w_t \sim [0, W_t].$$

- *Informação inicial:*

$$(\beta_0 | D_0, \phi) \sim (m_0, C_0).$$

- *Passo 1: Priori de $\lambda_t = F_t' \beta_t$:*

Tome,

$$\begin{pmatrix} \lambda_t \\ \beta_t \end{pmatrix} \Big| D_{t-1} \sim \left[\begin{pmatrix} f_t \\ a_t \end{pmatrix}, \begin{pmatrix} q_t & F_t' R_t \\ R_t F_t & R_t \end{pmatrix} \right]$$

onde

$$f_t = F_t' a_t; \quad q_t = F_t' R_t F_t; \quad a_t = G_t m_{t-1}; \quad R_t = G_t C_{t-1} G_t' + W_t.$$

Então,

$$(\lambda_t | D_{t-1}, \phi) \sim (f_t, q_t) \quad e \quad (\beta_t | D_{t-1}, \phi) \sim (a_t, R_t).$$

Considerando-se já conhecidos os momentos de

$$(\beta_{t-1} | D_{t-1}) \sim (m_{t-1}, C_{t-1}).$$

- *Passo 2: Previsão de 1 passo a frente*

a) Especificação da priori de μ_t , $P(\mu_t | D_{t-1}, \phi)$:

Seja:

$$\begin{aligned} P(\mu_t | D_{t-1}, \phi) &= c(r_t, s_t) \exp\{r_t \mu + s_t a(\mu)\} \\ &= c(r_t, s_t) \left[\frac{\Gamma(\phi)}{\Gamma(\mu_t \phi) \Gamma((1 - \mu_t) \phi)} \right]^{s_t} \exp\{\mu_t r_t\}, \end{aligned} \quad (4.5)$$

com $c(r_t, s_t)$ como em $P(\mu | \phi)$ e onde s_t e r_t são especificados de modo a satisfazerem as equações:

$$E[g(\mu_t) | D_{t-1}, \phi] = f_t \quad e \quad V[g(\mu_t) | D_{t-1}, \phi] = q_t.$$

b) Relacionando $E(\mu_t | D_{t-1})$ e $V(\mu_t | D_{t-1})$ com $E(\lambda_t | D_{t-1}, \phi)$ e $V(\lambda_t | D_{t-1}, \phi)$:

$$E(\mu_t | D_{t-1}, \phi) = E \left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \Big| D_{t-1} \right].$$

No cálculo acima, a expansão da função de ligação em *série de Taylor* é uma alternativa para obter a relações entre os momentos (os métodos utilizados para as aproximações são descritos em maiores detalhes na seção 5.1.1). No caso da expansão em série de Taylor de 1ª ordem, temos a seguinte relação:

$$\begin{aligned}
E(\mu_t|D_{t-1}, \phi) &= E \left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_{t-1} \right] \approx E \left[\frac{e^{f_t}}{1 + e^{f_t}} + (\lambda_t - f_t) \frac{e^{f_t}}{(1 + e^{f_t})^2} \right] \\
&\approx \frac{e^{f_t}}{1 + e^{f_t}}.
\end{aligned}$$

e

$$\begin{aligned}
V(\mu_t|D_{t-1}, \phi) &= V \left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_{t-1} \right] \approx V \left[\frac{e^{f_t}}{1 + e^{f_t}} + (\lambda_t - f_t) \frac{e^{f_t}}{(1 + e^{f_t})^2} \right] \\
&\approx q_t \frac{(e^{f_t})^2}{(1 + e^{f_t})^4}.
\end{aligned}$$

Problema: De modo a encontrar os valores s_t e r_t que satisfaçam

$$E(g(\mu_t)|D_{t-1}, \phi) = f_t \quad \text{e} \quad V(g(\mu_t)|D_{t-1}, \phi) = q_t,$$

é necessário obter os valores de $E(\mu_t|D_{t-1}, \phi)$ e $V(\mu_t|D_{t-1}, \phi)$. Suponha que tais valores sejam

$$\mu(s_t, r_t) = \hat{E}(\mu_t|D_{t-1}, \phi) \quad \text{e} \quad \sigma^2(s_t, r_t) = \hat{V}(\mu_t|D_{t-1}, \phi),$$

tem-se

$$\mu(s_t, r_t) \cong \frac{e^{f_t}}{1 + e^{f_t}} \quad \text{e} \quad \sigma^2(s_t, r_t) \cong q_t \frac{(e^{f_t})^2}{(1 + e^{f_t})^4}, \quad (4.6)$$

em que s_t e r_t resolvem as equações acima.

Os cálculos dos hiperparâmetros s_t e r_t envolvem contas muito complicadas. Como forma de contornar este problema, no Capítulo 5 o modelo é formulado utilizando outra opção de *priori* para $(\mu_t|D_{t-1}, \phi)$.

Capítulo 5

Modelo Dinâmico Beta com priori Beta

No contexto da abordagem dos *Modelos Lineares Dinâmicos Generalizados* (McCullagh e Nelder, 1994), a escolha da distribuição a priori é feita de modo que seja o mais conveniente possível para operar-se analiticamente as equações envolvidas no processo ite-rativo de estimação dos parâmetros do modelo. Na maioria dos casos, a priori conjugada é a escolha mais adequada, devido a possibilidade da obtenção de soluções em forma exata, o que na maioria dos casos simplifica os procedimentos computacionais. No entanto, como discutiu-se no Capítulo 4, o uso da priori conjugada não possibilitou tal simplificação do problema. Neste capítulo apresenta-se a formulação do Modelo Dinâmico Beta utilizando uma priori não conjugada, justamente com o intuito de simplificar os cálculos.

5.1 Modelo Dinâmico Beta com ϕ conhecido

Seguem abaixo os passos de estimação dos parâmetros do Modelo Dinâmico Beta para o caso de ϕ conhecido.

– **Equação das observações:**

$$P(Y_t|\phi, \mu_t) = \frac{\Gamma(\phi)}{\Gamma(\mu_t\phi)\Gamma((1-\mu_t)\phi)} y^{\mu_t\phi-1} (1-y)^{(1-\mu_t)\phi-1}. \quad (5.1)$$

– **Priori:**

$$(\mu_t | D_{t-1}, \phi) \sim \text{Beta}(s_t, r_t).$$

Considere a função logito de ligação logito:

$$\lambda_t = g(\mu_t) = F_t' \beta_t = \log \left(\frac{\mu_t}{1 - \mu_t} \right) \Rightarrow \mu_t = \frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)}.$$

– **Equação das evoluções:**

$$\beta_t = G_t \beta_{t-1} + w_t; \quad w_t \sim (0, W_t).$$

– **Informação Inicial:**

$$(\beta_0 | D_0, \phi) \sim (m_0, C_0).$$

5.1.1 Inferência para o Modelo Dinâmico Beta

– **Passo 1:** Priori de $\lambda_t = F_t' \beta_t$.

Considerando que

$$(\beta_{t-1} | D_{t-1}, \phi) \sim (m_{t-1}, C_{t-1}),$$

seja

$$f_t = F_t' a_t; \quad q_t = F_t' R_t F_t; \quad a_t = G_t m_{t-1}; \quad R_t = G_t C_{t-1} G_t' + W_t.$$

As prioris de λ_t e β_t são dadas por:

$$(\lambda_t | D_{t-1}, \phi) \sim (f_t, q_t) \quad e \quad (\beta_t | D_{t-1}, \phi) \sim (a_t, R_t).$$

Com isso, toma-se a distribuição conjunta:

$$\left(\begin{array}{c} \lambda_t \\ \beta_t \end{array} \middle| D_{t-1}, \phi \right) \sim \left[\begin{array}{c} \left(\begin{array}{c} f_t \\ a_t \end{array} \right), \left(\begin{array}{cc} q_t & F_t' R_t \\ R_t F_t & R_t \end{array} \right) \end{array} \right]$$

– **Passo 2:** Priori para μ_t .

2a Especificação da priori de μ_t , $P(\mu_t|D_{t-1}, \phi)$:

Como $\mu_t \in (0, 1)$, considere uma priori $Beta(s_t, r_t)$ onde s_t e r_t sejam especificados de modo a satisfazer as esquações

$$E[g(\mu_t)|D_{t-1}, \phi] = f_t \quad e \quad V[g(\mu_t)|D_{t-1}, \phi] = q_t,$$

devidas a $(\lambda_t|D_{t-1}, \phi) \sim (f_t, q_t)$.

2b Relacionando $E(\mu_t|D_{t-1}, \phi)$ e $V(\mu_t|D_{t-1}, \phi)$ com $E(\lambda_t|D_{t-1}, \phi)$ e $V(\lambda_t|D_{t-1}, \phi)$:

$$E(\mu_t|D_{t-1}) = E \left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_{t-1}, \phi \right].$$

No cálculo acima, considere a expansão de Taylor de 1ª ordem ao redor de x_0 para a função:

$$f(x) = \frac{e^x}{1 + e^x}.$$

Logo,

$$\begin{aligned} f(x) &\approx f(x_0) + (x - x_0)f'(x_0) \\ &\approx \frac{e^{x_0}}{1 + e^{x_0}} + (x - x_0) \frac{e^{x_0}}{(1 + e^{x_0})^2}. \end{aligned}$$

Seja $x_0 = f_t$. Então,

$$\frac{e^x}{1 + e^x} \approx \frac{e^{f_t}}{1 + e^{f_t}} + (x - f_t) \frac{e^{f_t}}{(1 + e^{f_t})^2}.$$

Logo,

$$\begin{aligned} E(\mu_t|D_{t-1}, \phi) &= E \left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_{t-1}, \phi \right] \\ &\approx E \left[\frac{e^{f_t}}{1 + e^{f_t}} + (\lambda - f_t) \frac{e^{f_t}}{(1 + e^{f_t})^2} \right] \\ &\approx \frac{e^{f_t}}{1 + e^{f_t}} + \frac{e^{f_t}}{(1 + e^{f_t})^2} (E[\lambda_t|D_{t-1}, \phi] - f_t) \\ &\approx \frac{e^{f_t}}{1 + e^{f_t}}. \end{aligned}$$

Similarmente,

$$\begin{aligned} V(\mu_t|D_{t-1}, \phi) &\approx V \left[\frac{e^{f_t}}{1 + e^{f_t}} + (\lambda - f_t) \frac{e^{f_t}}{(1 + e^{f_t})^2} \middle| D_{t-1}, \phi \right] \\ &\approx \left(\frac{e^{f_t}}{(1 + e^{f_t})^2} \right)^2 q_t. \end{aligned}$$

Buscando melhorar a aproximação dos momentos $E(\mu_t|D_{t-1})$ e $V(\mu_t|D_{t-1})$, considere, também, a expansão de Taylor de 2ª ordem ao redor de x_0 para a função $f(x) = \frac{e^x}{1+e^x}$.

$$\begin{aligned} f(x) &\approx f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!} f''(x_0) \\ &\approx \frac{e^{x_0}}{1 + e^{x_0}} + (x - x_0) \frac{e^{x_0}}{(1 + e^{x_0})^2} + \frac{(x - x_0)^2}{2!} \frac{e^{x_0}(1 - e^{x_0})}{(1 + e^{x_0})^3}. \end{aligned}$$

Com $x_0 = f_t$,

$$\frac{e^x}{1 + e^x} \approx \frac{e^{f_t}}{1 + e^{f_t}} + (x - f_t) \frac{e^{f_t}}{(1 + e^{f_t})^2} + \frac{(x - f_t)^2}{2!} \frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3}.$$

Logo,

$$\begin{aligned} E(\mu_t|D_{t-1}, \phi) &= E \left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_{t-1}, \phi \right] \\ &\approx E \left[\frac{e^{f_t}}{1 + e^{f_t}} + (\lambda - f_t) \frac{e^{f_t}}{(1 + e^{f_t})^2} + \frac{(\lambda - f_t)^2}{2!} \frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3} \right] \\ &\approx \frac{e^{f_t}}{1 + e^{f_t}} + \frac{e^{f_t}}{(1 + e^{f_t})^2} (E[\lambda_t|D_{t-1}, \phi] - f_t) \\ &\quad + \frac{V[\lambda_t|D_{t-1}, \phi]}{2!} \frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3} \\ &\approx \frac{e^{f_t}}{1 + e^{f_t}} + \frac{q_t}{2} \frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3}. \end{aligned}$$

$$\begin{aligned} V(\mu_t|D_{t-1}, \phi) &\approx V \left[\frac{e^{f_t}}{1 + e^{f_t}} + (\lambda - f_t) \frac{e^{f_t}}{(1 + e^{f_t})^2} + \frac{(\lambda - f_t)^2}{2!} \frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3} \middle| D_{t-1}, \phi \right] \\ &\approx \left(\frac{e^{f_t}}{(1 + e^{f_t})^2} \right)^2 q_t + \frac{(8q_t f_t^2 - q_t^2)}{4} \left(\frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3} \right)^2. \end{aligned}$$

c) Obtenção dos hiperparâmetros de $(\mu_t|D_{t-1}, \phi) \sim Beta(s_t, r_t)$:

c.1 - Obtendo s_t e r_t via aproximação de segunda ordem de $E(\mu_t|D_{t-1}, \phi)$:

Como $(\mu_t|D_{t-1}, \phi) \sim Beta(s_t, r_t)$,

$$E(\mu_t|D_{t-1}, \phi) = \frac{s_t}{s_t + r_t} \quad e \quad V(\mu_t|D_{t-1}, \phi) = \frac{s_t r_t}{(s_t + r_t)^2 (s_t + r_t + 1)}.$$

No caso do uso da expansão de Taylor de 2ª ordem o sistema é definido por:

$$\begin{cases} \frac{s_t}{s_t + r_t} = \frac{e^{f_t}}{1 + e^{f_t}} + \frac{q_t}{2} \frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3} \\ \frac{s_t r_t}{(s_t + r_t)^2 (s_t + r_t + 1)} = \left(\frac{e^{f_t}}{(1 + e^{f_t})^2} \right)^2 q_t + \frac{(8q_t f_t^2 - q_t^2)}{4} \left(\frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3} \right)^2 \end{cases} \quad (5.2)$$

Para facilitar a solução do sistema, considere os valores da expansão em termos de h_t e v_t :

$$\begin{aligned} h_t &= \frac{e^{f_t}}{1 + e^{f_t}} + \frac{q_t}{2} \frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3} \quad e \\ v_t &= \left(\frac{e^{f_t}}{(1 + e^{f_t})^2} \right)^2 q_t + \frac{(8q_t f_t^2 - q_t^2)}{4} \left(\frac{e^{f_t}(1 - e^{f_t})}{(1 + e^{f_t})^3} \right)^2. \end{aligned}$$

Definindo a expansão em termos de h_t e v_t , independentemente da ordem utilizada para aproximar os momentos $E(\mu_t|D_{t-1}, \phi)$ e $V(\mu_t|D_{t-1}, \phi)$, a solução do sistema para definir os parâmetros da priori $P(\mu_t|D_{t-1}, \phi)$ é dada por:

$$s_t = \frac{(1 - h_t)h_t^2}{v_t} - h_t; \quad r_t = \frac{(1 - h_t)^2 h_t}{v_t} - (1 - h_t).$$

c.2 - Obtendo s_t e r_t via momentos *log-beta*:

Uma metodologia alternativa é descrita (vide West e Harrison (1997) pg. 529 e 530) para encontrar estimativas aproximadas de s_t e r_t quando $(\mu_t|D_{t-1}, \phi)$ segue uma distribuição $Beta(s_t, r_t)$ e quando a função de ligação, λ_t , é a transformação logística de μ_t . Neste caso, $E(\lambda_t|D_{t-1}, \phi)$ e $V(\lambda_t|D_{t-1}, \phi)$ são determinados da forma:

$$\begin{aligned} E(\lambda_t|D_{t-1}, \phi) &= E \left[\log \left(\frac{\mu_t}{1 - \mu_t} \right) \middle| D_{t-1}, \phi \right] = \psi(s_t) - \psi(r_t) = f_t, \\ V(\lambda_t|D_{t-1}, \phi) &= V \left[\log \left(\frac{\mu_t}{1 - \mu_t} \right) \middle| D_{t-1}, \phi \right] = \psi'(s_t) + \psi'(r_t) = q_t, \end{aligned}$$

onde $\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ e $\psi'(z)$ são, respectivamente, as funções *digama* e *trigama* (mais informações sobre aproximações desta função podem ser encontradas em Muqattash e Yahdi, 2006).

– **Passo 3:** Previsão de 1 passo a frente, $P(Y_t|D_{t-1}, \phi)$:

$$P(Y_t|D_{t-1}, \phi) = \int P(Y_t|\mu_t, D_{t-1}, \phi)P(\mu_t|D_{t-1})d\mu_t.$$

Esta integral não possui forma fechada e, portanto, faz-se necessário o uso de aproximações numéricas para resolvê-la. Uma possibilidade é utilizar os métodos de Newton-Cotes (aproximações por quadratura) (vide Abramovitz e Stegun, 1972, Corbit, 1996 e Migon e Gamerman, 1999, pg. 144). Intervalos de credibilidade podem ser aproximados por *intervalos de máxima densidade a posteriori* (HDP) para $(Y_t|D_{t-1}, \phi)$ (vide Migon e Gamerman, 1999, pg 99). Entretanto, como na maioria dos casos, o interesse está unicamente nos primeiros momentos de $(Y_t|D_{t-1})$. Estes podem ser encontrados diretamente utilizando as propriedades condicionais de média e variância juntamente com o fato de $\mu_t \sim \text{Beta}(s_t, r_t)$. Sendo assim, a previsão de um passo a frente de Y_t é estimada da forma:

$$\begin{aligned} \tilde{Y}_t &= E(Y_t|D_{t-1}, \phi) = E[E(Y_t|\mu_t, D_{t-1}, \phi)] \\ &= \frac{s_t}{s_t + r_t}, \end{aligned} \quad (5.3)$$

e variância:

$$\begin{aligned} V(Y_t|D_{t-1}, \phi) &= E[V(Y_t|\mu_t, D_{t-1}, \phi)] + V[E(Y_t|\mu_t, D_{t-1}, \phi)] \\ &= E \left[\frac{\mu_t - \mu_t^2}{1 + \phi} \middle| D_{t-1}, \phi \right] + V[\mu_t|D_{t-1}, \phi] = \\ &= \frac{1}{1 + \phi} \left[\frac{s_t}{s_t + r_t} \left(1 - \frac{s_t}{s_t + r_t} \right) \right] + \frac{\phi}{1 + \phi} \frac{s_t r_t}{(s_t + r_t)^2 (s_t + r_t + 1)} \end{aligned} \quad (5.4)$$

Estes momentos são de grande utilidade para medidas comparativas. Por exemplo, a partir de $E(Y_t|D_{t-1}, \phi)$, podemos obter medidas de precisão para a previsão, como o desvio médio absoluto, DMA e o erro quadrático médio, EQM , onde: $DMA = \frac{1}{T} \sum_{t=1}^T |e_t|$ e $EQM = \frac{1}{T} \sum_{t=1}^T e_t^2$, com $e_t = Y_t - E(Y_t|D_{t-1}, \phi)$.

A variância, $V(Y_t|D_{t-1}, \phi)$ pode ser utilizada para obter intervalos de confiança de previsão de um passo a frente. Outra medida utilizada para comparação de modelos é a *densidade preditiva observada* ou *verossimilhança observada*, descrita por:

$$p(Y_1, \dots, Y_T|D_0, \phi) = \prod_{t=1}^T p(Y_t|D_{t-1}, \phi). \quad (5.5)$$

– **Passo 4** - Posteriori de μ_t - $P(\mu_t|D_t)$:

Pelo Teorema de Bayes, a posteriori de μ_t é:

$$\begin{aligned} P(\mu_t|D_t, \phi) &\propto P(Y_t|\mu_t, D_{t-1}, \phi)P(\mu_t|D_{t-1}, \phi) \\ &\propto \frac{\Gamma(\phi)}{\Gamma(\mu_t\phi)\Gamma((1-\mu_t)\phi)} y_t^{\mu_t\phi-1} (1-y_t)^{(1-\mu_t)\phi-1} \\ &\quad \times \frac{\Gamma(s_t+r_t)}{\Gamma(s_t)\Gamma(r_t)} \mu_t^{s_t-1} (1-\mu_t)^{r_t-1}. \end{aligned} \quad (5.6)$$

Observe que $P(\mu_t|D_t, \phi)$ não tem forma conhecida. Portanto, faz-se necessário utilizar um método aproximado para estimar os momentos de 1ª e 2ª ordem de $P(\mu_t|D_t, \phi)$. Uma possibilidade é utilizar a *moda a posteriori*, que é o ponto que maximiza $P(\mu_t|D_t, \phi)$, para estimar $E(\mu_t|D_t, \phi)$, e menos a inversa da matriz Hessiana para estimar $V(\mu_t|D_t, \phi)$.

– **Aproximações para $E(\mu_t|D_t, \phi)$ e $V(\mu_t|D_t, \phi)$ via moda a posteriori:**

Seja:

$$\begin{aligned} L = P(\mu_t|D_t, \phi) &\propto \frac{\Gamma(\phi)}{\Gamma(\mu_t\phi)\Gamma((1-\mu_t)\phi)} y_t^{\mu_t\phi-1} (1-y_t)^{(1-\mu_t)\phi-1} \\ &\quad \times \frac{\Gamma(s_t+r_t)}{\Gamma(s_t)\Gamma(r_t)} \mu_t^{s_t-1} (1-\mu_t)^{r_t-1}, \end{aligned}$$

$$\begin{aligned} l = \log(L) &\propto -\log(\Gamma(\phi\mu_t)) - \log(\Gamma(\phi(1-\mu_t))) + (\mu_t\phi - 1) \log(y_t) \\ &\quad + [(1-\mu_t)\phi - 1] \log(1-y_t) + (s_t - 1) \log \mu_t + (r_t - 1) \log(1-\mu_t). \end{aligned}$$

Seja m o ponto em $\mu_t \in (0, 1)$ que maximiza l , este é denominado a *moda a posteriori* de μ_t . Seguem abaixo as derivadas de 1ª e 2ª ordem e estimativa

da variância V :

$$\frac{\partial l}{\partial \mu_t} \propto -\phi\psi(\phi\mu_t) + \phi\psi[\phi(1 - \mu_t)] + \phi \log y_t - \phi \log(1 - y_t) + \frac{s_t - 1}{\mu_t} - \frac{r_t - 1}{1 - \mu_t},$$

$$\frac{\partial^2 l}{\partial \mu_t^2} \propto -\phi^2\psi'(\phi\mu_t) - \phi^2\psi'[\phi(1 - \mu_t)] - \frac{s_t - 1}{\mu_t^2} - \frac{r_t - 1}{(1 - \mu_t)^2},$$

$$V \approx - \left[\frac{\partial^2 l}{\partial \mu_t^2} \Big|_{\mu_t=m} \right]^{-1},$$

Desta forma, temos os estimadores dos momentos:

$$m = \hat{E}(\mu_t | D_t)$$

$$V = \hat{V}(\mu_t | D_t).$$

– **Aproximações para $E(\mu_t | D_t, \phi)$ e $V(\mu_t | D_t, \phi)$ via método de Laplace:**

A precisão na estimação de $E(\mu_t | D_t, \phi)$ e $V(\mu_t | D_t, \phi)$ pode ser melhorada com o uso de aproximações de Laplace, introduzidas por Tierney e Kadane (1986). A seguir descrevemos o método e o aplicamos ao caso do modelo beta dinâmico.

Seja a distribuição a posteriori $\pi(\theta)$, a priori $p(\theta)$ associada a θ e $L(x | \theta)$ a verossimilhança. Utilizando o teorema de Bayes, os momentos a posteriori podem ser obtidos através de

$$E(h(\theta)) = \int h(\theta)\pi(\theta)d\theta = \frac{\int h(\theta)L(x | \theta)p(\theta)}{\int L(x | \theta)p(\theta)d\theta} = \frac{\int \exp(K^*(\theta))d\theta}{\int \exp(K(\theta))d\theta},$$

onde $h(\theta)$ é uma função estritamente positiva,

$$K^*(\theta) = \log(h(\theta)) + \log(L(x | \theta)) + \log(p(\theta))$$

e

$$K(\theta) = \log(L(x | \theta)) + \log(p(\theta)).$$

Aplicando a expansão de Taylor de segunda ordem em K^* , tem-se

$$K^*(\theta) \approx K^*(m^*) - \frac{1}{2}(\theta - m^*)'(V^*)^{-1}(\theta - m^*),$$

de modo que, considerando a aproximação de Laplace proposta por Tierney e Kadane (1986),

$$E(h(\theta)) \approx \left(\frac{|V^*|}{|V|} \right)^{1/2} \exp(K^*(m^*) - K(m)),$$

onde m^* é o ponto que maximiza $K^*(\theta)$, V^* é a inversa da matriz hessiana associada a $K^*(\theta)$, avaliada no ponto m^* . Definições similares se aplicam a m e V .

Considerando-se $\theta = \mu_t$, $h(\mu_t) = \mu_t$ e, posteriormente, $h(\mu_t) = \mu_t^2$, $\pi(\mu_t) = P(\mu_t | \phi, D_t)$, e

$$L(x | \mu_t, \phi) = \frac{\Gamma(\phi)}{\Gamma(\phi\mu_t)\Gamma(\phi(1-\mu_t))} x^{\phi\mu_t-1} (1-x)^{\phi(1-\mu_t)-1},$$

tem-se

$$\begin{aligned} V &= - \left[\frac{\partial^2 K(\mu_t)}{\partial \mu_t^2} \Big|_{\mu_t=m} \right]^{-1} \quad \text{and} \quad V^* = - \left[\frac{\partial^2 K^*(\mu_t)}{\partial \mu_t^2} \Big|_{\mu_t=m^*} \right]^{-1}, \\ d_t &= \hat{E}[\mu_t | D_t, \phi] \approx \left(\frac{|V^*|}{|V|} \right)^{1/2} \exp[K^*(m^*) - K(m)], \\ l_t &= \hat{V}[\mu_t | D_t, \phi] \approx \hat{E}[\mu_t^2 | D_t, \phi] - (\hat{E}[\mu_t | D_t, \phi])^2. \end{aligned}$$

Como não é possível encontrar m e m^* analiticamente, utilizamos o método de Newton-Raphson.

Utilizando os valores l_t e d_t de maneira análoga ao segundo item do Passo

2, com a **expansão de primeira ordem**, tem-se:

$$\begin{aligned} d_t &\approx E[\mu_t|D_t, \phi] = E \left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_t, \phi \right] \\ &\approx \frac{e^{f_t^*}}{1 + e^{f_t^*}}, \end{aligned}$$

$$\begin{aligned} l_t &\approx V[\mu_t|D_t, \phi] = V \left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_t, \phi \right] \\ &\approx \left(\frac{e^{f_t^*}}{(1 + e^{f_t^*})^2} \right)^2 q_t^*. \end{aligned}$$

Neste caso, os momentos de $(\lambda_t|D_t, \phi)$, f_t^* e q_t^* são definidos, em termos de d_t e l_t , por:

$$f_t^* = \log \left(\frac{d_t}{1 - d_t} \right) \quad \text{e} \quad q_t^* = l_t \left(\frac{e^{f_t^*}}{(1 + e^{f_t^*})^2} \right)^{-2} \quad (5.7)$$

Com a **expansão de 2ª ordem**, os momentos d_t e l_t são expressos por:

$$\begin{aligned} d_t &\approx E[\mu_t|D_t, \phi] = E\left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_t, \phi\right] \\ &\approx \frac{e^{f_t^*}}{1 + e^{f_t^*}} + \frac{q_t^* e^{f_t} (1 - e^{f_t})}{2 (1 + e^{f_t})^3}, \\ l_t &\approx V[\mu_t|D_t, \phi] = V\left[\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} \middle| D_t, \phi\right] \\ &\approx \left(\frac{e^{f_t^*}}{(1 + e^{f_t^*})^2}\right)^2 q_t^* + \frac{8q_t^* (f_t^*)^2 - (q_t^*)^2}{4} \left(\frac{e^{f_t} (1 - e^{f_t})}{(1 + e^{f_t})^3}\right)^2. \end{aligned}$$

Então, f_t^* e q_t^* são obtidos a partir da solução deste sistema de equações não lineares. A solução deste sistema não pode ser expressa de forma analítica, portanto foi utilizado um procedimento iterativo por meio da função “*dfsane*”, disponível na biblioteca *BB* do software *R*, para obter uma solução aproximada. A unicidade da solução pode ser garantida, uma vez que q_t^* está definido entre os reais positivos.

– **Passo 5** - Estrutura Condicional para $(\beta_t|\lambda_t, D_{t-1}, \phi)$:

$$P(\beta_t|D_t, \phi) = \int P(\beta_t|\lambda_t, D_{t-1}, \phi)P(\lambda_t|D_t, \phi)d\lambda_t$$

Utilizando o método de estimação Linear Bayes, encontra-se:

$$\begin{cases} \hat{E}[\beta_t|\lambda_t, D_{t-1}, \phi] = a_t + R_t F_t (\lambda_t - f_t) / q_t, \\ \hat{V}[\beta_t|\lambda_t, D_{t-1}, \phi] = R_t - R_t F_t F_t' R_t / q_t. \end{cases}$$

– **Passo 6**: Atualização de β_t :

$$P(\beta_t|D_t, \phi) = \int P(\beta_t|\lambda_t, D_{t-1}, \phi)P(\lambda_t|D_t, \phi)d\lambda_t.$$

Segue que

$$E(\beta_t|D_t, \phi) = E[E\{\beta_t|\lambda_t, D_{t-1}, \phi\}|D_t, \phi] \quad \text{e}$$

$$V(\beta_t|D_t, \phi) = V[E\{\beta_t|\lambda_t, D_{t-1}, \phi\}|D_t, \phi] + E[V\{\beta_t|\lambda_t, D_{t-1}, \phi\}|D_t, \phi].$$

Desta forma, os momentos a posteriori de $(\beta_t|D_t)$ podem ser estimados por

$$(\beta_t|D_t, \phi) \sim (m_t, C_t),$$

onde,

$$m_t = a_t + R_t F_t (f_t^* - f_t) / q_t$$

e

$$C_t = R_t - \frac{1}{q_t} [R_t F_t F_t' R_t (1 - q_t^* / q_t)].$$

Desta forma, temos o valor esperado de λ_t como $F_t' E(\beta_t | D_t) = F_t' m_t$.

Como

$$\lambda_t = g(\mu_t) = \log \left(\frac{\mu_t}{1 - \mu_t} \right).$$

O nível μ_t é estimado a partir da função inversa da função de ligação $g(\cdot)$ sobre o valor esperado de λ_t :

$$\tilde{\mu}_t = \frac{\exp(F_t' m_t)}{1 + \exp(F_t' m_t)}.$$

5.2 Modelo Dinâmico Beta com ϕ desconhecido

Na seção anterior o Modelo Dinâmico Beta foi descrito considerando-se conhecido o valor do parâmetro escala ϕ . A estimação de ϕ , a princípio, não deveria ser um problema muito complicado e, para esse propósito, alternativas razoáveis seriam utilizar ou a expressão (5.1), relacionada com a equação de observação, ou a expressão (5.6), que descreve a distribuição a posteriori de $(\mu_t | D_t, \phi)$, tendo em vista que ambas envolvem ϕ . No entanto, nos deparamos com algumas complicações.

Uma delas é a dificuldade em se obter uma função de verossimilhança que expresse a relação de dependência entre todos os μ_t 's. Uma forma de contornar esta dificuldade é desconsiderando-se a relação de dependência com o uso de uma função escrita como o produto das distribuições individuais de μ_t , como no caso de dados i.i.d.

Outra dificuldade é devido ao fato dos valores da função descrita tanto pela expressão (5.1) quanto pela expressão (5.6) tenderem a crescer monotonicamente com ϕ . Na parte superior da Figura (5.1) é ilustrada a superfície de uma função de verossimilhança descrita para dados i.i.d. gerados a partir de

$y \sim \text{Beta}(\phi\mu, \phi(1 - \mu))$ para $\mu = 0,4$ e $\phi = 50$. Percebe-se que a superfície é bem achatada no eixo em que os valores de ϕ variam, dificultando no momento de maximizar a função com respeito a ϕ correspondente ao problema. Na parte inferior da Figura (5.1) observa-se um corte da superfície para $\mu = 0,4$, evidenciando, mais uma vez, que a maximização da verossimilhança com respeito a ϕ pode ser problemática. A literatura ainda carece de maiores informações sobre o problema e como tratá-lo.

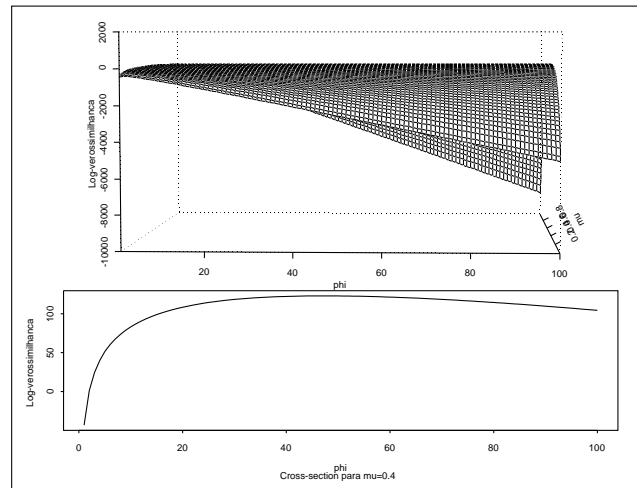


Figura 5.1: Função de log-verossimilhança gerada de dados com $\phi = 50$ e $\mu = 0,4$

5.2.1 Procedimento Ad-Hoc para estimar ϕ em dois estágios

A maior dificuldade em estimar ϕ por meio das expressões (5.1) ou (5.6) ocorre devido a presença da função gama envolvendo o parâmetro ϕ . Como uma alternativa, um procedimento *ad-hoc* é utilizado para estimar tal parâmetro.

Um procedimento completamente bayesiano para estimar ϕ é bastante mais complexo e estamos trabalhando para desenvolvê-lo (vide Seção 7.2). No entanto, com o intuito de evitar o uso de procedimentos computacionalmente intensos, como o MCMC, por exemplo, foi proposta esta abordagem *ad-hoc*.

A idéia é baseada em utilizar uma função que, mesmo sem estar diretamente parametrizada em termos de ϕ , captura a sua influência indiretamente. Nota-se que, como explicitado anteriormente, μ_t e ϕ estão relacionados na forma $\mu_t = p_t/\phi$, onde p_t é o parâmetro da distribuição Beta com a parametrização

convencional. Desta forma, considerando a priori $(\mu_t|D_{t-1}, \phi)$, que depende de ϕ indiretamente, mas possui a vantagem de não incluir a função gama envolvendo ϕ , é apresentado, a seguir, um procedimento em dois estágios para estimar tal parâmetro.

- **Grid:** Considera-se um conjunto de pares de valores para (W_i, ϕ_j) , com W representando a variância dos erros de evolução, $i = 1, \dots, a$ e $j = 1, \dots, b$.
- **(a) $(\phi|W)$:** Nesta etapa, deseja-se estimar ϕ para um valor fixo de W . Ou seja, para um ponto fixo da coordenada W_t é obtido o valor da função L_P para todos os pontos ϕ , onde

$$\begin{aligned} L_P(\phi) &= \prod_{t=1}^n P(\mu_t|D_{t-1}) \\ &= \prod_{t=1}^n \frac{\Gamma(s_t + r_t)}{\Gamma(s_t)\Gamma(r_t)} \mu_t^{s_t-1} (1 - \mu_t)^{r_t-1} \\ &\approx \prod_{t=1}^n \phi^{s_t+r_t-2} (\phi - p_t)^{r_t-1}, \quad \text{onde } \mu_t = p_t/\phi. \end{aligned}$$

São armazenados os valores $\hat{\phi}_{(i)}$ que maximizam L_P no ponto fixo $W = W_i$, e, por fim, o valor estimado de ϕ é dado por:

$$\hat{\phi} = \text{mediana}\{\hat{\phi}_{(1)}, \dots, \hat{\phi}_{(a)}\}.$$

- **(b) $(W|\phi)$:** Para estimar W condicionalmente com respeito a ϕ , o procedimento consiste em minimizar o quadrado médio dos erros descrito pela expressão (5.8) sobre os pontos do grid (W_i, ϕ_j) :

$$EQM_Y(\phi) = \frac{1}{T} \sum_{t=1}^T [Y_t - \tilde{Y}_t]^2, \quad (5.8)$$

onde \tilde{Y}_t corresponde ao valor esperado da previsão um passo a frente de $(Y_t|D_{t-1}, \phi)$.

No Capítulo 7 são apresentadas algumas formulações para estimar o parâmetro ϕ no processo inferencial do ajuste do modelo. No entanto, ainda faz-se necessário alguns ajustes para verificar o funcionamento destes métodos.

Capítulo 6

Simulação e Aplicação

Neste capítulo discutiremos, inicialmente, o relacionamento entre a variância do erro observacional, V_t , e a variância do erro de evolução, W_t , para o Modelo Dinâmico Beta. Em seguida, são apresentadas as análises com dados simulados *Beta* para o caso de ϕ conhecido. São também comparadas diferentes aproximações para estimar os momentos a posteriori que são utilizados no processo de inferência do Modelo Dinâmico Beta. Ao final, analisamos alguns conjuntos de dados reais.

6.1 A razão sinal-ruído no Modelo Dinâmico Beta

Nesta seção analisamos a relação entre a variância do erro observacional, V_t , e a variância do erro de evolução, W_t .

Os erros, w_t , do sistema de equações, controlam a evolução sob a influência de uma variância W_t que, quanto maior (menor) são os seus valores mais brusca (suave) é a evolução ao longo do tempo. O fato de $E(w_t) = 0$ garante um certo nível fixo em torno do qual os valores de β_t irão variar.

O comportamento das trajetórias de Y_t e β_t estão relacionadas com a magnitude da razão entre as variâncias, $r_t = W_t/V_t$ (razão sinal-ruído). Quando o valor de r_t é baixo, a maioria dos movimentos da série se dá devido às observações Y_t e

no caso de valores elevados de r_t , a maior parte dos movimentos se dá devido a ambas as variações nos Y_t 's e nos β_t 's.

Para os Modelos Dinâmicos Gaussianos, com variância observacional e de evolução constantes, um valor de $r_t = 0.05$, que corresponde a $W = V/20$, indica séries de tempo tipicamente suaves em torno de um nível localmente constante. Por outro lado, $r_t = 0.5$ corresponde a séries de tempo com movimentos mais bruscos ao longo do tempo.

No caso dos Modelos Dinâmicos Beta, esta relação não acontece desta mesma forma, tendo em vista que a natureza dos dados analisados no caso gaussiano e no beta são muito diferentes. Além disso, cabe ressaltar o fato de que a variância observacional, V_t depende do nível μ_t (que não é o caso no modelo gaussiano) o que torna esta relação ainda mais complexa. Com isso, os valores correspondentes a r_t no Modelo Dinâmico Beta podem não ter a mesma interpretação de sinal-ruído do Modelo Dinâmico Gaussiano.

Nos Modelos Dinâmicos Beta, a variância, V_t , apresenta um limite superior dado por $V_t \leq \frac{1}{4(1+\phi)}$. O limite superior é alcançado quando $\mu_t = 1/2$. Na medida em que μ_t se aproxima dos limites do espaço paramétrico (0 ou 1), V_t se aproxima de 0. A magnitude de V_t é pequena e decresce com o aumento de ϕ . O caso de $V_t = 0$ corresponde a observações de séries estáticas ao longo do tempo. Para as simulações, foram utilizados níveis de valores tais que $0.01 \leq \mu_t \leq 0.99$, portanto, $\frac{0.0099}{\phi+1} \leq V_t \leq \frac{1}{4(1+\phi)}$.

Com o intuito de estabelecer uma relação coerente entre V_t , que está definida no intervalo (0, 1), e W_t , que está definida no intervalo (0, ∞), foi feito um reescalonamento de W_t com base em uma transformação sobre o erro de evolução w_t , tal que $w_t^* = \frac{e^{w_t}}{1+e^{w_t}}$. Obtém-se que $W_t^* = W_t/16$ consiste em uma aproximação da variância de w_t^* . Portanto, $\frac{0.0099r_t}{\phi+1} \leq W_t \leq \frac{r_t}{4(1+\phi)}$, onde $r_t = W_t^*/V_t$, implicando, por exemplo, em $W_t^* \leq 0.0125/(1 + \phi)$ para um sinal-ruído de $r_t = 0.05$.

6.2 Lidando com observações $Y_t = 0$ ou $Y_t = 1$

Um problema a ser tratado nas séries de dados composicionais é o cuidado com os pontos muito próximos às extremidades do espaço paramétrico (0 ou 1). Sabe-se que a função de ligação *logito* não está definida em 0 e 1 e, portanto, os pontos próximos a esta região levam a pontos indefinidos na função de verossimilhança da série, impossibilitando sua maximização.

Uma forma intuitiva de contornar este problema é utilizando-se uma transformação para “encolher” o intervalo $[0, 1]$ de modo que os pontos não cheguem tão próximos de suas extremidades. Um método frequentemente utilizado é adicionar $1/(2N)$ em pontos próximos de *zero* e subtrair $1/(2N)$ dos pontos próximos de 1, onde N é o número de observações (MacMillan e Creelman, 2005). Esta transformação não altera os pontos internos do intervalo, mas pode trazer algum viés aos vizinhos dos valores extremos. Uma sugestão mais adequada, apresentada por Smithson e Verkuilen (2006), para uma série contínua em $[0, 1]$ é transformar a série numa variável definida no intervalo aberto $(0, 1)$ ponderada da forma:

$$y_t^* = \frac{y_t(N - 1) + s}{N},$$

onde N é o número de observações e s uma constante de valor entre 0 e 1. Do ponto de vista *bayesiano*, s funciona como no caso em que levamos em conta uma priori. Sendo assim, 0.5 é um valor razoável para s .

Utilizando-se desta transformação em y_t , o problema de modelar as séries com valores próximos das extremidades do espaço paramétrico é solucionado, permitindo que sejam estimadas séries com grandes oscilações, ou seja, com valores mais elevados de W_t , assim como possibilita estimar séries com maior dispersão, ou seja, valores mais elevados de V_t .

6.3 Analisando dados simulados - ϕ conhecido

Para avaliar o comportamento do modelo dinâmico Beta que foi desenvolvido no Capítulo 5, considerou-se dados gerados segundo dez casos distintos,

que são sumarizados na Tabela 6.1.

W_t	ϕ			
	15	25	50	100
0.01	Caso 1	Caso 2	Caso 3	Caso 4
0.09	Caso 5	Caso 6	Caso 7	-
0.15	Caso 8	Caso 9	-	-
0.20	Caso 10	-	-	-

Tabela 6.1: Combinações de séries avalidas

Destacam-se os casos 4, 9 e 10. No Caso 4 as séries são *bem comportadas*, pois tanto W_t (a variância dos erros de evolução) quanto V_t (a variância dos erros de observações) são pequenos. No Caso 9 as séries geradas são *razoavelmente estáveis*, com valores intermediários de V_t e W_t . No Caso 10 as séries geradas apresentam *oscilações mais bruscas*, uma vez que os valores de ambos, V_t e W_t são elevados.

Em geral, o valor de W_t é desconhecido. Como alternativa, utiliza-se uma estimativa por meio do *fator de desconto* δ (descrito no final do Capítulo 1).

Como um critério para avaliar a bondade do ajuste, utilizou-se o erro quadrático médio entre o valor real, μ_t , e o valor estimado do nível, $\tilde{\mu}_t$, isto é:

$$EQM = \frac{1}{T} \sum_{t=1}^T [\mu_t - \tilde{\mu}_t]^2 \quad (6.1)$$

Observando o erro quadrático médio (EQM) para os dez casos listados na Tabela 6.2, nota-se que nos ajustes com W_t conhecido são obtidos os menores valores de EQM . Em seguida destacam-se os ajustes feitos com fator de desconto $\delta = 0.8$. Estes foram os que apresentaram EQM mais próximas de quando o ajuste foi feito considerando-se os valores reais de W_t , principalmente nos quatro primeiros casos apresentados, em que as séries são simuladas com $W_t = 0.01$ (séries bem comportadas). Nota-se que quanto menor a dispersão dos dados (maior o tamanho de ϕ), maior é a diferença entre os ajustes com W_t conhecido. Na medida em que são utilizados valores mais elevados para o fator de desconto, como $\delta = 0.9$, o EQM tende a ser maior. No entanto,

para séries bem comportadas (como $W_t = 0.01$ e $W_t = 0.09$) os valores ainda apresentam EQM próximos ao obtido com W_t conhecido. Quando utiliza-se $\delta = 0.95$, o EQM aumenta muito, principalmente para séries com valores mais elevados de W_t . De acordo com a Tabela 6.2, $\delta = 0.95$ só é viável para séries com comportamento bastante estável, como em $W_t = 0.01$. No caso de δ 's com valor inferior a 0.8, a magnitude de W_t deixa de ter muita relevância de modo que a previsão é passa a ser quase que totalmente baseada nas observações Y_t 's, desconsiderando grande parte do histórico passado da série.

Casos	$EQM \times 10^2$			
	W_t conhecido	$\delta = 0.8$	$\delta = 0.9$	$\delta = 0.95$
Caso 1: $\phi = 15; W_t = 0.01$	0.3087	0.3108	0.3450	0.4309
Caso 2: $\phi = 25; W_t = 0.01$	0.2368	0.2408	0.3086	0.4167
Caso 3: $\phi = 50; W_t = 0.01$	0.1370	0.1580	0.2518	0.3879
Caso 4: $\phi = 100; W_t = 0.01$	0.0901	0.1319	0.2362	0.3786
Caso 5: $\phi = 15; W_t = 0.09$	0.5117	0.7617	1.3850	2.2440
Caso 6: $\phi = 25; W_t = 0.09$	0.3310	0.5740	1.1215	1.8872
Caso 7: $\phi = 50; W_t = 0.09$	0.2354	0.6545	1.2238	3.1989
Caso 8: $\phi = 15; W_t = 0.15$	0.5892	1.0145	1.7429	3.0818
Caso 9: $\phi = 25; W_t = 0.15$	0.4439	0.9310	1.7255	2.4439
Caso 10: $\phi = 15; W_t = 0.20$	0.5284	1.2071	2.2525	3.2485

Tabela 6.2: Erro quadrático médio das séries simuladas.

Para séries mais estáveis, como o Caso 4, onde $\phi = 100$ e $W_t = 0.01$, valores próximos de 0.9 e 0.95 já apresentam resultados satisfatórios para estimar a série (vide Figuras 6.1 e 6.2). No caso de séries com oscilações mais elevadas, como o Caso 9, onde $\phi = 25$ e $W_t = 0.15$, o fator de desconto $\delta = 0.8$ apresentou estimativas mais próximas das evoluções reais da série, já com $\delta = 0.95$ os níveis obtidos após a suavização (vide Apêndice G) ficaram bem distantes dos níveis reais (vide Figuras 6.3 e 6.4). No Caso 10, onde $\phi = 15$ e $W = 0.2$, foi necessário o uso da transformação de variável mencionado na seção 6.2. A partir da Figura 6.5 percebe-se que a principal diferença ao utilizar os diferentes valores de δ se dá nas regiões próximas das extremidades, no período entre os tempos $t = 50$

e $t \approx 150$, onde ocorrem as oscilações mais bruscas. Pela Figura 6.6 nota-se que $\delta = 0.8$ foi o fator de desconto que proporcionou um comportamento mais próximo em relação à estimativa com W_t conhecido. Com $\delta = 0.95$ a série estimada não é capaz de ajustar as fortes oscilações presentes na série real (vide Figura 6.6 d)).

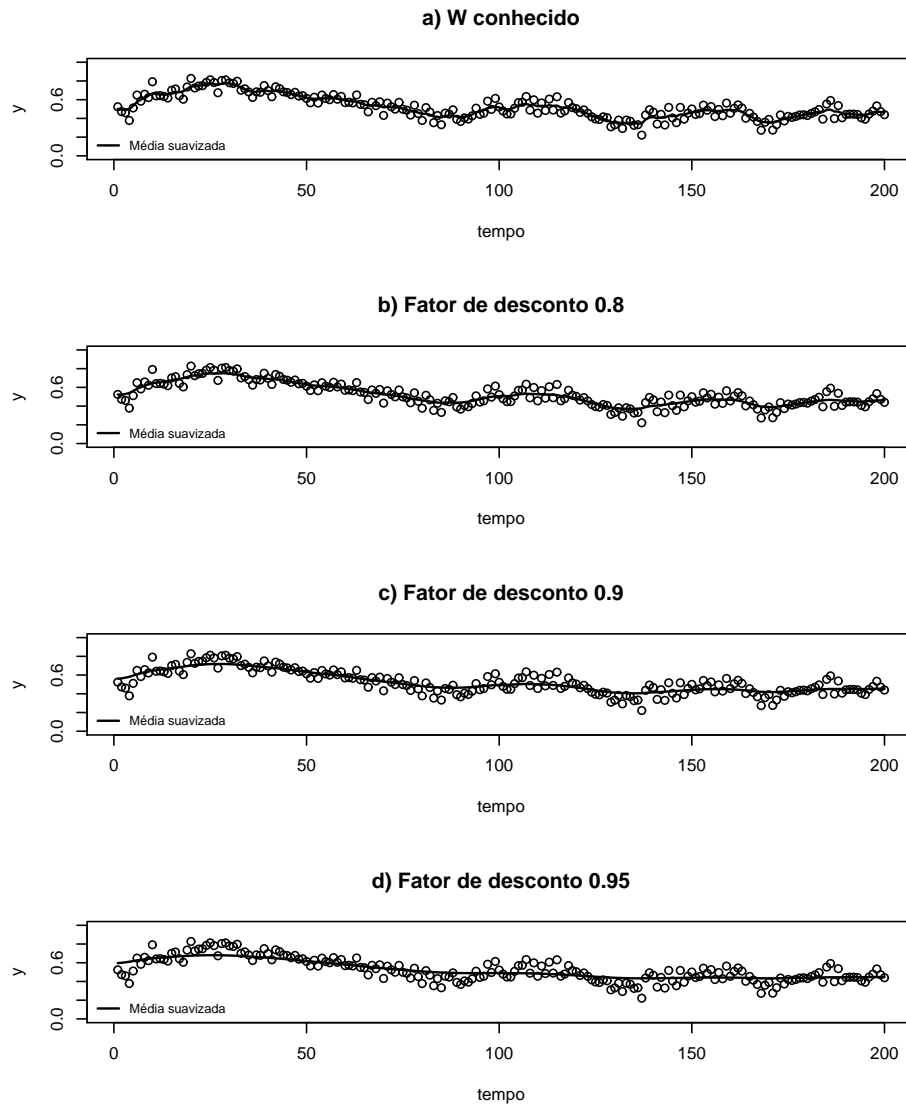


Figura 6.1: Caso 4 - Série com $W = 0.01$ e $\phi = 100$.

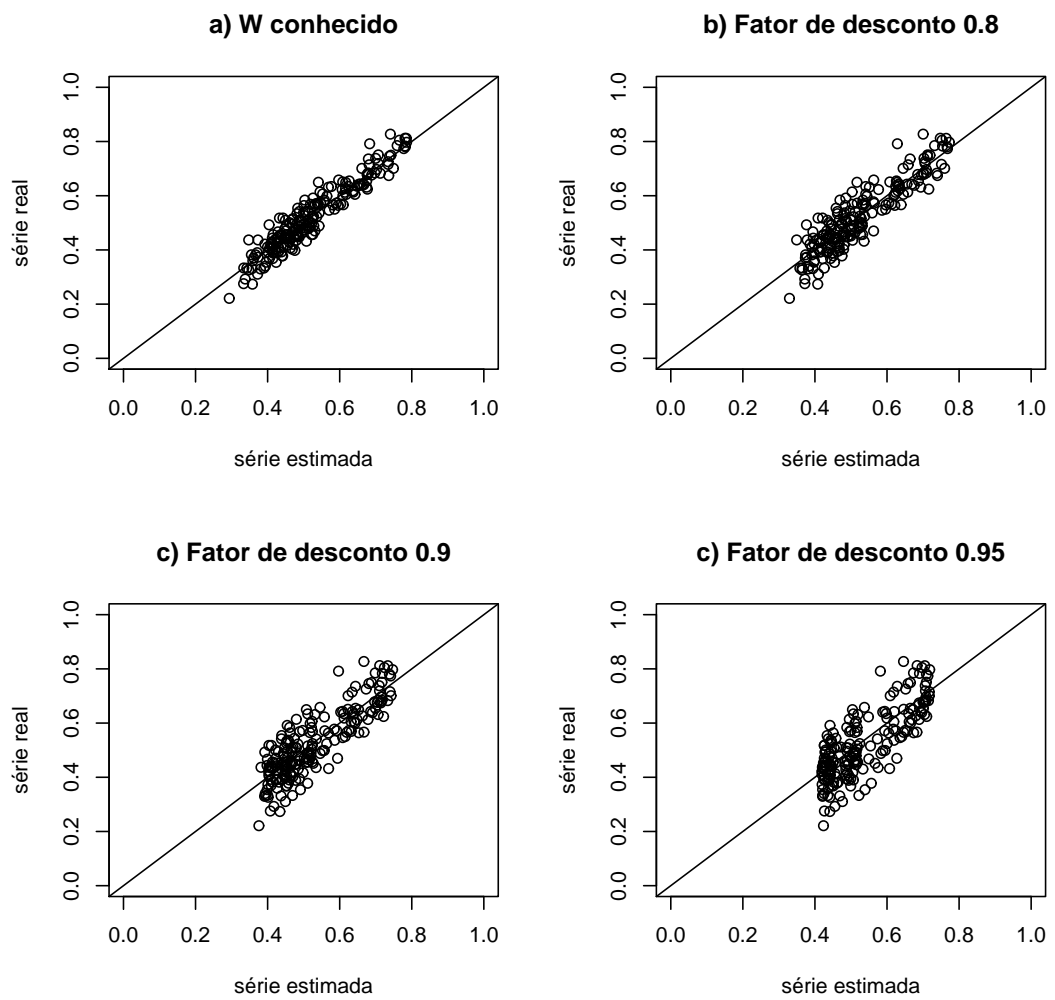


Figura 6.2: Caso 4 - Gráfico de Dispersão da série com $W = 0.01$ e $\phi = 100$.

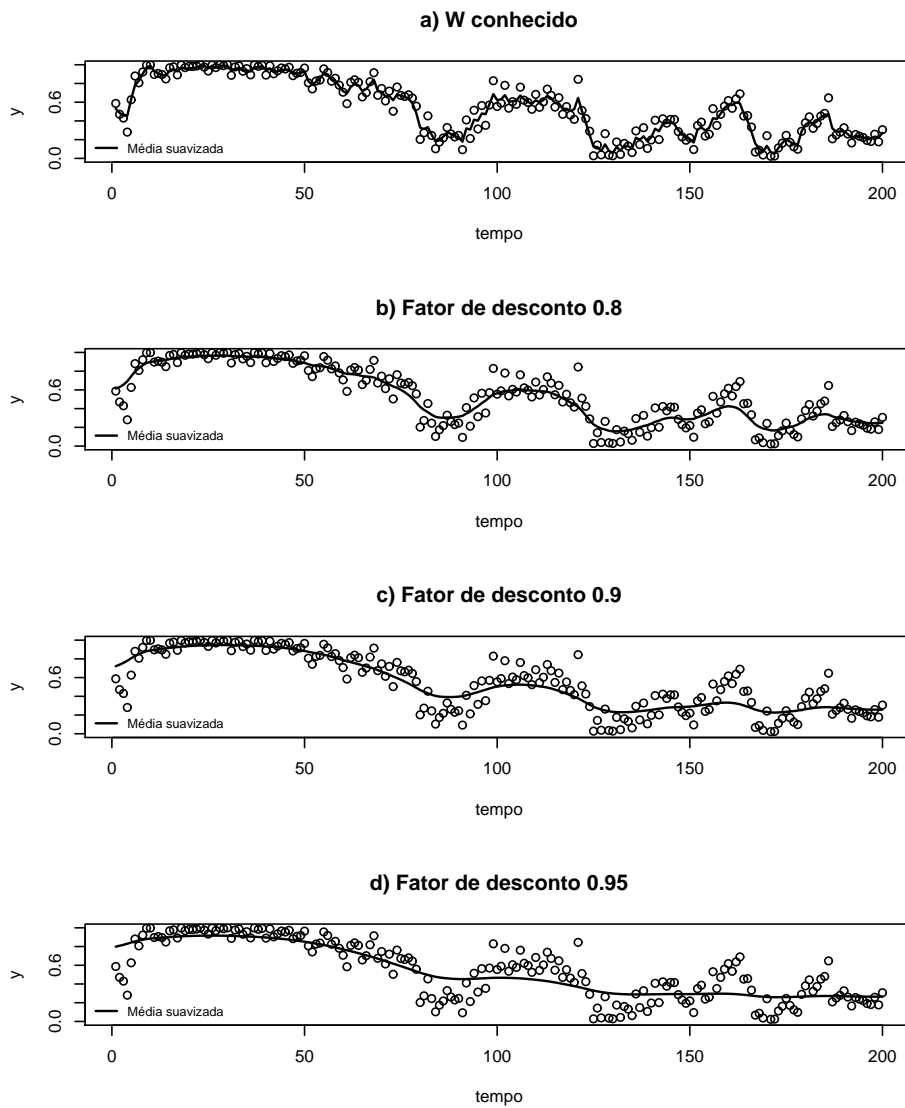


Figura 6.3: Caso 9 - Série com $W = 0.15$ e $\phi = 25$.

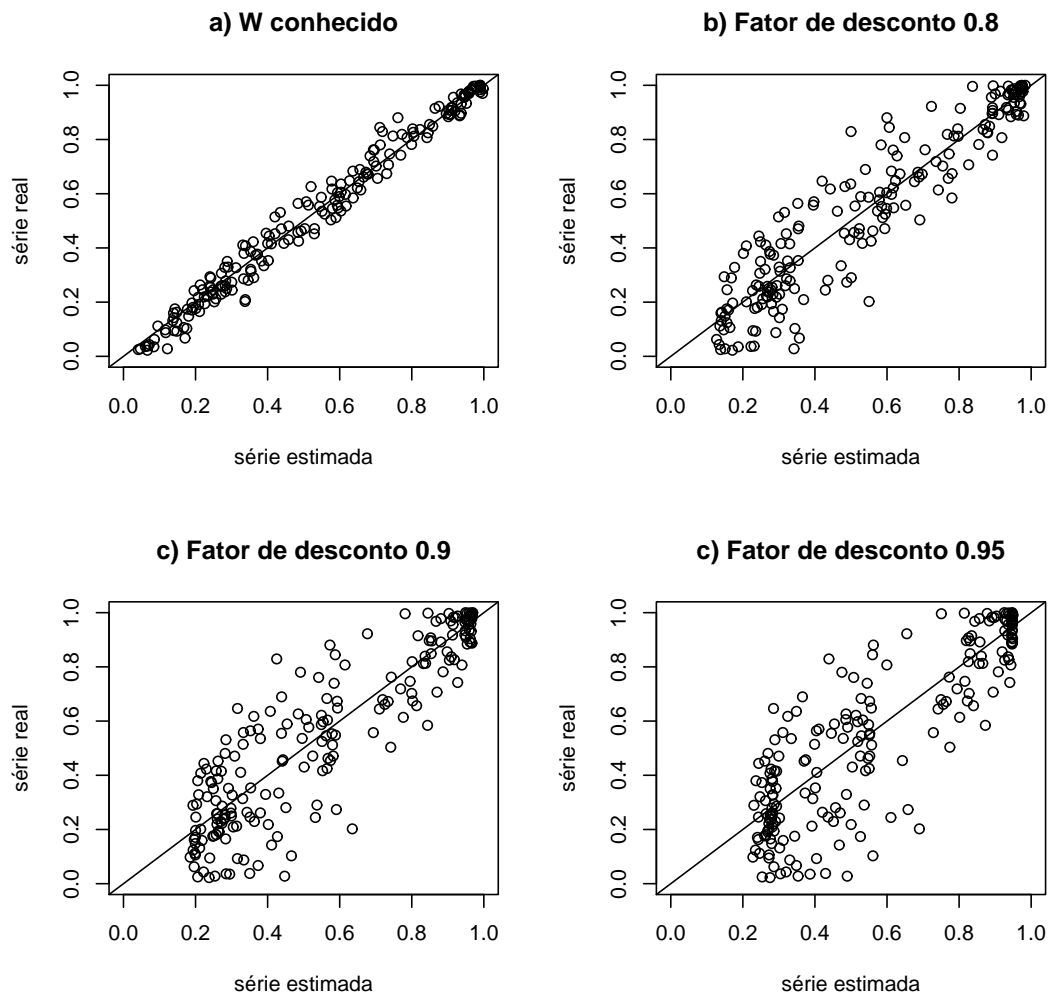


Figura 6.4: Caso 9 - Gráfico de dispersão da série com $W = 0.15$ e $\phi = 25$.

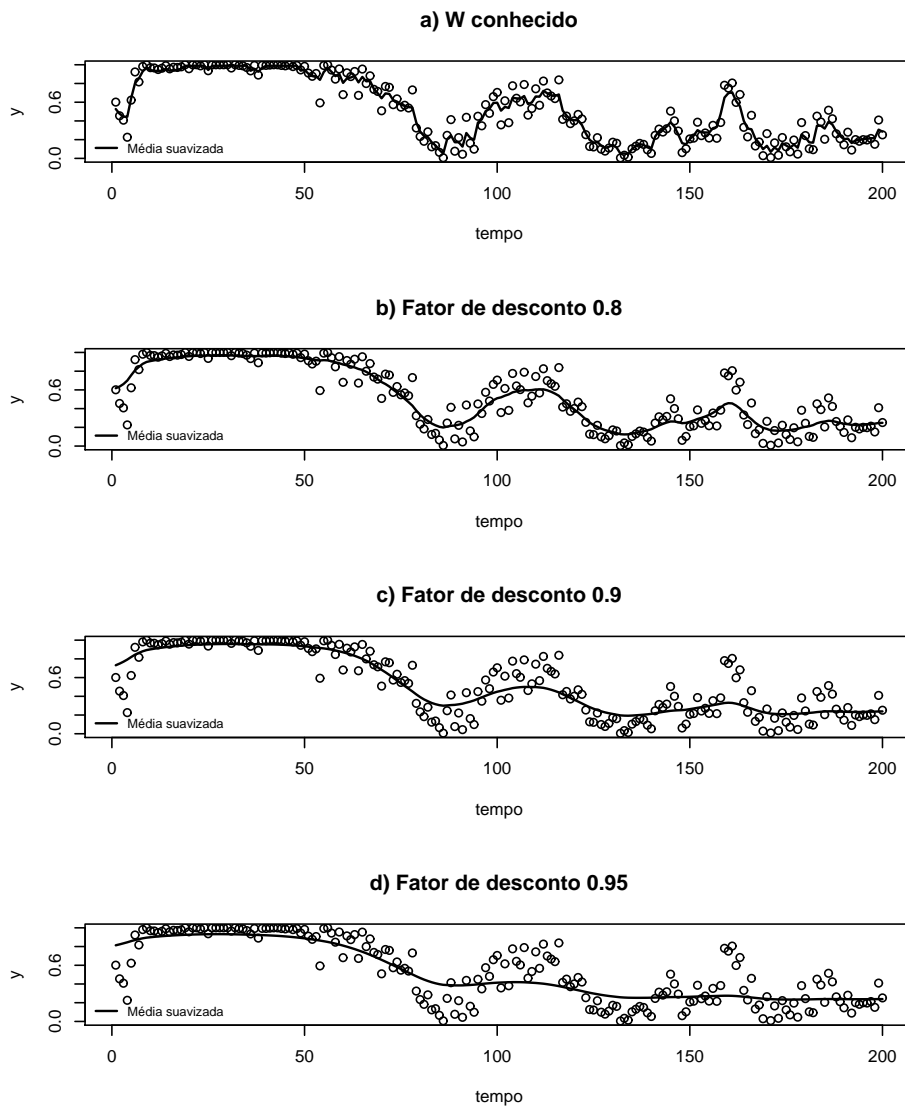


Figura 6.5: Caso 10 - Série com $W = 0.20$ e $\phi = 15$.

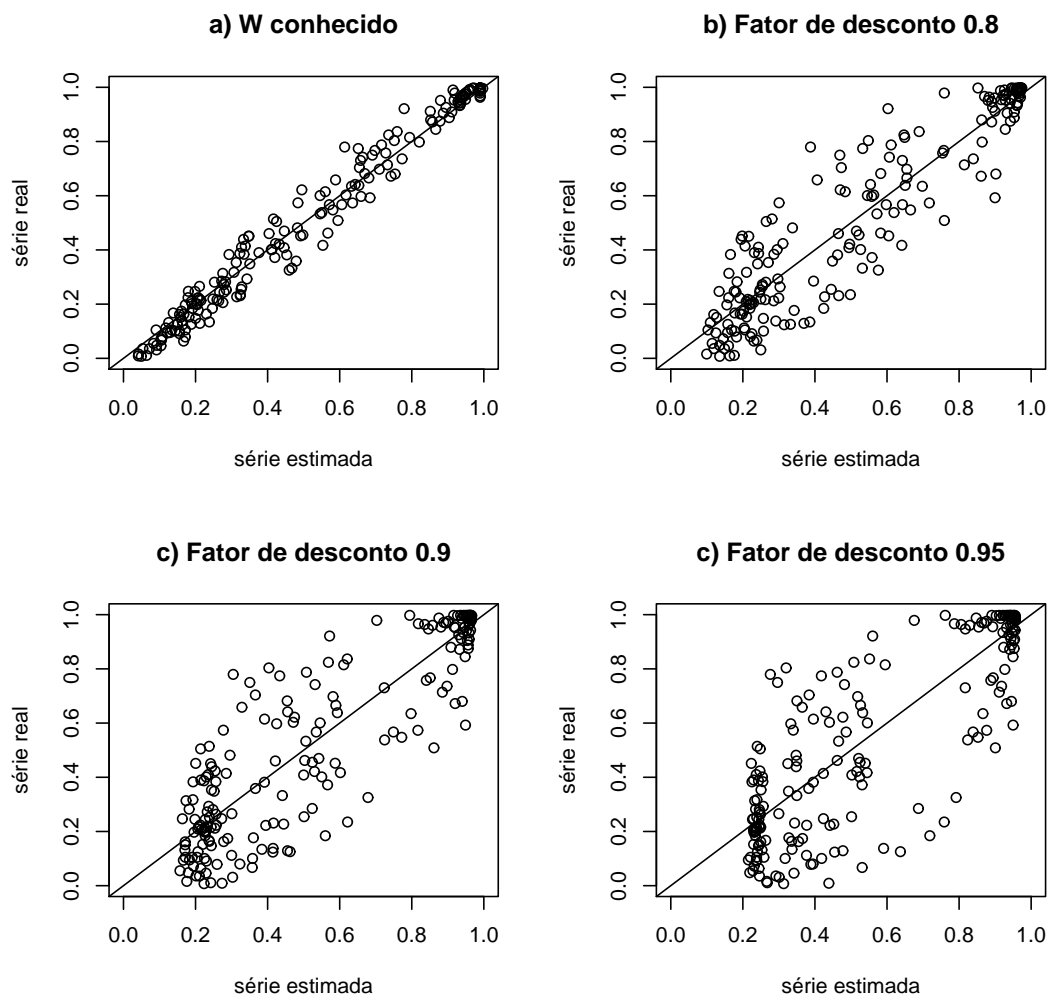


Figura 6.6: Caso 10 - Gráfico de dispersão da série com $W = 0.2$ e $\phi = 15$.

6.4 Comparações Entre Métodos de Estimação

Na Seção 5.1 (Passo 4) foram apresentados dois métodos para estimar os momentos a posteriori de μ_t , isto é, $E(\mu_t|D_t, \phi)$ e $V(\mu_t|D_t, \phi)$, o método da moda a posteriori e o método de Laplace, que é supostamente mais acurado que o primeiro.

Com o intuito de verificar se há ganho significativo ao utilizar a aproximação de Laplace e a expansão de Taylor de 2ª ordem nas estimativas dos momentos de $(\mu_t|D_t)$, foram ajustados modelos dinâmicos Beta utilizando ambas as aproximações. Para avaliar possíveis diferenças, foram calculadas estatísticas descritivas dos resíduos obtidos por ambos os métodos, além das análises gráficas dos resíduos por meio de gráficos de dispersão e gráficos das séries estimadas. Os resíduos da série são obtidos através de:

$$e_t = \mu_t - \tilde{\mu}_t$$

onde μ_t representa o valor real da série e $\tilde{\mu}_t$ o seu valor estimado no tempo t .

Como critério na comparação entre os dois tipos de aproximações utilizadas na estimação dos momentos de $(\mu_t|D_t, \phi)$, utilizou-se o erro quadrático médio (6.1).

As análises que seguem foram feitas para os casos 4, 9 e 10. No Caso 4 (séries bem comportadas), os valores de EQM apresentados na Tabela 6.3 evidenciam que o procedimento mais complexo, com aproximação de Laplace e expansão de 2ª ordem, proporcionam estimativas mais precisas de μ_t . No entanto, tal vantagem é bastante modesta. A comparação gráfica das séries estimadas sob os dois procedimentos e os diagramas de dispersão em que são confrontados μ_t e $\tilde{\mu}_t$ (vide Figura 6.7) tornam claro que, para o Caso 4, não há ganho significativo na precisão das estimativas quando utiliza-se o procedimento mais complexo.

No Caso 9 (séries razoavelmente estáveis), os valores de EQM apresentados na Tabela 6.3, indicam ligeira superioridade nas estimativas obtidas com o procedimento mais complexo. No entanto, os gráficos inclusos na Figura 6.8, novamente evidenciam que o uso do procedimento mais simples pode ser adotado sem prejudicar o ajuste de forma significativa.

No Caso 10 (séries com fortes oscilações), já é possível notar alguma diferença entre as séries. Ao observar a Figura 6.9 nota-se que há uma dispersão menor dos pontos quando a estimativa é feita utilizando-se as aproximações de Laplace e expansão de Taylor de 2ª ordem. Neste caso, o ganho ao utilizar um procedimento mais complexo torna-se relevante, o que leva a concluir que o uso de aproximações mais robustas é interessante quando se trabalha com séries que apresentam oscilações mais expressivas. Pela Figura 6.9 é possível notar que graficamente as séries se comportam de forma ainda um tanto similar pelos dois procedimentos e a diferença é percebida principalmente nas regiões próximas as extremidades do espaço paramétrico (0 e 1).

Casos	Método de Estimação	min	mediana	max	média	$EQM \times 10^2$
Caso 4	Aproximação de Laplace	-0.095	-0.0010	0.076	-0.004	0.090
	Moda a posteriori	-0.096	-0.0004	0.075	-0.004	0.093
Caso 9	Aproximação de Laplace	-0.208	-0.003	0.241	-0.0002	0.443
	Moda a posteriori	-0.212	-0.001	0.247	-0.0007	0.453
Caso 10	Aproximação de Laplace	-0.262	-0.003	0.226	-0.0099	0.528
	Moda a posteriori	-0.273	-0.007	0.248	-0.0171	0.699

Tabela 6.3: Algumas estatísticas para os erros $e_t = [\mu_t - \tilde{\mu}_t]$ e erro quadrático médio (EQM) entre os níveis reais e obtidos pelo Modelo Dinâmico Beta estimados usando tanto a moda a posteriori e aproximações de Taylor de primeira ordem ou aproximação de Laplace e aproximações de Taylor segunda ordem para os momentos $(\mu_t | D_t, \phi)$.

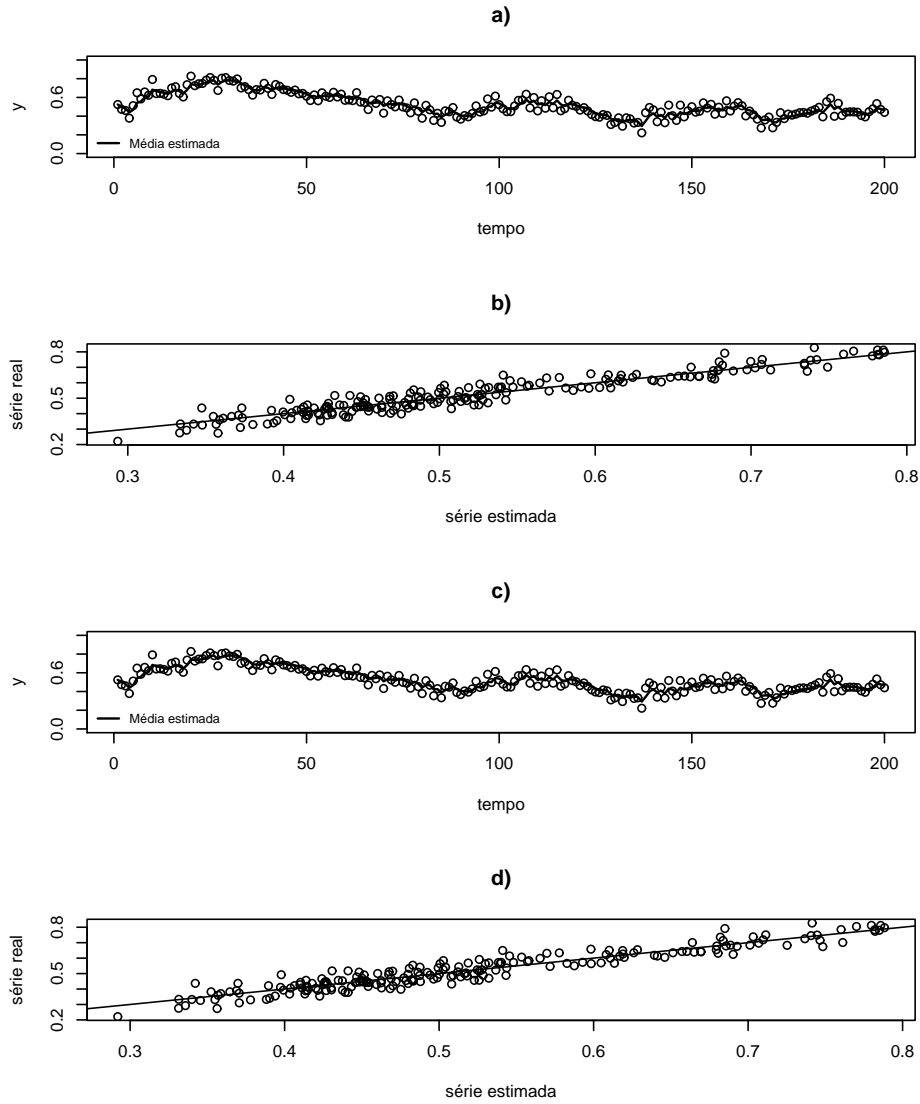


Figura 6.7: Caso 4 - Série com $W = 0.01$ e $\phi = 100$: a) Série estimada via aproximação de Laplace e expansão de 2ª ordem; b) Dispersão da série estimada via aproximação de Laplace e expansão de 2ª ordem; c) Série estimada via moda a posteriori e expansão de 1ª ordem; d) Dispersão da série estimada via moda a posteriori e expansão de 1ª ordem.

Como recomendação prática sugerimos o uso da aproximação da moda a posteriori na estimação dos momentos $E(\mu_t|D_t, \phi)$ e $V(\mu_t|D_t, \phi)$.

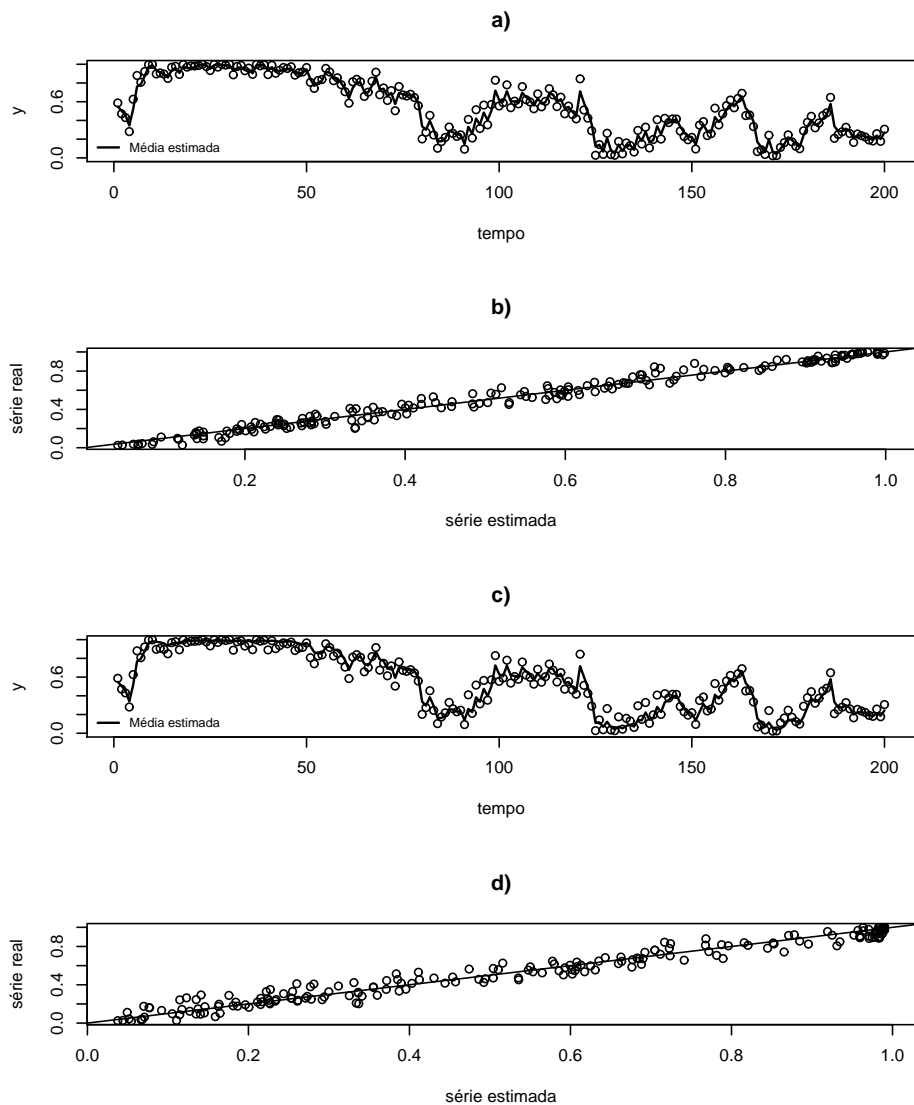


Figura 6.8: Caso 9 - Série com $W = 0.15$ e $\phi = 25$: a) Série estimada via aproximação de Laplace e expansão de 2^a ordem; b) Dispersão da série estimada via aproximação de Laplace e expansão de 2^a ordem; c) Série estimada via moda a posteriori e expansão de 1^a ordem; d) Dispersão da série estimada via moda a posteriori e expansão de 1^a ordem.

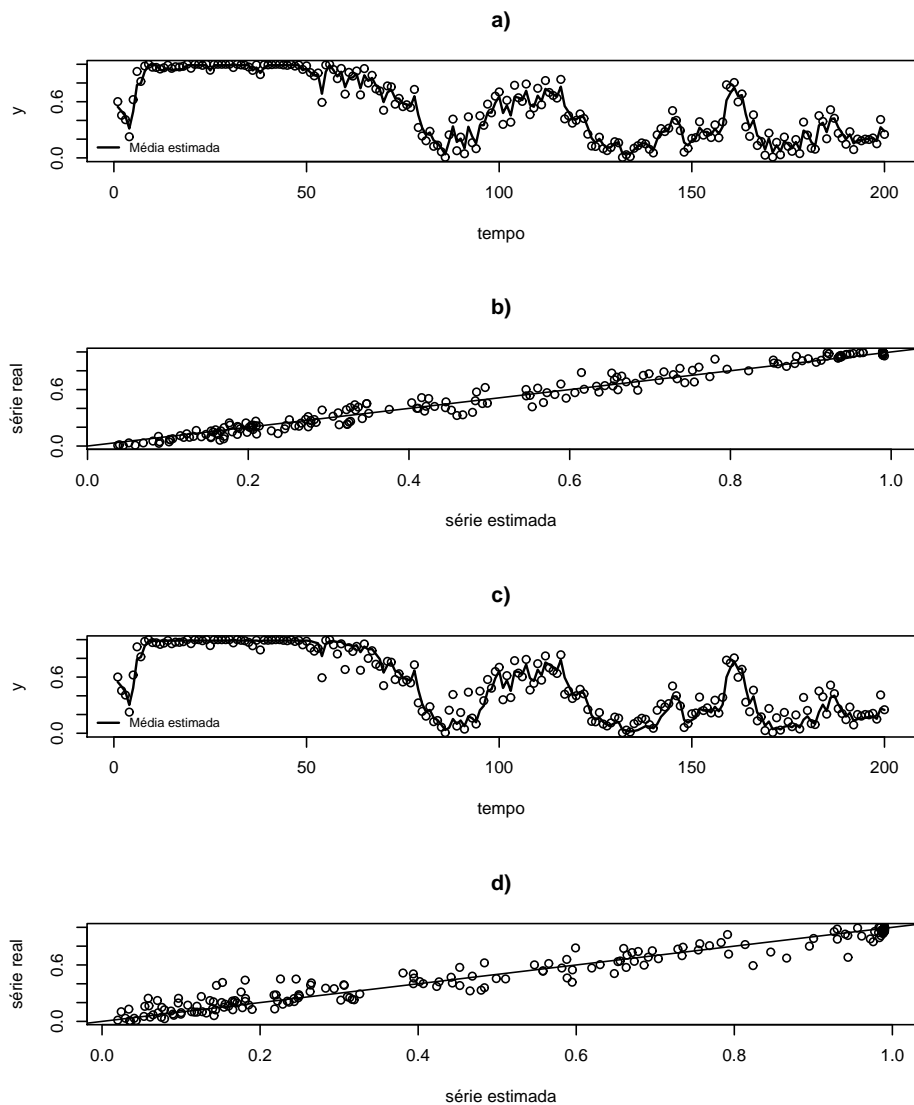


Figura 6.9: Caso 10 - Série com $W = 0.2$ e $\phi = 15$: a) Série estimada via aproximação de Laplace e expansão de 2ª ordem; b) Dispersão da série estimada via aproximação de Laplace e expansão de 2ª ordem; c) Série estimada via moda a posteriori e expansão de 1ª ordem; d) Dispersão da série estimada via moda a posteriori e expansão de 1ª ordem.

6.5 Analisando dados simulados - ϕ desconhecido

Considerando a metodologia descrita na Seção 5.2, observou-se que quando o valor de W é conhecido, o procedimento *ad-hoc* conduz a valores próximos do ϕ real. Mesmo para valores elevados de W (como 0.2, por exemplo) e valores baixos de ϕ (como 15), correspondendo aos casos de séries mal comportadas, foi possível obter estimativas próximas do ϕ real.

Como em geral não se sabe o valor de W_t , busca-se estimar ϕ em um “grid” de valores fixos para W . Neste “grid” foram atribuídos valores variando de 0.01 a 0.25, e para em cada W_t , foi encontrado uma estimativa $\tilde{\phi}$ de ϕ como o ponto que maximiza a *pseudo-verossimilhança* L_p . Desta forma, são obtidos 25 possíveis valores para $\hat{\phi}$ (um correspondente a cada valor de W_t pertencente ao “grid”). Para chegar à estimativa final de ϕ , um segundo ajuste é realizado levando em consideração $(\tilde{\phi}_1, \dots, \tilde{\phi}_{25})$, os 25 valores obtidos para $\hat{\phi}$. Neste segundo estágio da estimativa de ϕ o modelo é ajustado considerando 4 alternativas:

- valor mínimo de $(\tilde{\phi}_1, \dots, \tilde{\phi}_{25})$;
- média de $(\tilde{\phi}_1, \dots, \tilde{\phi}_{25})$: $\hat{\phi} = (\sum_{i=1}^{25} \tilde{\phi}_i)/25$;
- mediana de $(\tilde{\phi}_1, \dots, \tilde{\phi}_{25})$;
- valor máximo de $(\tilde{\phi}_1, \dots, \tilde{\phi}_{25})$.

Dentre estes valores estimados, $\hat{\phi}$ é escolhido como aquele entre que melhor se ajusta a série no segundo estágio.

A Figura 6.10 mostra o impacto de utilizar $\hat{\phi}$ estimado, respectivamente, pelo valor mínimo, médio, mediano e máximo de $(\tilde{\phi}_1, \dots, \tilde{\phi}_{25})$. Percebe-se que o Modelo Dinâmico Beta é bastante robusto em relação aos valores das estimativas de ϕ , resultado bastante importante para os analistas que irão trabalhar com o Modelo Dinâmico Beta.

Na Tabela 6.4 são apresentadas algumas estatísticas para os resíduos e o erro quadrático médio para os casos simulados onde o ajuste foi feito utilizando W_t conhecido em ambos os casos de modo que em um é utilizado ϕ conhecido e no outro $\hat{\phi}$ estimado como descrito na Seção 5.2. Percebe-se, pelas estatísticas dos

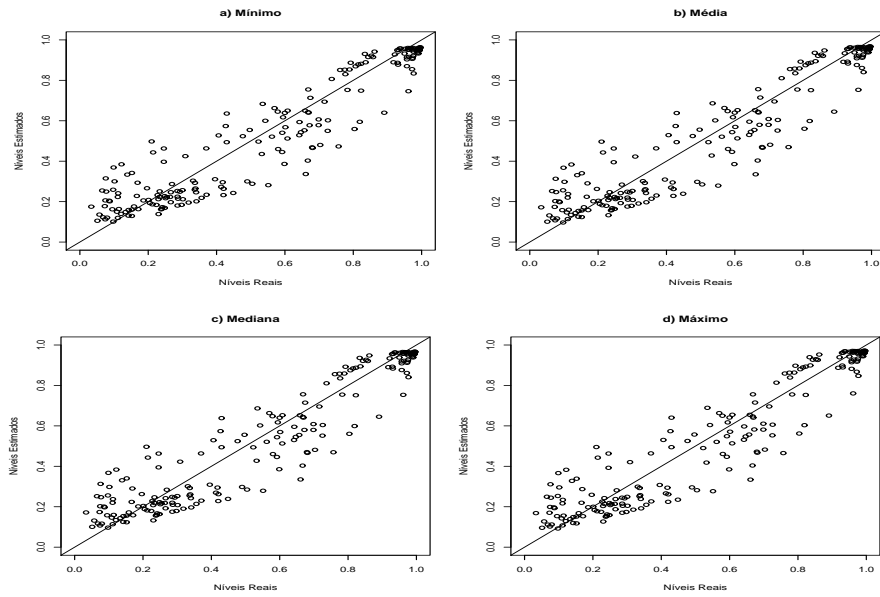


Figura 6.10: Caso 10 - Gráficos de dispersão dos níveis reais e estimados para a séries Beta geradas quando ϕ é estimado utilizando $\hat{\phi}_{\min}$, $\hat{\phi}_{\text{média}}$, $\hat{\phi}_{\text{mediana}}$ e $\hat{\phi}_{\text{máx}}$.

resíduos, que o modelo ajustado com $\hat{\phi}$ apresentam comportamentos bastante semelhantes aos ajustes realizados com o parâmetro ϕ verdadeiro. Nota-se que o EQM pouco aumenta ao utilizar o ajuste com $\hat{\phi}$ nos casos em que $W_t = 0.01$. No entanto, na medida que cresce o valor de W_t , a perda ao utilizar $\hat{\phi}$ se torna mais expressiva.

Os valores apresentados na Tabela 6.5 correspondem às estimativas $\hat{\phi}_{\min}$, $\hat{\phi}_{\text{médio}}$, $\hat{\phi}_{\text{mediano}}$ e $\hat{\phi}_{\text{máx}}$ juntamente com seus respectivos EQM . Nota-se que o parâmetro ϕ é bastante robusto, de modo que o erro quadrático médio pouco varia para os diferentes valores estimados. Ao observar as duas últimas colunas da Tabela 6.5, é possível perceber que nos casos com em que ϕ apresenta valores mais elevados, ou seja, em séries com baixa dispersão (Casos 3, 4 e 7).

Casos	ϕ	min	mediana	max	média	$EQM \times 10^2$
Caso 1	Conhecido	-0.137500	0.011870	0.187100	0.017450	0.3140462
	Estimado	-0.137500	0.011870	0.187100	0.017450	0.3236349
Caso 2	Conhecido	-0.119200	0.011810	0.160500	0.015690	0.2415315
	Estimado	-0.115200	0.010140	0.156000	0.015720	0.2437977
Caso 3	Conhecido	-0.091600	0.005729	0.124800	0.007919	0.1393307
	Estimado	-0.090760	0.005171	0.123500	0.007920	0.1393087
Caso 4	Conhecido	-0.073980	0.003302	0.096870	0.006113	0.0918570
	Estimado	-0.076840	0.004007	0.097330	0.006098	0.0924318
Caso 5	Conhecido	-0.237400	0.018030	0.250600	0.009815	0.5170380
	Estimado	-0.305700	0.014650	0.254300	0.006504	0.7935885
Caso 6	Conhecido	-0.145500	0.009657	0.208800	0.005431	0.3335225
	Estimado	-0.211800	0.013990	0.237700	0.003277	0.5358845
Caso 7	Conhecido	-0.104400	0.002698	0.152800	0.005626	0.2362833
	Estimado	-0.210200	0.003893	0.189500	0.003284	0.4703508
Caso 8	Conhecido	-0.235800	0.005453	0.263700	0.004609	0.589311
	Estimado	-0.294300	0.012140	0.268800	0.003055	0.838767
Caso 9	Conhecido	-0.238900	0.006779	0.209900	0.002317	0.4448096
	Estimado	-0.370500	0.013880	0.242300	0.000833	0.9199885
Caso 10	Conhecido	-0.222500	0.005312	0.264300	0.012600	0.533168
	Estimado	-0.200300	0.009745	0.290000	0.011560	0.562797

Tabela 6.4: Comparação dos resíduos com ϕ conhecido e ϕ estimado nos Modelos Dinâmicos Beta

Casos		Estatísticas de $\hat{\phi}$ e $EQM \times 10^2$				ϕ e W_t conhecidos	ϕ conhecido $\delta = 0.8$
		com ϕ e W_t desconhecidos ($\delta = 0.8$)					
		$\hat{\phi}_{min}$	$\hat{\phi}_{m\u00e9dio}$	$\hat{\phi}_{mediano}$	$\hat{\phi}_{m\u00e1x}$		
Caso 1	EQM	0.3142	0.3155	0.3142	0.3241	0.3140	0.3108
$\phi = 15; W_t = 0.01$	$\hat{\phi}$	14	14.84	14	22		
Caso 2	EQM	0.2440	0.2448	0.2445	0.2430	0.2415	0.2408
$\phi = 25; W_t = 0.01$	$\hat{\phi}$	22	23.52	23	32		
Caso 3	EQM	0.1596	0.1599	0.1598	0.1605	0.1370	0.1580
$\phi = 50; W_t = 0.01$	$\hat{\phi}$	39	44.24	42	56		
Caso 4	EQM	0.1332	0.1335	0.1334	0.1339	0.0918	0.1319
$\phi = 100; W_t = 0.01$	$\hat{\phi}$	70	108.24	88	200		
Caso 5	EQM	0.7712	0.7711	0.7712	0.7712	0.5170	0.7617
$\phi = 15; W_t = 0.09$	$\hat{\phi}$	13	13.44	13	17		
Caso 6	EQM	0.5846	0.5832	0.5836	0.5811	0.3335	0.5740
$\phi = 25; W_t = 0.09$	$\hat{\phi}$	19	20.35	20	23		
Caso 7	EQM	0.6556	0.6536	0.6536	0.6524	0.2363	0.6545
$\phi = 50; W_t = 0.09$	$\hat{\phi}$	27	34.96	35	44		
Caso 8	EQM	1.0253	1.0242	1.0244	1.0264	0.5893	1.0145
$\phi = 15; W_t = 0.15$	$\hat{\phi}$	11	13.36	13	19		
Caso 9	EQM	0.9435	0.9426	0.9427	0.9398	0.4448	0.9314
$\phi = 25; W_t = 0.15$	$\hat{\phi}$	18	19.20	19	22		
Caso 10	EQM	1.2153	1.2141	1.2141	1.2145	0.5331	1.2071
$\phi = 15; W_t = 0.20$	$\hat{\phi}$	11	13.79	14	19		

Tabela 6.5: Comparação das EQM 's e valores estimados de ϕ para os diferentes casos e características de W_t e ϕ nos Modelos Dinâmicos Beta

6.6 Modelo Dinâmico Beta versus Modelo Linear Dinâmico Normal Matriz-Variado

Com o intuito de verificar o ganho do *Modelo Dinâmico Beta (MDB)* comparado com o *Modelo Linear Dinâmico Normal Matriz-Variado (MLDNM)* com transformação razão-log, descrito no Capítulo 2, as séries de proporções simuladas foram estimadas por meio dos dois modelos. No caso do modelo *Normal Matriz Variado*, foi utilizada a transformação *logística*:

$$Z_t = \log \left(\frac{Y_t}{1 - Y_t} \right), \quad \text{onde } Y_t \sim \text{Beta}(\phi\mu_t, \phi(1 - \mu_t)).$$

Utilizou-se uma aproximação Normal para Z_t , de modo que o procedimento descrito no Capítulo 2 foi aplicado aos dados simulados com distribuição Beta. As Figuras 6.11 a 6.13 ilustram os ajustes para dados simulados sob os casos 4, 9 e 10 (vide Seção 6.3) utilizando os modelos dinâmicos normal matriz-variado e beta. Na parte superior de cada figura estão graficados os ajustes para o modelo dinâmico normal matriz-variado enquanto na parte inferior os ajustes para o modelo dinâmico beta. Ambos os ajustes foram realizados sob ϕ e W_t conhecidos. Em todos os casos observa-se, como esperado, um melhor ajuste quando o modelo dinâmico beta é utilizado.

De acordo com a Tabela 6.6, a vantagem do Modelo Dinâmico Beta é destacada, principalmente, nos modelos em que a variância dos erros da equação de observações, V_t , é pequena, ou seja, para grandes valores de ϕ . Na medida que a razão sinal-ruído diminui, o ganho do modelo beta torna-se ainda mais evidente em relação ao modelo Normal (vide Caso 4). Quando a razão sinal-ruído é alta (Casos 9 e 10), os modelos Beta deixam de apresentar uma vantagem tão evidente em seus ajustes. No entanto, ainda assim, o modelo dinâmico beta apresentou menor erro quadrático médio e valores distribuídos de forma mais simétrica em torno de *zero*, como era de se esperar.

Casos	Modelo	mínimo	mediana	máximo	média	$EQM \times 10^2$
Caso 4	MLDNM	-0.125	-0.0021	0.153	0.0001	0.268
	MDB	-0.076	0.0010	0.095	0.0040	0.090
Caso 9	MLDNM	-0.328	-0.0005	0.238	-0.0012	0.592
	MDB	-0.241	0.0026	0.208	-0.0002	0.443
Caso 10	MLDNM	-0.206	-0.0034	0.288	0.0109	0.730
	MDB	-0.226	0.0036	0.262	0.0100	0.5285

Tabela 6.6: Algumas Estatísticas dos erro $e_t = \mu_t - \tilde{\mu}_t$ e o erro quadrático médio (EQM) para comparar os ajustes das séries beta analisadas pelo Modelo Dinâmico Beta ou pelo Modelo Linear Dinâmico Normal Matriz-Variado.

6.7 Aplicação do Modelo Dinâmico Beta Em Dados Reais

6.7.1 Taxa Mensal de Desemprego no Brasil

O Instituto Brasileiro de Geografia e Estatística (IBGE) implementou em 1980 a Pesquisa Mensal de Emprego (PME), mas desde de 2002 uma nova metodologia tem sido adotada. A PME é uma pesquisa mensal relativa à força de trabalho e renda. As cidades metropolitanas mais relevantes no Brasil são incluídas nesta pesquisa, como São Paulo, Rio de Janeiro, Belo Horizonte, Porto Alegre, Recife e Salvador. A série pode ser obtida no site:

<http://www.ibge.gov.br>.

A taxa de desemprego mensal foi analisada baseada na série obtida pela PME a partir de março de 2002 até julho de 2009 (89 observações disponíveis). As taxas são expressas em porcentagem, de modo que o Modelo Dinâmico Beta foi utilizado no ajuste dos dados. Foram obtidos os seguintes valores estimados, considerando-se um modelo dinâmico de 1ª ordem em que $\beta_t = G_t\beta_{t-1} + \omega_t$ com $G_t = 1$ e $F_t = 1$.

– $\hat{W} = 0.02$; $\hat{\phi} = 131$; Fator de desconto $\delta = 0.7$.

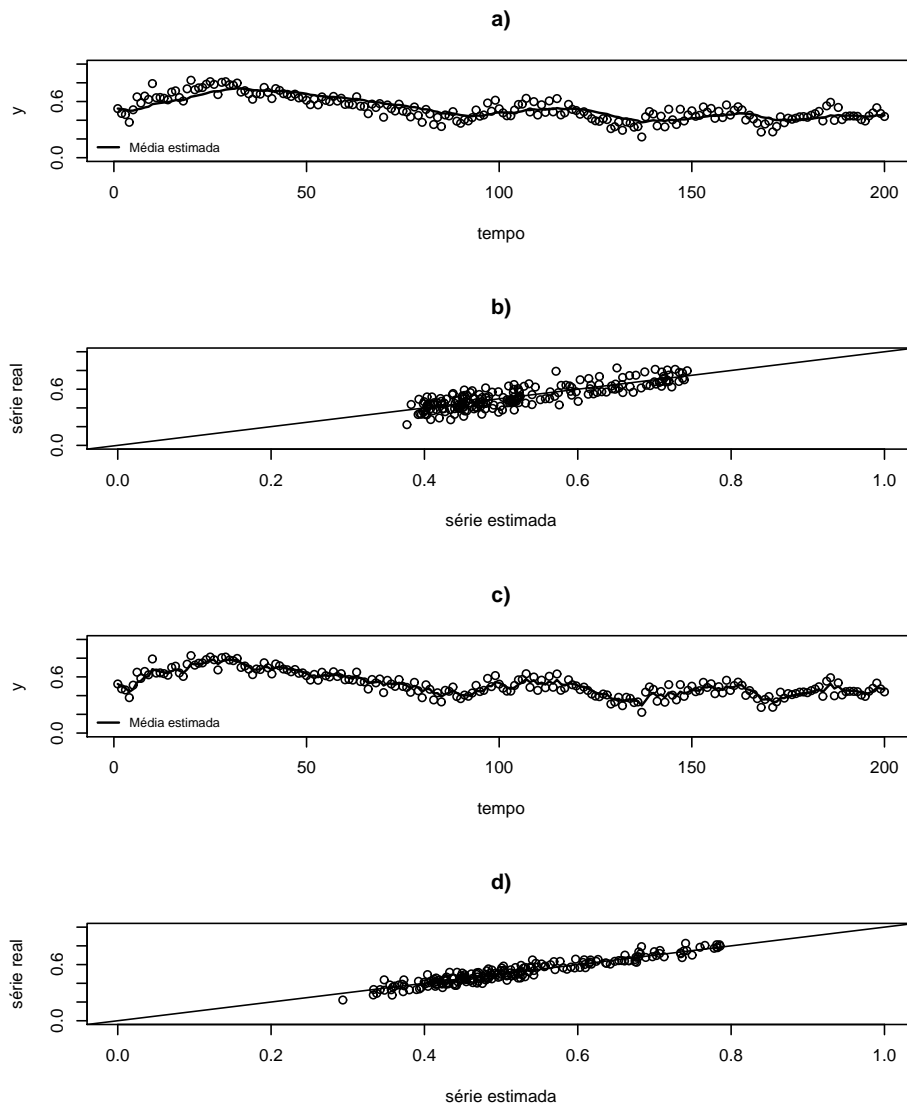


Figura 6.11: Caso 4 - Série com $W = 0.01$ e $\phi = 100$: a) Série estimada via MLDNM; b) Dispersão da série estimada via MLDNM; c) Série estimada via MDB; d) Dispersão da série estimada via MDB.

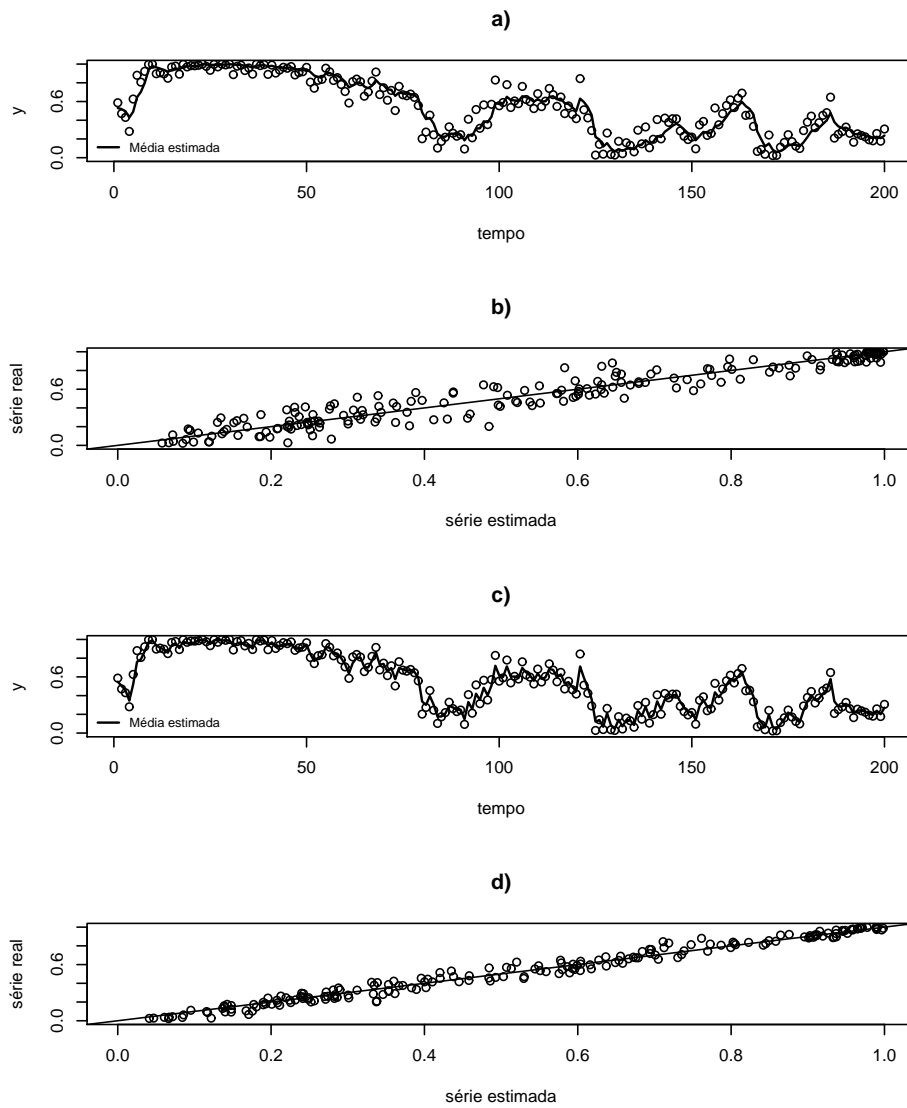


Figura 6.12: Caso 9 - Série com $W = 0.15$ e $\phi = 25$: a) Série estimada via MLDNM; b) Dispersão da série estimada via MLDNM; c) Série estimada via MDB; d) Dispersão da série estimada via MDB.

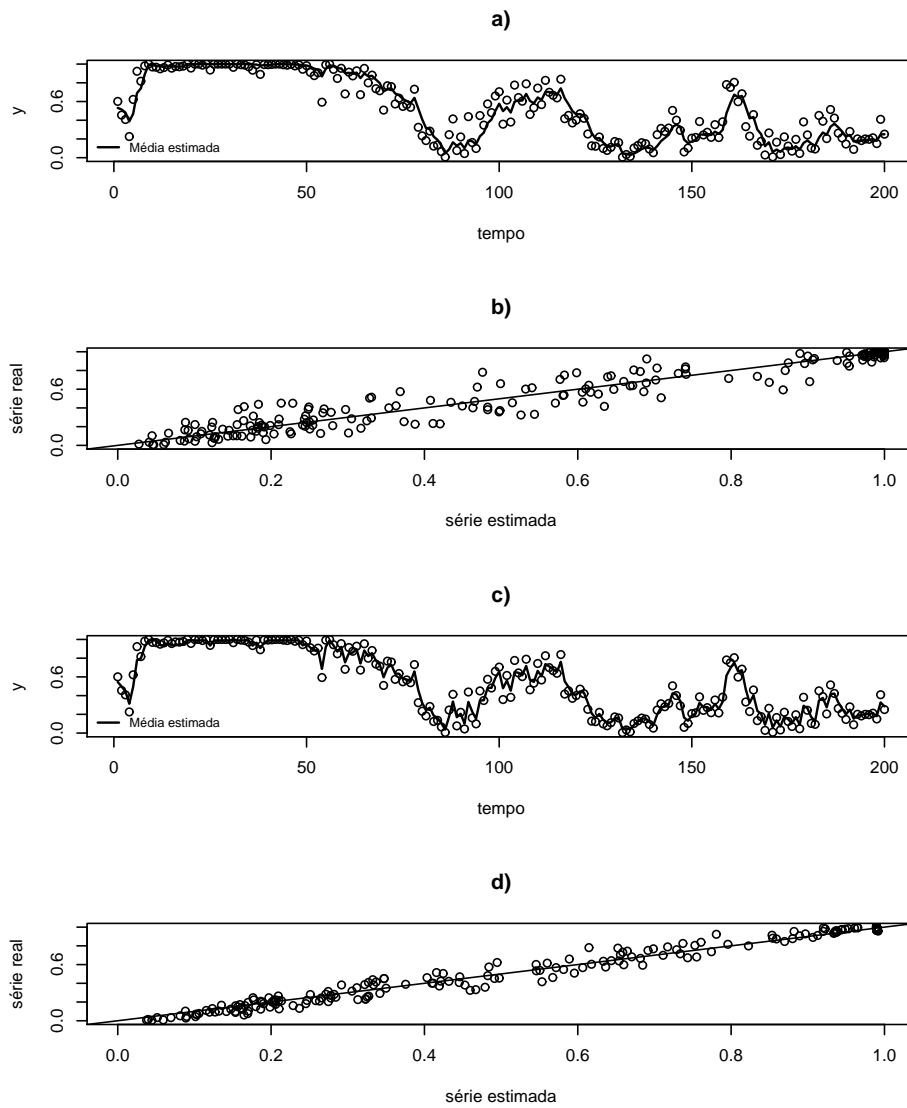


Figura 6.13: Caso 10 - Série com $W = 0.2$ e $\phi = 15$: a) Série estimada via MLDNM; b) Dispersão da série estimada via MLDNM; c) Série estimada via MDB; d) Dispersão da série estimada via MDB.

Esta primeira análise serviu para obter uma estimativa do valor de ϕ por meio da metodologia descrita na Seção 5.2.1. No entanto, é razoável assumir previamente que a taxa de desemprego apresenta um comportamento sazonal regido por ciclos anuais que ocorrem devido a festividades natalinas, férias escolares e fatores climáticos. É possível perceber quedas periódicas nos meses de novembro e dezembro devido ao aumento de empregos temporários. Além disso, a partir de fatores macroeconômicos as taxas podem apresentar um aumento ou declínio em um dado período. Levando estes pontos em consideração, foi adotado um Modelo Dinâmico Beta especificado por efeitos de tendência polinomial de segunda ordem e sazonal. O vetor de parâmetros é dado por:

$$\theta_t = (\beta_{t1}, \beta_{t2}, \psi_{t1}, \dots, \psi_{tp})',$$

onde o vetor de *tendência* é descrito por $(\beta_{t1}, \beta_{t2})'$, com β_{t1} representando a coordenada de nível, enquanto β_{t2} a coordenada que mede o grau de variação no nível. Para um ciclo sazonal de tamanho p , o seu efeito é descrito pelos parâmetros $(\psi_{t1}, \dots, \psi_{tp})'$.

A especificação do modelo é determinada por:

$$\left\{ F = \begin{pmatrix} E_2 \\ E_p \end{pmatrix}, G = \begin{pmatrix} J_2(1) & 0 \\ 0 & P \end{pmatrix}, W_t = \begin{pmatrix} W_{t,\beta} & 0 \\ 0 & W_{t,\psi} \end{pmatrix} \right\}.$$

onde W_t é uma matriz bloco-diagonal de variância-covariância e

$$E_p = (1, 0'_{p-1}), J_2(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad e \quad P = \begin{pmatrix} 0 & I_{p-1} \\ 1 & 0' \end{pmatrix}.$$

A matriz de permutações P é p -*cíclica*, ou seja $P^{np} = I_p$ e $P^{h+np} = P^h$, para $h = 0, \dots, p$ e qualquer inteiro $n \geq 0$.

A matriz W_t é descrita com auxílio dos blocos de *fatores de desconto*. Relembrando que $(\theta_{t-1} | D_{t-1}) \sim (m_{t-1}, C_{t-1})$, seja C_{Tt-1} a submatriz representando a relação de covariância a posteriori dos componentes de tendência no tempo $t = t - 1$ e C_{St-1} a submatriz de covariância a posteriori dos componentes de sazonalidade. Então, os blocos $W_{t,\beta}$ e $W_{t,\psi}$ em W_t são dados por:

$$W_{t,\beta} = \left(\frac{1 - \delta_T}{\delta_T} \right) J_2(1) C_{Tt-1} J_2(1)' \quad e \quad W_{t,\psi} = \left(\frac{1 - \delta_S}{\delta_S} \right) P C_{St-1} P',$$

onde δ_T e δ_S são os *fatores de desconto* associados a estes componentes. Mais detalhes sobre modelos sazonais podem ser encontrados no Capítulo 8 de West e Harrison (1997).

Para a série da taxa de desemprego mensal no Brasil foi utilizado $\delta_T = (0.90, 0.95)$, $\delta_S = 0.98$ e a mesma estimativa de ϕ do modelo de ordem 1, $\hat{\phi} = 131$. Pode ser observado, a partir da Figura 6.14, que as taxas de desemprego no Brasil apresentaram uma queda consistente desde 2002. De acordo com uma informação postada em 26 de janeiro de 2009, no site:

<http://brazilportal.wordpress.com/2009/01/26/>,

“A taxa de desemprego do Brasil atingiu o seu menor valor em sete anos”. A partir de uma análise econômica, isto ocorre, essencialmente, devido à contínua consolidação do ajuste macroeconômico, com a taxa de câmbio do real em 1999, juntamente com algumas medidas de políticas governamentais que foram efetivas para conter a inflação, déficits públicos e vulnerabilidade externa. A boa situação do setor bancário também ajudou a estabilizar a economia brasileira. Contudo, de acordo com alguns economistas, esta tendência favorável pode mudar com a crise econômica mundial.

A Figura 6.15 apresenta os efeitos suavizados de nível e tendência ao longo do tempo. Primeiramente é apresentada a estimativa suavizada da componente de nível, que apresenta pouco movimento ao longo do período. No entanto, é observada uma leve tendência de queda de nível. Logo abaixo segue o gráfico da estimativa da componente que mede o grau de variação do nível. Neste gráfico, observa-se que no período de 2003 a 2004 há uma mudança na direção do grau de variação, justamente no período em que a componente de nível atinge o seu valor máximo. Analisando conjuntamente os gráficos, percebe-se, nitidamente, a existência de um efeito de tendência linear na série. Por último, é apresentado o gráfico da estimativa da componente de sazonalidade. Percebe-se que de fato há uma variação cíclica que ocorre de forma regular ao longo dos anos.

Valores observados e estimados de Taxa de Mensal de Desemprego

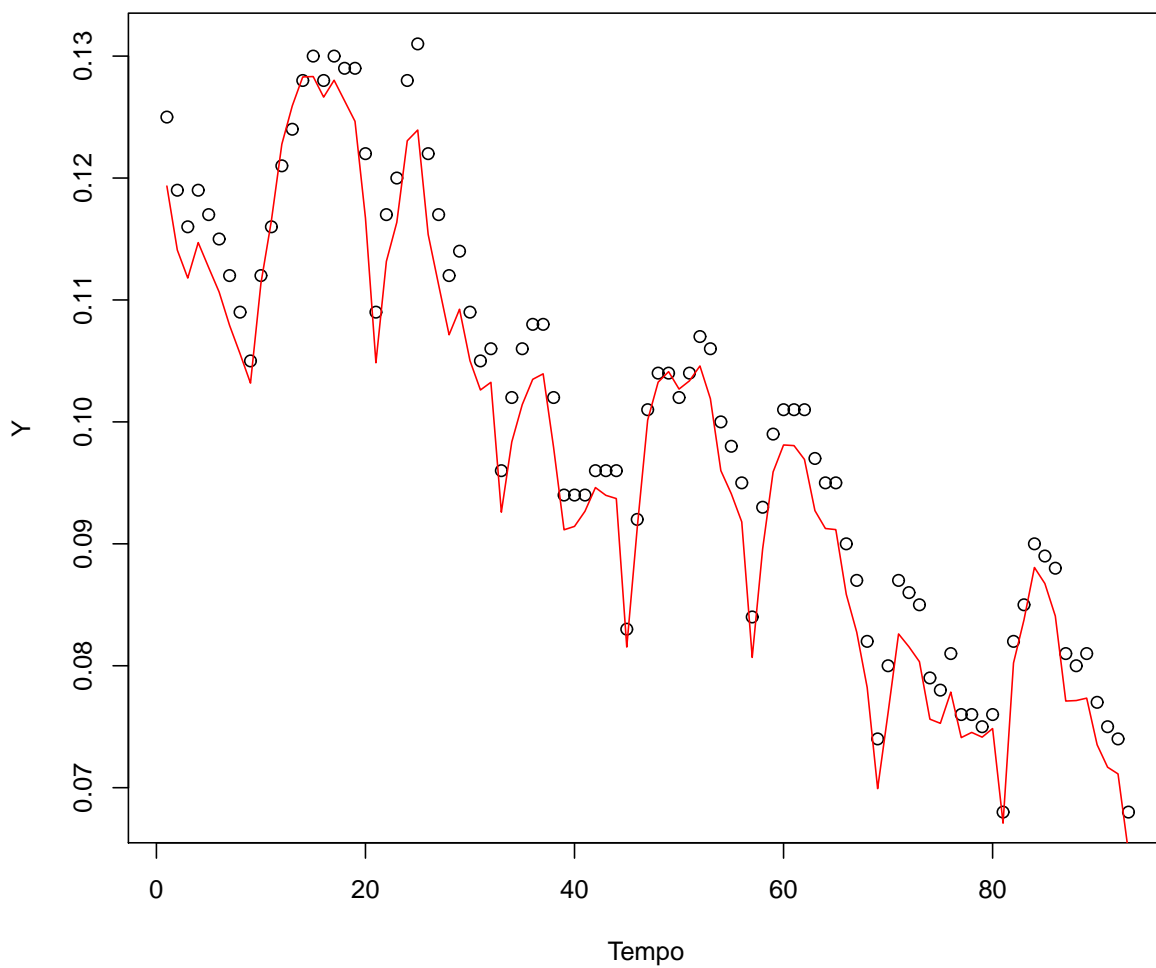


Figura 6.14: Taxa de desemprego mensal brasileira baseada nos dados da PME do período entre março de 2002 a julho de 2009. Fonte: PME/IBGE.

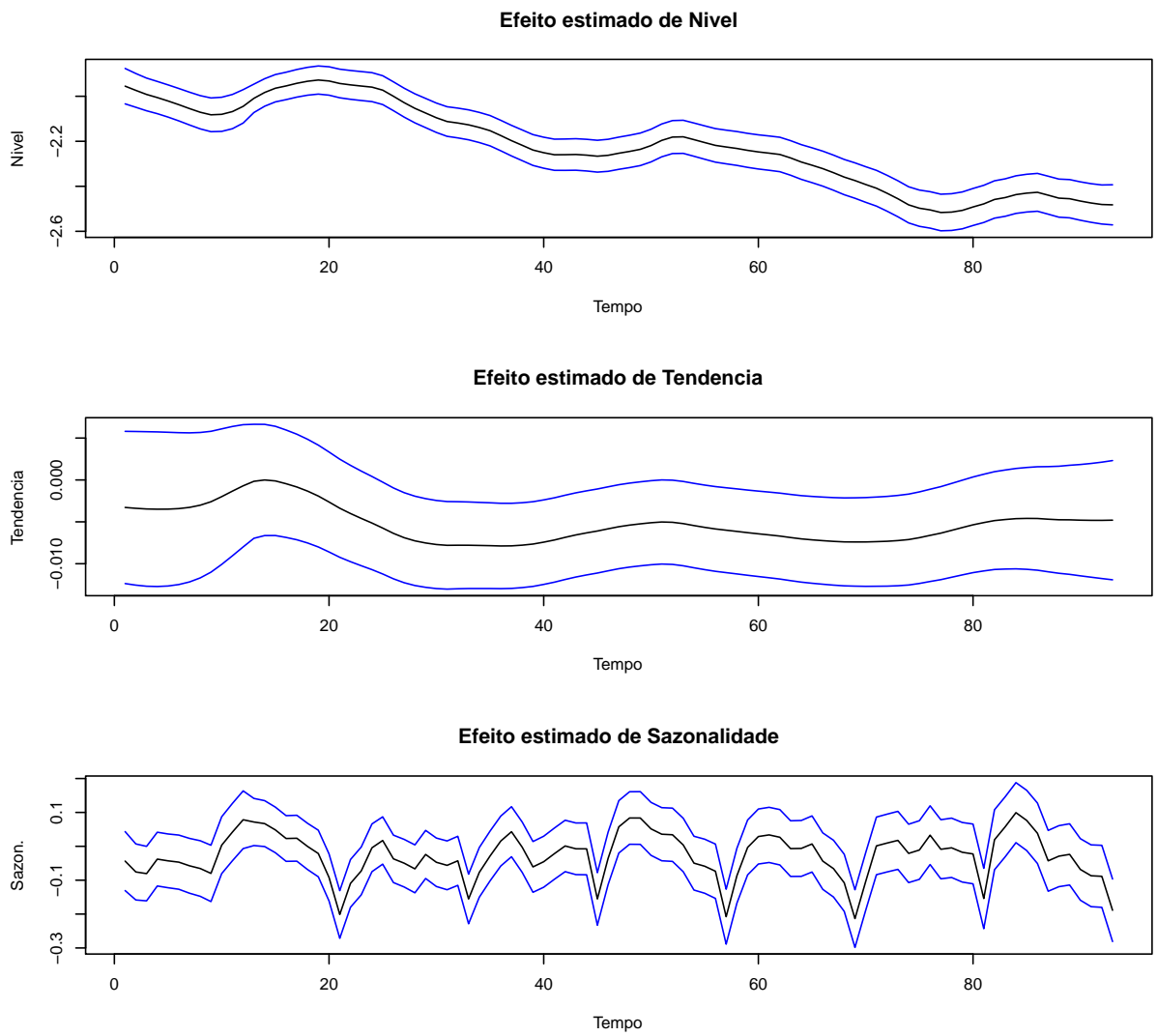


Figura 6.15: Gráfico dos efeitos de nível, tendência e sazonalidade estimados para a taxa de desemprego mensal brasileira baseada nos dados da PME do período entre março de 2002 a julho de 2009. Fonte: PME/IBGE.

6.7.2 Proporção de argila em diferentes profundidades de água em um lago ártico

Coakley e Rust (1968) apresentam a composição de 39 amostras de sedimentos em termos de proporções de areia, argila e lodo. As amostras foram retiradas a diferentes níveis de profundidade. Este conjunto de dados tem sido utilizado especificamente para análise de dados composicionais (Aitchison, 1982; Aitchison, 2003). Em Aitchison (2003) pode ser encontrado alguns comentários sobre a relação de dependência entre a composição dos sedimentos e o nível de profundidade na água. Neste trabalho o mesmo conjunto de dados foi analisado modelando a proporção de argila em diferentes (crescentes) níveis de profundidade de água.

A princípio, não há razões para incorporar outro efeito na especificação do modelo além dos componentes de tendência. Portanto, foi utilizado um Modelo Dinâmico Beta polinomial de segunda ordem para o efeito de tendência. Desta forma, a especificação do modelo é dada por:

$$\{F = E_2, G = J_2(1), W_t = W_{t,\beta}\}$$

Para este conjunto de dados foi utilizado $\delta_T = (0.85, 0.99)$ e $\hat{\phi} = 34$ (estimado conforme a metodologia descrita na Seção 5.2.1). Pode-se observar, na Figura 6.16, que a proporção de argila cresce continuamente com o aumento da profundidade da água, corroborando com o que foi discutido em Aitchison (2003).

Na Figura 6.17 é, primeiramente, apresentada a estimativa suavizada do efeito de nível, onde é observado um crescimento contínuo. Em seguida, é apresentada estimativa suavizada da componente que mede o grau de variação do nível, onde é observado um comportamento estável ao longo dos diferentes níveis de profundidade. Percebe-se que o crescimento da proporção de argila ao longo dos níveis de profundidade ocorre principalmente devido ao comportamento crescente da componente do efeito de nível.

Valores observados e estimados de proporcao de argila

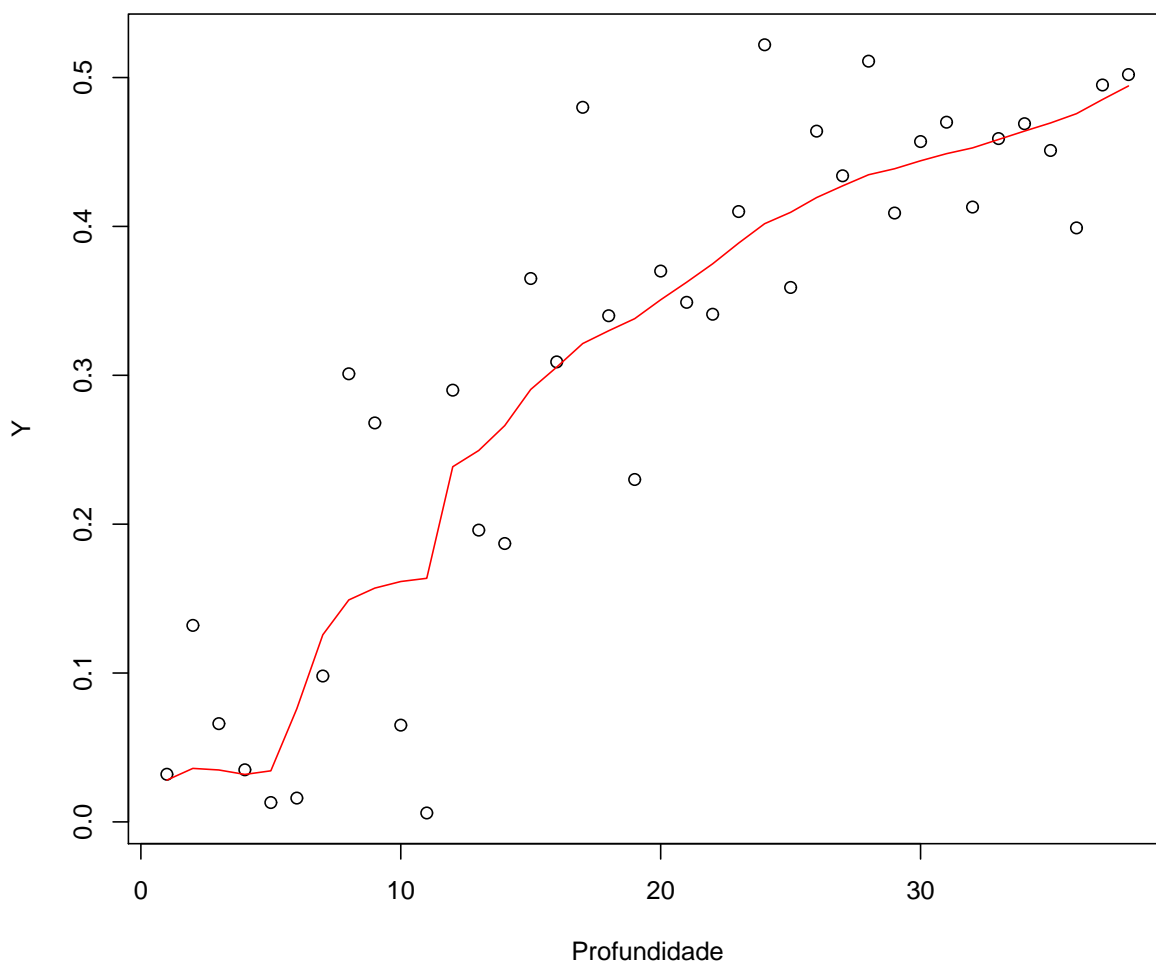


Figura 6.16: Proporção de argila em diferentes níveis de profundidade de água em um lago artigo estimado via Modelo Dinâmico Beta.

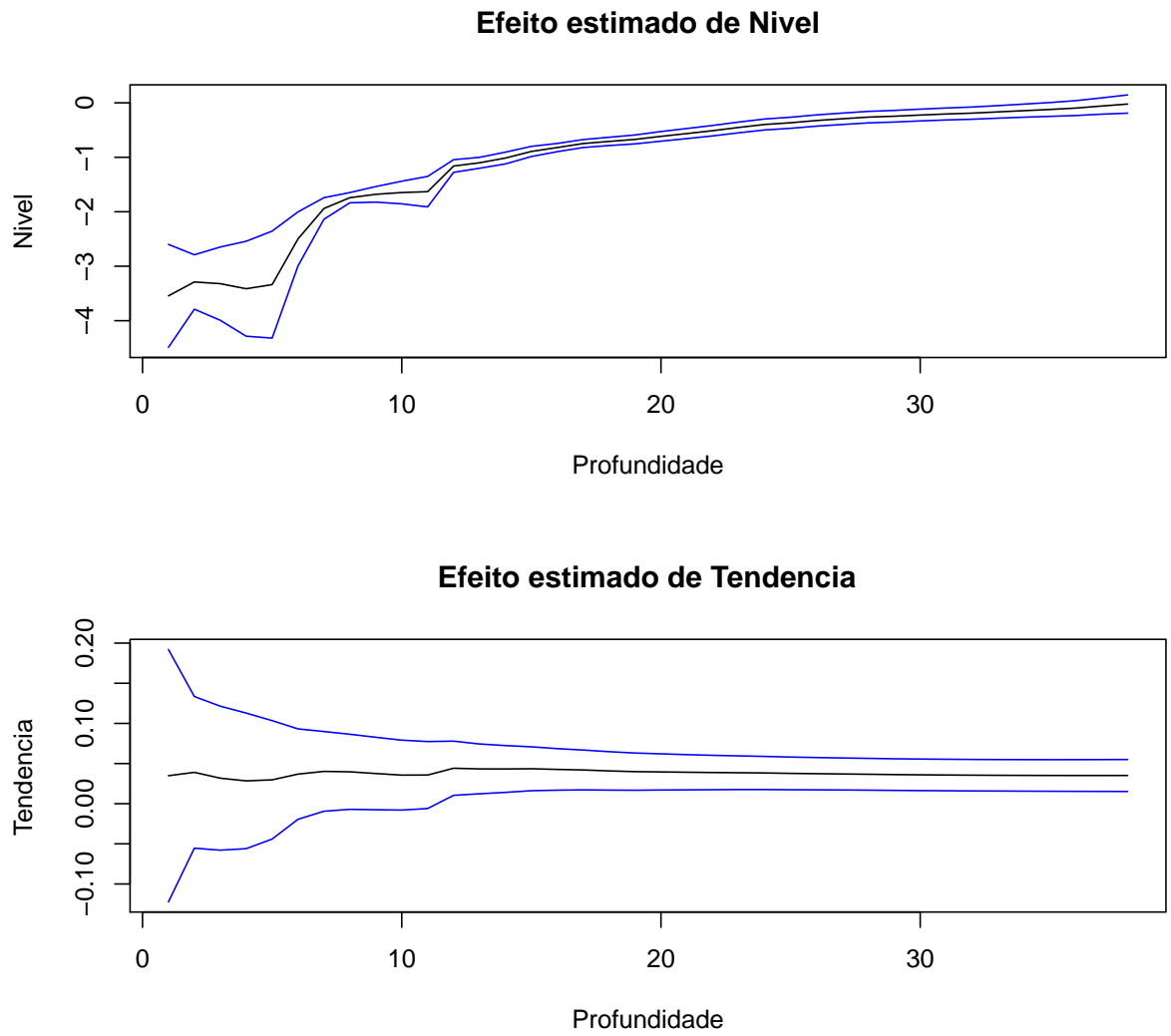


Figura 6.17: Gráfico dos efeitos de nível e tendência estimados para a proporção de argila em diferentes níveis de profundidade de água em um lago ártico.

6.7.3 Proporção de machos em uma população de gambás

Um Modelo Dinâmico Beta de 1ª ordem foi aplicado a proporção de machos em uma população de gambás brasileiros (*Gracilinamus microtarsus*) em um estudo ecológico. Os machos da espécie *Gracilinamus microtarsus* apresentam uma característica peculiar de sobrevivência denominada *semelparidade parcial*. De acordo com Boonstra (2005), a sobrevivência pós-reprodutiva define a posição de dada espécie em um contínuo da *semelparidade* (uma condição em que machos morrem após uma temporada reprodutiva e as fêmeas sobrevivem por mais um ano), para *iteroparidade*, onde os machos possuem múltiplas oportunidades de acasalamento ao longo de suas vidas adultas. A Semelparidade parcial tem sido demonstrada para *G. microtarsus* em Martins et al. (2006a,b) e em da-Silva et al. (2008).

Usando o método de amostragem por captura-recaptura (vide Williams et al., 2002), Martins et al. (2006a,b) monitoraram indivíduos desta espécie em uma região de cerrado de agosto de 2000 a fevereiro 2003, com o período de janeiro de 2001 a fevereiro de 2002, representando a Coorte 2000 (14 amostras) e o período entre dezembro de 2001 a fevereiro de 2003, representando a Coorte 2001 (15 amostras). da-Silva et al. (2008) mostrou que ambas as taxas de sobrevivência e recaptura da espécie em estudo podem ser expressas por uma função de covariáveis (gênero e tempo). Na medida em que as taxas de sobrevivência dos machos decresce, o mesmo ocorre com as taxas de captura.

Neste estudo foram analisadas as proporções de machos capturados nas duas coortes, um total de 31 meses. A série de tempo para as proporções de gambás machos é relativamente pequena e apresenta muitas oscilações. Foi feito, inicialmente, um ajuste simples especificado por um Modelo Dinâmico Beta polinomial de ordem 1, onde: $\{F_t = 1, G_t = 1, W_t = \hat{W}\}$. Seguindo a metodologia descrita na seção 5.2.1, foram estimados $\hat{W} = 0.25$, $\hat{\phi} = 3$. No ajuste considerando W_t estimado via fator de desconto, foi utilizado $\delta_T = 0.6$. Como pode ser observado na Figura 6.18, devido a escassêz de dados e a falta de uma especificação mais elaborada do modelo, o ajuste não foi muito satisfatório, com certa tendência a subestimar a proporção de machos. Entretanto, o modelo foi capaz de capturar

corretamente os ciclos de reprodução descritos na Coorte 2000 e Coorte 2001 e suas implicações na capacidade de capturar a população de gambás machos.

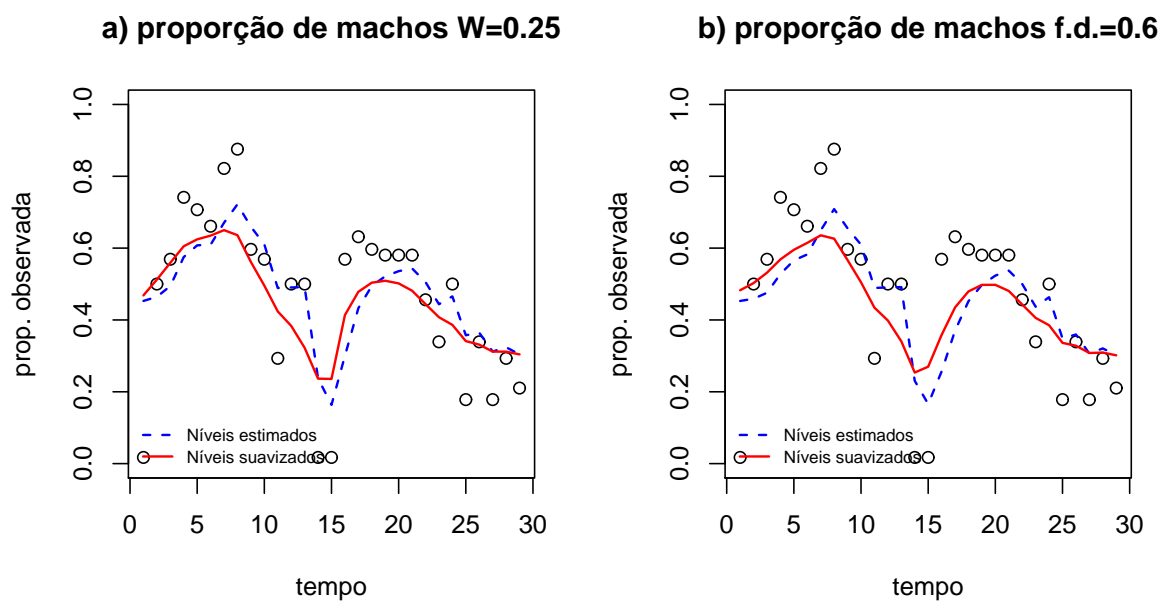


Figura 6.18: Proporção de gambás (*Gracilnamus microtarsus*) machos capturados nas amostras observadas. Em a) a série foi ajustada considerando $\hat{W} = 0.25$ e $\hat{\phi} = 3$. Em b) a série foi ajustada considerando o fator de desconto $\delta = 0.6$ e $\hat{\phi} = 3$.

Capítulo 7

Conclusões e Trabalhos Futuros

7.1 Conclusões

Os exercícios de simulação com séries de tempo *Beta* e as aplicações apresentadas no capítulo anterior mostram que os *Modelos Dinâmicos Beta* podem ser muito úteis para a análise de comportamento e previsão de séries de tempo de dados agregados. Apesar do procedimento descrito em West e Harrison (1997), usando o *Modelo Linear Dinâmico Normal Matriz-Variado* com transformação razão-log, funcionar bem para a modelagem dados composicionais, este ainda apresenta algumas limitações que podem ser contornadas com o Modelo Dinâmico Beta. A principal delas é a suposição de normalidade na distribuição dos erros da equação de evolução, enquanto o Modelo Dinâmico Beta permite que esta hipótese seja relaxada. Outra importante vantagem é devido ao fato de com o Modelo Dinâmico Beta trabalharmos diretamente com a escala de proporções, enquanto no Modelo Linear Dinâmico Normal Matriz-Variado faz-se necessário o uso de transformações lineares, o que limita a interpretação das previsões e seus intervalos de credibilidade.

As comparações entre os métodos de estimação apresentadas no capítulo anterior mostram que é possível chegar a bons ajustes de série de dados agregados por meio de cálculos relativamente simples e sem a necessidade de procedimentos que exigem intenso esforço computacional.

Como trabalhos futuros objetiva-se, modelar o parâmetro de precisão ϕ conjuntamente com os níveis μ_t . de modo que seja descrito uma estrutura de correlação entre estes parâmetros. A partir da metodologia desenvolvida para dados com *distribuição Beta*, é possível chegar a uma metodologia capaz de incorporar maiores dimensões, com modelos estruturados com base na *distribuição Dirichlet*. Estas idéias são inicialmente introduzidas na próxima seção, mas ainda necessitam ser aperfeiçoadas e implementadas computacionalmente.

7.2 Sugestões de Trabalhos Futuros

7.2.1 Ajuste com ϕ integrado

Com o intuito de discutir novas formas de estimar o parâmetro de precisão ϕ . Levando em consideração as dificuldades apresentadas na Seção 5.2, deseja-se implementar uma maneira em que a distribuição de ϕ seja levada em consideração no procedimento inferencial, estabelecendo a relação conjunta entre (μ_t, ϕ) .

Seja $p(\phi|D_t)$ a distribuição a posteriori de ϕ . Portanto, a distribuição conjunta é dada por:

$$p(\mu_t, \phi|D_t) = p(\mu_t|\phi, D_t)p(\phi|D_t), \quad (7.1)$$

e a distribuição a posteriori de μ_t é dada por:

$$p(\mu_t|D_t) = \int p(\mu_t|\phi, D_t)p(\phi|D_t)d\phi, \quad (7.2)$$

com

$$\begin{aligned} p(\mu_t|D_t, \phi) &= \frac{P(Y_t|D_{t-1}, \mu_t, \phi)P(\mu_t|D_{t-1}, \phi)}{\int_0^1 P(Y_t|D_{t-1}, \mu_t, \phi)P(\mu_t|D_{t-1}, \phi)d\mu_t} \\ &= \frac{P(Y_t|D_{t-1}, \mu_t, \phi)P(\mu_t|D_{t-1}, \phi)}{P(Y_t|D_{t-1}, \phi)}. \end{aligned}$$

Enquanto a distribuição a posteriori de ϕ é dada por:

$$P(\phi|D_t) = \frac{P(Y_t|D_{t-1}, \phi)P(\phi|D_{t-1})}{\int P(Y_t|D_{t-1}, \phi)P(\phi|D_{t-1})d\phi}. \quad (7.3)$$

A distribuição a priori conjunta de (Y_t, ϕ) é dada por:

$$P(Y_t, \phi|D_{t-1}) = P(Y_t|D_{t-1}, \phi)P(\phi|D_{t-1}) \quad (7.4)$$

e a marginal:

$$P(Y_t|D_{t-1}) = \int P(Y_t, \phi|D_{t-1})d\phi. \quad (7.5)$$

Em termos do processo iterativo, os mesmos passos são seguidos para obter os momentos de $(\mu_t|D_t)$ que foram descritos na Seção 5.1.

O parâmetro ϕ pode ser interpretado como um “tamanho de amostra”. Portanto, para iniciar o processo iterativo, é atribuído uma priori a ϕ , no tempo $t = 0$, como uma distribuição uniforme discreta, $(\phi|D_0) \sim U(1, M)$, com M representando um valor alto, como 200, por exemplo. O estimador bayesiano de ϕ , considerando a perda quadrática, é dado por $E(\phi|D_T)$.

Em termos práticos, pode-se trabalhar com valores fixos de $\phi = \phi_j$, onde $\phi_j = 1, \dots, M$. Desta forma, mesmo que $P(Y_t|D_{t-1}, \phi)$ não possua forma fechada, a sua distribuição pode ser aproximada por um procedimento de integração varrendo $\phi_j = 1, \dots, M$. Ou seja,

$$P(Y_t|D_{t-1}, \phi_j) = \int_0^1 P(Y_t|D_{t-1}, \mu_t, \phi_j)P(\mu_t|D_{t-1}, \phi_j)d\mu_t$$

onde $P(Y_t|D_{t-1}, \mu_t, \phi_j) \sim \text{Beta}(\phi_j\mu_t, \phi_j(1-\mu_t))$ e $P(\mu_t|D_{t-1}, \phi_j) \sim \text{Beta}(r_t, s_t)$. Conseqüentemente,

$$P(\phi|D_t) = \frac{P(Y_t|D_{t-1}, \phi)P(\phi|D_{t-1})}{\int P(Y_t|D_{t-1}, \phi)P(\phi|D_{t-1})d\phi}.$$

Logo,

$$\begin{aligned} P(Y_t|D_{t-1}) &= \int P(Y_t|D_{t-1}, \phi)P(\phi|D_{t-1})d\phi \\ &\approx \sum_{j=1}^M P(Y_t|D_{t-1}, \phi_j)P(\phi_j|D_{t-1}) \\ &= \frac{1}{M} \sum_{h=1}^M P(Y_t|D_{t-1}, h). \end{aligned}$$

Portanto,

$$P(\phi|D_t) \approx \frac{P(Y_t|D_{t-1}, \phi)P(\phi|D_{t-1})}{\sum_{h=1}^M P(Y_t|D_{t-1}, h)}. \quad (7.6)$$

Com isso, a distribuição a posteriori marginal de μ_t é obtida por:

$$P(\mu_t|D_t) \approx \sum_{j=1}^M P(\mu_t|D_t, \phi_j)P(\phi_j|D_t),$$

onde deseja-se obter os valores dos momentos $E(\mu_t|D_t)$ e $V(\mu_t|D_t)$.

Sintetizando o procedimento:

- Construir uma matriz $T \times M$ para os valores de $P(Y_t|D_{t-1}, \phi_j)$, com $j = 1, \dots, M$ e $t = 1, \dots, T$.
- Calcular a posteriori de ϕ :

$$P(\phi|D_t) = \frac{P(Y_t|D_{t-1}, \phi)P(\phi|D_{t-1})}{\sum_{h=1}^M P(Y_t|D_{t-1}, h)},$$

com

$$p(\phi = j|D_t) = \frac{P(Y_t|D_{t-1}, j)P(\phi|D_{t-1})}{\sum_{h=1}^M P(Y_t|D_{t-1}, h)}.$$

- Obter a matriz $T \times M$ para os valores de $P(\phi = j|D_t)$.
- E então,

$$P(\mu_t|D_t) \approx \sum_{j=1}^M P(\mu_t|D_t, \phi_j)P(\phi_j|D_t).$$

- Por fim, é utilizado um procedimento de integração para estimar $E(\mu_t|D_t)$ e $V(\mu_t|D_t)$:

$$\begin{aligned}\tilde{\mu}_t &= E(\mu_t|D_t) = \int \mu_t P(\mu_t|D_t) d\mu_t; \\ \tilde{V}_t &= V(\mu_t|D_t) \approx \int (\mu_t - \tilde{\mu}_t)^2 P(\mu_t|D_t) d\mu_t.\end{aligned}$$

7.2.2 Modelo Dinâmico Beta com μ_t e ϕ Desconhecidos

Nesta seção, a idéia é de incorporar uma distribuição de probabilidade a ϕ , de modo que esta esteja de alguma forma correlacionada com os níveis μ_t . Para isto, o sistema de equações é estabelecido da seguinte forma:

- *Equação das Observações:*

$$\begin{aligned}P(Y_t|\phi, \mu_t) &= P(Y_t|\phi, \mu_t, D_{t-1}) \\ &= \frac{\Gamma(\phi)}{\Gamma(\mu_t\phi)\Gamma(\phi(1-\mu_t))} y_t^{\mu_t\phi-1} (1-y_t)^{\phi(1-\mu_t)-1}\end{aligned}$$

- *Função de Ligação:*

$$g(\mu_t) = \log\left(\frac{\mu_t}{1-\mu_t}\right) = \lambda_t = F_t^T \beta_t.$$

Logo,

$$\mu_t = \frac{e^{\lambda_t}}{1 + e^{\lambda_t}}$$

e $\lambda_t \in \mathbb{R}$.

– *Equações de Evolução:*

$$\beta_t = G_t \beta_{t-1} + \omega_t; \quad \omega_t \sim (0, W_t)$$

$$\phi \sim (z, E)$$

onde ω_t independe de $\phi \forall t$ e os ω_t 's são independentes entre si.

– *Informação Inicial:*

$$(\phi, \beta_0 | D_0) \sim \left[\begin{array}{c} \left(\begin{array}{c} z_0 \\ m_0 \end{array} \right), \left[\begin{array}{cc} E_0 & \Sigma_{\phi\beta_0} \\ \Sigma_{\phi\beta_0} & C_0 \end{array} \right] \end{array} \right]$$

Tem-se:

G_t é matriz $j \times j$ e F_t é vetor $k \times 1$

Passo 1:

– *Priori em t* - Priori conjunta de $(\phi, \beta_t, \lambda_t | D_{t-1})$.

Tome

$$\left(\begin{array}{c} \phi \\ \beta_t \\ \lambda_t \end{array} \middle| D_{t-1} \right) \sim \left[\begin{array}{c} \left(\begin{array}{c} z_t \\ a_t \\ f_t \end{array} \right), \left(\begin{array}{ccc} \Sigma_{\phi\phi} & \Sigma_{\phi\beta_t} & \Sigma_{\phi\lambda_t} \\ \Sigma_{\beta_t\phi} & \Sigma_{\beta_t\beta_t} & \Sigma_{\beta_t\lambda_t} \\ \Sigma_{\lambda_t\phi} & \Sigma_{\lambda_t\beta_t} & \Sigma_{\lambda_t\lambda_t} \end{array} \right) \end{array} \right]$$

– *Médias:*

$$z_t = E(\phi | D_{t-1})$$

$$a_t = E(\beta_t | D_{t-1}) = E[G_t \beta_{t-1} + \omega_t | D_{t-1}] = G_t m_{t-1}$$

$$f_t = E(\lambda_t | D_{t-1}) = E[F_t^T \beta_t | D_{t-1}] = F_t^T a_t$$

– *Variâncias e Covariâncias:*

$$\Sigma_{\phi\phi} = \text{Var}(\phi | D_{t-1}) = E_t$$

$$\Sigma_{\beta_t\beta_t} = \text{Var}(G_t \beta_{t-1} + \omega_t | D_{t-1}) = G_t C_{t-1} G_t^T + W_t$$

$$\Sigma_{\lambda_t\lambda_t} = \text{Var}(F_t^T \beta_t | D_{t-1}) = F_t^T \Sigma_{\beta_t\beta_t} F_t = q_t$$

$$\Sigma_{\phi\lambda_t} = \text{Cov}(\phi, F_t^T \beta_t | D_{t-1}) = \Sigma_{\phi\beta_t} F_t$$

$$\Sigma_{\lambda_t\phi_t} = \text{Cov}(F_t^T \beta_t, \phi | D_{t-1}) = F_t^T \Sigma_{\beta_t\phi}$$

$$\begin{aligned}
\Sigma_{\beta_t \phi_t} &= Cov(G_t \beta_{t-1} + \omega_t; \phi | D_{t-1}) \\
&= Cov(G_t \beta_{t-1}; \phi | D_{t-1}) + Cov(\omega_t; \phi | D_{t-1}) \\
&= G_t \Sigma_{\beta_{t-1} \phi} \\
\Sigma_{\phi \beta_t} &= \Sigma_{\beta_t \phi}^T
\end{aligned}$$

Passo 2: Atualização de λ_t - Obtenção de $E(\lambda_t | D_t)$ e $V(\lambda_t | D_t)$

1. Especificação da Priori de $(\mu_t | D_{t-1})$
2. Relacionamento entre $E(\mu_t | D_{t-1})$ e $V(\mu_t | D_{t-1})$ com $E(\lambda_t | D_{t-1})$ e $V(\lambda_t | D_{t-1})$
3. Especificação dos hiperparâmetros de $(\mu_t | D_{t-1})$
4. Atualização de μ_t : obtenção dos momentos de $(\mu_t | D_t)$
5. Relacionamento dos momentos de $(\mu_t | D_t)$ com os momentos de $(\lambda_t | D_t)$
6. Obtenção dos valores atualizados $f_t^* = E(\lambda_t | D_t)$ e $q_t^* = V(\lambda_t | D_t)$

Passo 3: Estrutura Condicional de $(\phi, \beta_t | \lambda_t, D_{t-1})$:

$$\begin{aligned}
P(\phi, \beta_t, \lambda_t | D_t) &\propto P(\phi, \beta_t, \lambda_t, y_t, D_{t-1}) \\
&\propto P(y_t | \phi, \lambda_t, D_{t-1}) P(\phi, \beta_t, \lambda_t | D_{t-1}) \\
&\propto P(\phi, \beta_t | \lambda_t, D_{t-1}) \underbrace{P(Y_t | \phi, \lambda_t, D_{t-1}) P(\lambda_t | D_{t-1})}_{P(\lambda_t | D_t)} \\
&\propto P(\phi, \beta_t | \lambda_t, D_{t-1}) P(\lambda_t | D_t)
\end{aligned}$$

$$P(\phi, \beta_t | D_t) = \int P(\phi, \beta_t | \lambda_t, D_{t-1}) P(\lambda_t | D_t) d\lambda_t$$

Passo 4: Obtenção de $E(\phi, \beta_t | \lambda_t, D_{t-1})$ e $V(\phi, \beta_t | \lambda_t, D_{t-1})$

Partindo das propriedades da distribuição Normal Multivariada para encontrar os valores esperados, deve-se considerar que:

$$X = \begin{pmatrix} X_1 \\ \dots \\ X_2 \end{pmatrix} = \begin{pmatrix} \phi_t \\ \beta_t \\ \dots \\ \lambda_t \end{pmatrix} \Bigg| D_{t-1}, \xi = \begin{pmatrix} \xi_1 \\ \dots \\ \xi_2 \end{pmatrix} = \begin{pmatrix} z_t \\ a_t \\ \dots \\ f_t \end{pmatrix} \quad \text{e} \quad H = \begin{bmatrix} H_{11} & \vdots & H_{12} \\ \dots & \dots & \dots \\ H_{21} & \vdots & H_{22} \end{bmatrix}$$

onde

$$H_{11} = \begin{bmatrix} \Sigma_{\phi_t \phi_t} & \Sigma_{\phi_t \beta_t} \\ \Sigma_{\beta_t \phi} & \Sigma_{\beta_t \beta_t} \end{bmatrix}; \quad H_{12} = \begin{bmatrix} \Sigma_{\phi, \lambda_t} \\ \Sigma_{\beta_t \lambda_t} \end{bmatrix}$$

$$H_{21} = \begin{bmatrix} \Sigma_{\lambda_t \phi} & \Sigma_{\lambda_t \beta_t} \end{bmatrix} \quad \text{e} \quad H_{22} = \Sigma_{\lambda_t \lambda_t} = q_t.$$

Com base na teoria Normal (vide apêndice E),

$$(X_1|X_2) \sim N(\xi_1(X_2), H_{11}(X_2))$$

$$\text{onde } \xi_1(X_2) = \xi_1 + H_{12}H_{22}^{-1}(X_2 - \xi_2)$$

$$\text{e } H_{11}(X_2) = H_{11} - H_{12}H_{22}^{-1}H_{21}.$$

Portanto,

$$E(\phi, \beta_t | \lambda_t, D_{t-1}) = \begin{pmatrix} z_t \\ a_t \end{pmatrix} + \frac{1}{q_t} H_{12}(\lambda_t - f_t)$$

$$V(\phi, \beta_t | \lambda_t, D_{t-1}) = H_{11} - \frac{H_{12}H_{21}}{q_t}$$

Passo 5: Atualização de (ϕ, β_t) - Obtenção de $E[\phi, \beta_t | D_t]$ e $V[\phi, \beta_t | D_t]$

Na obtenção de $E[\phi, \beta_t | D_t]$ e $V[\phi, \beta_t | D_t]$ considera-se que (*Linear Bayes*, vide apêndice):

$$\begin{aligned} E[\phi, \beta_t | D_t] &= E[E(\phi, \beta_t | \lambda_t, D_{t-1}) | D_t] \\ &= \begin{pmatrix} z_t \\ a_t \end{pmatrix} + \frac{1}{q_t} H_{12}(E(\lambda_t | D_t) - f_t) \\ &= \begin{pmatrix} z_t \\ a_t \end{pmatrix} + \frac{1}{q_t} H_{12}(f_t^* - f_t) \end{aligned}$$

e

$$\begin{aligned} V[\phi, \beta_t | D_t] &= V[E(\phi, \beta_t | \lambda_t, D_{t-1}) | D_t] + E[V(\phi, \beta_t | \lambda_t, D_{t-1}) | D_t] \\ &= \frac{1}{q_t^2} H_{12}H_{21}V(\lambda_t | D_t) + H_{11} - \frac{H_{12}H_{21}}{q_t} \\ &= H_{11} - \frac{1}{q_t} H_{12}H_{21} \left[1 - \frac{q_t^*}{q_t} \right]. \end{aligned}$$

No entanto, foi observado que ao utilizar esta metodologia, o valor encontrado para $E(\phi | D_t)$ está fortemente ligado ao valor inicial estabelecido para a priori:

$$(\phi, \beta_0 | D_0) \sim \left[\begin{array}{c} \left(\begin{array}{c} z_0 \\ m_0 \end{array} \right), \left[\begin{array}{cc} \Sigma_{\phi\phi} & \Sigma_{\phi\beta_0} \\ \Sigma_{\beta_0\phi} & \Sigma_{\beta_0\beta_0} \end{array} \right] \end{array} \right]$$

Sendo assim, faz-se necessário ter um bom conhecimento inicial em relação a ϕ para incorporá-lo em sua priori.

7.2.3 Modelo Dinâmico Dirichlet

Generalizando o caso do modelo *Beta* para dimensões $q > 2$, tem-se então um modelo com distribuição *Dirichlet*. Supondo que Y tenha dimensão $q = 3$:

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \sim Dirichlet(a, b, c),$$

A função densidade da distribuição Dirichlet é dada por:

$$f_Y(y; a, b, c) = \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} y_1^{a-1} y_2^{b-1} (1-y_1-y_2)^{c-1}; \quad 0 < y_1, y_2 < 1, \quad a, b, c > 0.$$

Considerando-se uma parametrização alternativa, faça:

$$\mu_1 = \frac{a}{\phi}, \quad \mu_2 = \frac{b}{\phi}, \quad \mu_3 = 1 - \mu_1 - \mu_2 = \frac{c}{\phi}, \quad \text{e} \quad \phi = a + b + c,$$

reparametrizando em termos de μ_i ,

$$f_Y(y; \mu_1, \mu_2, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu_1\phi)\Gamma(\mu_2\phi)\Gamma((1-\mu_1-\mu_2)\phi)} y_1^{\mu_1\phi-1} y_2^{\mu_2\phi-1} \times (1-y_1-y_2)^{(1-\mu_1-\mu_2)\phi-1},$$

onde $0 < \mu_1, \mu_2 < 1, \quad \phi > 0$.

Desta forma,

$$E(Y_{it}) = \mu_{it}, \quad V(Y_{it}) = \frac{(\mu_{it})(1-\mu_{it})}{1+\phi} \quad \text{e} \quad Cov(Y_{it}, Y_{jt}) = -\frac{(\mu_{it})(\mu_{jt})}{1+\phi}.$$

Contextualizando no modelo dinâmico

– Equação das observações

$$(Y_t | \phi, \mu_{1t}, \mu_{2t}) \sim Dirichlet(\mu_{1t}\phi, \mu_{2t}\phi, (1-\mu_{1t}-\mu_{2t})\phi)$$

com ϕ conhecido e

$$g(\mu_t) = F_t' \beta_t = \lambda_t,$$

onde F_t é matrix $n \times 2$ de constantes conhecidas, β_t é vetor coluna de n elementos,

$$\lambda_t = \begin{pmatrix} \lambda_{1t} \\ \lambda_{2t} \end{pmatrix} \text{ e } \mu_t = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix}.$$

Considere a função de ligação logística,

$$\lambda_t = g(\mu_t) = \begin{pmatrix} \lambda_{1t} \\ \lambda_{2t} \end{pmatrix} = \begin{pmatrix} \log \left(\frac{\mu_{1t}}{1 - \mu_{1t} - \mu_{2t}} \right) \\ \log \left(\frac{\mu_{2t}}{1 - \mu_{1t} - \mu_{2t}} \right) \end{pmatrix}. \quad (7.7)$$

Observe que

$$\exp(\lambda_{jt}) = \frac{\mu_{jt}}{1 - \mu_{1t} - \mu_{2t}}, \quad j = 1, 2. \quad (7.8)$$

Da expressão acima, tem-se

$$\frac{\exp(\lambda_{1t})}{\exp(\lambda_{2t})} = \frac{\mu_{1t}}{\mu_{2t}}, \quad (7.9)$$

e, também, que

$$1 - \mu_{1t} - \mu_{2t} = \frac{\mu_{1t}}{\exp(\lambda_{1t})} = \frac{\mu_{2t}}{\exp(\lambda_{2t})}. \quad (7.10)$$

Substituindo os termos, obtém-se:

$$1 - \mu_{1t} = \mu_{2t} + \frac{\mu_{1t}}{\exp(\lambda_{1t})},$$

e, considerando a razão μ_{1t}/μ_{2t} , segue que:

$$\mu_{1t} = \frac{\exp(\lambda_{1t})}{1 + \exp(\lambda_{1t}) + \exp(\lambda_{2t})}. \quad (7.11)$$

Desta forma,

$$\mu_t = \begin{pmatrix} \frac{\exp(\lambda_{1t})}{1 + \exp(\lambda_{1t}) + \exp(\lambda_{2t})} \\ \frac{\exp(\lambda_{2t})}{1 + \exp(\lambda_{1t}) + \exp(\lambda_{2t})} \end{pmatrix}. \quad (7.12)$$

– **Equação das evoluções:**

$$\beta_t = G_t \beta_{t-1} + w_t; \quad w_t \sim (0, W_t),$$

onde G_t é matrix $n \times n$ de constantes conhecidas.

– **Informação Inicial:**

$$(\beta_0 | D_0, \phi) \sim (m_0, C_0).$$

Aprofundando, inicialmente, no caso de ϕ conhecido.

– **Passo 1:** Priori de $\lambda_t = F_t' \beta_t$.

Tome

$$\begin{pmatrix} \lambda_t \\ \beta_t \end{pmatrix} | D_{t-1}, \phi \sim \left[\begin{pmatrix} f_t \\ a_t \end{pmatrix}, \begin{pmatrix} Q_t & F_t' R_t \\ R_t F_t & R_t \end{pmatrix} \right]$$

onde

$$f_t = F_t' a_t; \quad Q_t = F_t' R_t F_t; \quad a_t = G_t m_{t-1}; \quad R_t = G_t C_{t-1} G_t' + W_t$$

Então,

$$(\lambda_t | D_{t-1}, \phi) \sim (f_t, Q_t) \quad e \quad (\beta_t | D_{t-1}, \phi) \sim (a_t, R_t),$$

e considera-se que

$$(\beta_{t-1} | D_{t-1}, \phi) \sim (m_{t-1}, C_{t-1}).$$

– **Passo 2:** Previsão de 1 passo a frente $P(Y_t | D_{t-1}, \phi)$.

a) Especificação da priori de μ_t , $P(\mu_t | D_{t-1}, \phi)$:

Como $\mu_{jt} \in (0, 1)$, $j = 1, 2$, e $\mu_{3t} = 1 - \mu_{1t} - \mu_{2t}$, considere uma priori

Dirichlet(s_t, r_t, u_t) onde s_t, r_t e u_t são especificados de modo a satisfazer as equações

$$E[g(\mu_t)|D_{t-1}, \phi] = f_t \quad e \quad V[g(\mu_t)|D_{t-1}, \phi] = Q_t,$$

devidas a $(\lambda_t|D_{t-1}, \phi) \sim (f_t, Q_t)$.

b) Relacionando $E(\mu_t|D_{t-1}, \phi)$ e $V(\mu_t|D_{t-1}, \phi)$ com $E(\lambda_t|D_{t-1}, \phi)$ e $V(\lambda_t|D_{t-1}, \phi)$:

$$E(\mu_{it}|D_{t-1}, \phi) = E \left[\frac{\exp(\lambda_{it})}{1 + \exp(\lambda_{1t}) + \exp(\lambda_{2t})} \middle| D_{t-1}, \phi \right].$$

No cálculo acima, considere a expansão de Taylor de 1ª ordem ao redor de x_0 para a função:

$$f(x) = \begin{pmatrix} \frac{e^{x_1}}{1+e^{x_1}+e^{x_2}} \\ \frac{e^{x_2}}{1+e^{x_1}+e^{x_2}} \end{pmatrix}.$$

Logo,

$$f(x) \approx f(x_0) + \nabla f(x_0)(x - x_0)$$

$$\approx \begin{pmatrix} \frac{e^{x_{01}}}{1+e^{x_{01}}+e^{x_{02}}} \\ \frac{e^{x_{02}}}{1+e^{x_{01}}+e^{x_{02}}} \end{pmatrix} + \begin{pmatrix} \frac{e^{x_{01}}(1+e^{x_{02}})}{(1+e^{x_{01}}+e^{x_{02}})^2} & -\frac{e^{x_{01}}e^{x_{02}}}{(1+e^{x_{01}}+e^{x_{02}})^2} \\ -\frac{e^{x_{01}}e^{x_{02}}}{(1+e^{x_{01}}+e^{x_{02}})^2} & \frac{e^{x_{02}}(1+e^{x_{01}})}{(1+e^{x_{01}}+e^{x_{02}})^2} \end{pmatrix} \begin{pmatrix} x_1 - x_{01} \\ x_2 - x_{02} \end{pmatrix}.$$

Seja $x_0 = f_t$, então,

$$\begin{pmatrix} \frac{e^{x_1}}{1+e^{x_1}+e^{x_2}} \\ \frac{e^{x_2}}{1+e^{x_1}+e^{x_2}} \end{pmatrix} \approx \begin{pmatrix} \frac{e^{f_{1t}}}{1+e^{f_{1t}}+e^{f_{2t}}} \\ \frac{e^{f_{2t}}}{1+e^{f_{1t}}+e^{f_{2t}}} \end{pmatrix} + \begin{pmatrix} \frac{e^{f_{1t}}(1+e^{f_{2t}})}{(1+e^{f_{1t}}+e^{f_{2t}})^2} & -\frac{e^{f_{1t}}e^{f_{2t}}}{(1+e^{f_{1t}}+e^{f_{2t}})^2} \\ -\frac{e^{f_{1t}}e^{f_{2t}}}{(1+e^{f_{1t}}+e^{f_{2t}})^2} & \frac{e^{f_{2t}}(1+e^{f_{1t}})}{(1+e^{f_{1t}}+e^{f_{2t}})^2} \end{pmatrix} \begin{pmatrix} x_1 - f_{1t} \\ x_2 - f_{2t} \end{pmatrix}.$$

Logo,

$$\begin{aligned}
E(\mu_t | D_{t-1}, \phi) &= E \left[\left(\begin{array}{c} \frac{\exp(\lambda_{1t})}{1 + \exp(\lambda_{1t}) + \exp(\lambda_{2t})} \\ \frac{\exp(\lambda_{2t})}{1 + \exp(\lambda_{1t}) + \exp(\lambda_{2t})} \end{array} \middle| D_{t-1}, \phi \right) \right] \\
&\approx E \left[\left(\begin{array}{c} \frac{e^{f_{1t}}}{1 + e^{f_{1t}} + e^{f_{2t}}} + (\lambda_{1t} - f_{1t}) \frac{e^{f_{1t}}(1 + e^{f_{2t}})}{(1 + e^{f_{1t}} + e^{f_{2t}})^2} - (\lambda_{2t} - f_{2t}) \frac{e^{f_{1t}}e^{f_{2t}}}{(1 + e^{f_{1t}} + e^{f_{2t}})^2} \\ \frac{e^{f_{2t}}}{1 + e^{f_{1t}} + e^{f_{2t}}} + (\lambda_{2t} - f_{2t}) \frac{e^{f_{2t}}(1 + e^{f_{1t}})}{(1 + e^{f_{1t}} + e^{f_{2t}})^2} - (\lambda_{1t} - f_{1t}) \frac{e^{f_{1t}}e^{f_{2t}}}{(1 + e^{f_{1t}} + e^{f_{2t}})^2} \end{array} \right) \right] \\
&\approx \left(\begin{array}{c} \frac{e^{f_{1t}}}{1 + e^{f_{1t}} + e^{f_{2t}}} \\ \frac{e^{f_{2t}}}{1 + e^{f_{1t}} + e^{f_{2t}}} \end{array} \right) = \left(\begin{array}{c} h_{1t} \\ h_{2t} \end{array} \right).
\end{aligned}$$

Similarmente,

$$V(\mu_t | D_{t-1}, \phi) \approx (\nabla f(f_t))' V[\lambda_t | D_{t-1}, \phi] (\nabla f(f_t)) = V_t = \begin{pmatrix} V_{11t} & V_{12t} \\ V_{21t} & V_{22t} \end{pmatrix},$$

com,

$$\nabla f(f_t) = \begin{pmatrix} \frac{e^{f_{1t}}(1 + e^{f_{2t}})}{(1 + e^{f_{1t}} + e^{f_{2t}})^2} & -\frac{e^{f_{1t}}e^{f_{2t}}}{(1 + e^{f_{1t}} + e^{f_{2t}})^2} \\ -\frac{e^{f_{1t}}e^{f_{2t}}}{(1 + e^{f_{1t}} + e^{f_{2t}})^2} & \frac{e^{f_{2t}}(1 + e^{f_{1t}})}{(1 + e^{f_{1t}} + e^{f_{2t}})^2} \end{pmatrix},$$

c) Obtendo s_t , r_t e u_t :

Como $(\mu_t | D_{t-1}, \phi) \sim \text{Dirichlet}(s_t, r_t, u_t)$,

$$\begin{aligned}
E(\mu_{1t} | D_{t-1}) &= \frac{s_t}{s_t + r_t + u_t}; \\
E(\mu_{2t} | D_{t-1}) &= \frac{r_t}{s_t + r_t + u_t}; \\
V(\mu_{1t} | D_{t-1}) &= \frac{s_t(r_t + u_t)}{(s_t + r_t + u_t)^2(s_t + r_t + u_t + 1)}; \\
V(\mu_{2t} | D_{t-1}) &= \frac{r_t(s_t + u_t)}{(s_t + r_t + u_t)^2(s_t + r_t + u_t + 1)}; \\
Cov(\mu_{1t}, \mu_{2t} | D_{t-1}) &= -\frac{s_t r_t}{(s_t + r_t + u_t)^2(s_t + r_t + u_t + 1)}.
\end{aligned}$$

Sejam,

$$h_{1t} = E(\mu_{1t}|D_{t-1}, \phi), \quad h_{2t} = E(\mu_{2t}|D_{t-1}, \phi) \quad \text{e}$$

$$V_t = \begin{pmatrix} V_{11t} & V_{12t} \\ V_{12t} & V_{22t} \end{pmatrix} = V(\mu_t|D_{t-1}, \phi),$$

e considere o sistema de equações,

$$\begin{cases} \frac{s_t}{s_t+r_t+u_t} = h_{1t}; & \frac{r_t}{s_t+r_t+u_t} = h_{2t} \\ \frac{s_t(r_t+u_t)}{(s_t+r_t+u_t)^2(s_t+r_t+u_t+1)} = V_{11t}; & \frac{r_t(s_t+u_t)}{(s_t+r_t+u_t)^2(s_t+r_t+u_t+1)} = V_{22t} \\ -\frac{s_tr_t}{(s_t+r_t+u_t)^2(s_t+r_t+u_t+1)} = V_{12t} \end{cases}$$

Portanto, os hiperparâmetros da Priori Dirichlet $P(\mu_t | D_{t-1})$ são dados por:

$$s_t = -h_{1t} \left(\frac{h_{1t}h_{2t}}{V_{12t}} + 1 \right); \quad r_t = -h_{2t} \left(\frac{h_{1t}h_{2t}}{V_{12t}} + 1 \right);$$

$$u_t = -(1 - h_{1t} - h_{2t}) \left(\frac{h_{1t}h_{2t}}{V_{12t}} + 1 \right).$$

d) Previsão de 1 passo a frente, $P(Y_t|D_{t-1})$:

$$P(Y_t|D_{t-1}, \phi) = \int P(Y_t|\mu_t, \phi, D_{t-1})P(\mu_t|\phi, D_{t-1})d\mu_t.$$

Assim como no caso do *modelo beta*, esta integral não possui forma fechada. Neste caso, é necessário utilizar algum método de aproximação, como o uso da esperança condicional, onde:

$$\tilde{Y}_t = E[E(Y_t|D_{t-1}, \phi)|D_t]$$

– **Passo 3** - Posteriori de μ_t - $P(\mu_t|D_t, \phi)$:

Pelo Teorema de Bayes, a posteriori de μ_t é:

$$\begin{aligned}
P(\mu_t|D_t, \phi) &\propto P(Y_t|\mu_t, \phi, D_{t-1})P(\mu_t|D_{t-1}, \phi) \\
&\propto \frac{\Gamma(\phi)}{\Gamma(\mu_{1t}\phi)\Gamma(\mu_{2t}\phi)\Gamma((1-\mu_{1t}-\mu_{2t})\phi)} y_{1t}^{\mu_{1t}\phi-1} y_{2t}^{\mu_{2t}\phi-1} \\
&\quad \times (1-y_{1t}-y_{2t})^{(1-\mu_{1t}-\mu_{2t})\phi-1} \\
&\quad \times \frac{\Gamma(s_t+r_t+u_t)}{\Gamma(s_t)\Gamma(r_t)\Gamma(u_t)} \mu_{1t}^{s_t-1} \mu_{2t}^{r_t-1} (1-\mu_{1t}-\mu_{2t})^{u_t-1}.
\end{aligned}$$

Observe que $P(\mu_t|D_t, \phi)$ não tem forma conhecida. Portanto, faz-se necessário utilizar um método computacional na estimação dos momentos de tal distribuição. Uma possibilidade é a obtenção de uma aproximação da moda a posteriori de $(\mu_t|D_t, \phi)$ (estimador de máxima verossimilhança) e da respectiva inversa da matriz de informação de Fisher. Este procedimento segue de forma análoga ao caso do *modelo dinâmico beta* descrito no Capítulo 5, podendo ainda fazer uso de aproximações de Laplace para aprimorar a estimativa dos momentos.

$$\begin{aligned}
L = P(\mu_{t1}|D_t) &\propto \frac{\Gamma(\phi)}{\Gamma(\mu_{1t}\phi)\Gamma(\mu_{2t}\phi)\Gamma((1-\mu_{1t}-\mu_{2t})\phi)} y_{1t}^{\mu_{1t}\phi-1} y_{2t}^{\mu_{2t}\phi-1} \\
&\quad \times (1-y_{1t}-y_{2t})^{(1-\mu_{1t}-\mu_{2t})\phi-1} \\
&\quad \times \frac{\Gamma(s_t+r_t+u_t)}{\Gamma(s_t)\Gamma(r_t)\Gamma(u_t)} \mu_{1t}^{s_t-1} \mu_{2t}^{r_t-1} (1-\mu_{1t}-\mu_{2t})^{u_t-1}
\end{aligned}$$

A log-verossimilhança é dada por:

$$\begin{aligned}
l = \log(L) &\propto -\log(\Gamma(\phi\mu_{1t})) - \log(\Gamma(\phi\mu_{2t})) - \log(\Gamma(\phi(1-\mu_{1t}-\mu_{2t}))) \\
&\quad + (\mu_{1t}\phi - 1)\log(y_{1t}) + (\mu_{2t}\phi - 1)\log(y_{2t}) \\
&\quad + [(1-\mu_{1t}-\mu_{2t})\phi - 1]\log(1-y_{1t}-y_{2t}) \\
&\quad + (s_t - 1)\log \mu_{1t} + (r_t - 1)\log \mu_{2t} \\
&\quad + (u_t - 1)\log(1-\mu_{1t}-\mu_{2t}).
\end{aligned}$$

Sejam $\hat{E}(\mu_{1t}|D_t)$ e $\hat{E}(\mu_{2t}|D_t)$ os pontos em $(\mu_{1t}, \mu_{2t}) \in (0, 1)$ que maximizam l . No entanto, não é possível encontrar analiticamente o ponto de máximo e portanto, assim como no caso do Modelo Dinâmico Beta, é necessário utilizar algum método numérico de maximização.

Seguem abaixo as derivadas de 1ª e 2ª ordem e estimativa da variância de $(\mu_t|D_t)$:

$$\begin{aligned} \frac{\partial l}{\partial \mu_{1t}} &\propto -\phi\psi(\phi\mu_{1t}) + \phi\psi[\phi(1 - \mu_{1t} - \mu_{2t})] + \phi_{1t} - \phi \log(1 - y_{1t} - y_{2t}) \\ &\quad + \frac{s_t - 1}{\mu_{1t}} - \frac{u_t - 1}{1 - \mu_{1t} - \mu_{2t}}, \end{aligned}$$

$$\begin{aligned} \frac{\partial l}{\partial \mu_{2t}} &\propto -\phi\psi(\phi\mu_{2t}) + \phi\psi[\phi(1 - \mu_{1t} - \mu_{2t})] + \phi_{2t} - \phi \log(1 - y_{1t} - y_{2t}) \\ &\quad + \frac{r_t - 1}{\mu_{2t}} - \frac{u_t - 1}{1 - \mu_{1t} - \mu_{2t}}, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \mu_{1t}^2} &\propto -\phi^2\psi'(\phi\mu_{1t}) - \phi^2\psi'[\phi(1 - \mu_{1t} - \mu_{2t})] \\ &\quad - \frac{s_t - 1}{\mu_{1t}^2} - \frac{u_t - 1}{(1 - \mu_{1t} - \mu_{2t})^2}, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \mu_{2t}^2} &\propto -\phi^2\psi'(\phi\mu_{2t}) - \phi^2\psi'[\phi(1 - \mu_{1t} - \mu_{2t})] \\ &\quad - \frac{r_t - 1}{\mu_{2t}^2} - \frac{u_t - 1}{(1 - \mu_{1t} - \mu_{2t})^2}, \end{aligned}$$

$$\frac{\partial^2 l}{\partial \mu_{1t} \partial \mu_{2t}} \propto -\phi^2\psi'[\phi(1 - \mu_{1t} - \mu_{2t})] - \frac{u_t - 1}{(1 - \mu_{1t} - \mu_{2t})^2},$$

$$\hat{V}[\mu_t|D_t] \approx - \left[\frac{\partial^2 l}{\partial \mu_t^2} \Big|_{\mu_t = \hat{\mu}_t} \right]^{-1} = \begin{bmatrix} \frac{\partial^2 l}{\partial \mu_{1t}^2} & \frac{\partial^2 l}{\partial \mu_{1t} \partial \mu_{2t}} \\ \frac{\partial^2 l}{\partial \mu_{1t} \partial \mu_{2t}} & \frac{\partial^2 l}{\partial \mu_{2t}^2} \end{bmatrix},$$

onde $\hat{\mu}_t = (\hat{E}(\mu_{1t}|D_t), \hat{E}(\mu_{2t}|D_t))'$.

Considere que, após o uso do procedimento acima, obtem-se estimativas de

$$d_t = E[\mu_t|D_t, \phi] \quad e \quad O_t = V[\mu_t|D_t, \phi].$$

Utilizando tais valores e de maneira análoga ao segundo item do Passo 2,

sejam:

$$d_{it} = E[\mu_{it}|D_t, \phi] = E \left[\frac{\exp(\lambda_{it})}{1 + \sum_{j=1}^2 \exp(\lambda_{jt})} \middle| D_t, \phi \right] \\ \approx \frac{\exp\{E[\lambda_{it}|D_t, \phi]\}}{1 + \sum_{j=1}^2 \exp\{E[\lambda_{jt}|D_t, \phi]\}},$$

então,

$$E[\lambda_{it}|D_t, \phi] \approx \log \left(\frac{d_{it}}{1 - \sum_{j=1}^2 d_{jt}} \right) = f_{it}^*.$$

Para O_t , temos:

$$O_t = V[\mu_t|D_t, \phi] = V \left[\left(\frac{\exp(\lambda_{1t})}{1 + \sum_{j=1}^2 \exp(\lambda_{jt})} \right) \middle| D_t, \phi \right] \approx (\nabla f(f_t))' V[\lambda_t|D_{t-1}, \phi] (\nabla f(f_t)).$$

Então,

$$Q_t^* = V[\lambda_t|D_t, \phi] \approx [(\nabla f(f_t))']^{-1} O_t [\nabla f(f_t)]^{-1}.$$

– **Passo 4** - Estrutura Condicional para $(\beta_t|\lambda_t, D_{t-1}, \phi)$:

$$P(\beta_t|D_t, \phi) = \int P(\beta_t|\lambda_t, D_{t-1}, \phi) P(\lambda_t|D_t, \phi) d\lambda_t.$$

Utilizando o método de estimação Linear Bayes, temos:

$$\begin{cases} \hat{E}[\beta_t|\lambda_t, D_{t-1}, \phi] = a_t + R_t F_t Q_t^{-1} (\lambda_t - f_t), \\ \hat{V}[\beta_t|\lambda_t, D_{t-1}, \phi] = R_t - R_t F_t Q_t^{-1} F_t' R_t. \end{cases}$$

– **Passo 5**: Atualização de β_t :

$$P(\beta_t|D_t, \phi) = \int P(\beta_t|\lambda_t, D_{t-1}, \phi) P(\lambda_t|D_t, \phi) d\lambda_t.$$

Desta forma, os momentos a posteriori de $(\beta_t|D_t, \phi)$ podem ser estimados por

$$(\beta_t|D_t, \phi) \sim (m_t, C_t),$$

onde,

$$\begin{aligned} m_t &= E(\beta_t | D_t, \phi) = E[\hat{E}(\beta_t | \lambda_t, D_{t-1}, \phi) | D_t, \phi] \\ &= E[a_t + R_t F_t Q_t^{-1} (\lambda_t - f_t) | D_t, \phi] \\ &= a_t + R_t F_t Q_t^{-1} (f_t^* - f_t), \end{aligned}$$

e

$$\begin{aligned} C_t &= V(\beta_t | D_t, \phi) = V[\hat{E}(\beta_t | \lambda_t, D_{t-1}, \phi) | D_t, \phi] + E[\hat{V}(\beta_t | \lambda_t, D_{t-1}, \phi) | D_t, \phi] \\ &= V[a_t + R_t F_t Q_t^{-1} (\lambda_t - f_t) | D_t, \phi] + E[R_t - R_t F_t Q_t^{-1} F_t' R_t | D_t, \phi] \\ &= V[R_t F_t Q_t^{-1} \lambda_t | D_t, \phi] + (R_t - R_t F_t Q_t^{-1} F_t' R_t) \\ &= R_t F_t Q_t^{-1} Q_t^* Q_t^{-1} F_t' R_t + (R_t - R_t F_t Q_t^{-1} F_t' R_t) \\ &= R_t - R_t F_t Q_t^{-1} [I - Q_t^* Q_t^{-1}] F_t' R_t. \end{aligned}$$

Referências Bibliográficas

- Abramowitz, M. and Stegun, I. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 9th edition.
- Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of Royal Statistic Societ. Series B (Methodological)*, 44(2):139–177.
- Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*, volume 1. Chapman Hall/CRC, 2nd edition.
- Aitchison, J. (2003). *A Concise Guide to Compositional Data Analysis*. CDA Workshop, Girona.
- Aitchison, J. and Shen, S. (1980). Logistic-normal distributions: Some properties and uses. *Biometrika*, 67(2):261–272.
- Albi, M., Migon, H., and da Silva, C. (2009). Bayesian inference for beta distribution. Technical report, Departamento de Estatística - Universidade de Brasília.
- Boonstra, R. (2005). Equipped for life: the adaptive role of the stress axis in male mammals. *Journal of Mammalogy*, 86:2362–47.
- Branscum, A., Johnson, W., and Thurmon, M. (2007). Bayesian beta regression: Applications to household expenditure data and genetic distance between foot-and-mouth disease viruses. *Aust. N. Z. J. Stat.*, 49(3):287–301.
- Coakley, J. and Rust, B. (1968). Sedimentaion in an arctic lake. *J. Sedimentary Petrology*, 38:1290–1300.

- Corbit, D. (1996). Numerical integration: from trapezoids to rms: Objected-oriented numerical integration. *Dr. Dobb's J.*, 252:117–120.
- da Silva, C., Martins, E., and Reis, S. (2008). Bayesian capture-recapture analysis: An application in modeling semelparity of a neotropical didelphid marsupial. *Communication in Statistics Simulation and Computation*, 37:8298–41.
- Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *Journal of Applied Statistics*, 10:1–18.
- Gamerman, D. (1991). Dynamic bayesian models for survival data. *Applied Statistics*, 40:63–79.
- Gamerman, D. (1998). Markov chain monte carlo for dynamic generalized linear mixed models. *Biometrika*, 85:215–27.
- Gamerman, D. and Lopes, H. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, volume 1. Chapman Hall/CRC, 2nd edition.
- Gamerman, D. and West, M. (1987). A time series application of dynamic survival models in unemployment studies. *The Statistician*, 36:174–269.
- Godolphin, E. and Triantafyllopoulos, K. (2006). Decomposition of time series models in state-space form. *Computational Statistics and Data Analysis*, 50:2232–2246.
- Goldstein, M. (1976). Bayesian analysis of regression problems. *Biometrika*, 63:51–58.
- Gupta, A. and Varga, T. (1992). Characterization of matrix variate normal distributions. *Lourn. Multi. Analys.*, 41(1):80–88.
- Hartigan, J. (1969). Linear bayesian methods. *J. Royal Statistics Society (Ser. B)*, 31:446–454.
- Khoan, T. and Truc, T. (1994). A characterization of matrix variate normal distributions. *Internat. J. Math Math Sci.*, 17(2):341–346.

- Kieschnick, R. and McCullough, B. (2003). Regression analysis of variates observed on $(0,1)$: percentages, proportions, and fractions. *Statistical Modelling*, 3:193–213.
- Lindsey, J. and Lambert, P. (1995). Dynamic generalized linear models and repeated measurements. *Journal of Statistical Planning and Inference*, 47:129–139.
- MacMillan, N. and Creelman, C. (2005). *Detection theory: A user's guide*. Lawrence Erlbaum Associates, 2nd edition.
- Martins, E., Bonato, V., da Silva, C., and Reis, S. (2006a). Seasonality in reproduction, age structure and density of the gracile mouse opossum *gracilinanus microtarsus* (marsupialia: Didelphidae) in a brazilian cerrado. *Journal of Tropical Ecology*, 22:461–468.
- Martins, E., Bonato, V., da Silva, C., and Reis, S. (2006b). Partial semelparity in the neotropical didelphid marsupial *gracilinanus microtarsus*. *Journal of Mammalogy*, 87(5):915–920.
- McCullagh, P. and Nelder, J. (1994). *Generalized Linear Models*, volume 37. Monographs on Statistics and Applied Probability, 2nd edition.
- Migon, H. (1984). *An approach to non-linear Bayesian forecasting problems with applications*. PhD thesis, University of Warwick, Department of Statistics.
- Migon, H. and Gamerman, D. (1999). *Statistical Inference: An Integrated Approach*. Hodder Arnold Publication, London, first edition.
- Migon, H. and Harrison, P. (1985). An application of non-linear bayesian forecasting to television advertising. In *Bayesian Statistics 2*. J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith (Eds.). North-Holland, Amsterdam, and Valencia University Press.
- Muqattash, I. and Yahdi, M. (2006). Infinite family of approximations of the digamma functions. *Math. Comput. Modelling*, 43:287–301.

- Nelder, J. and Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society*, 135:370–384.
- Paolino, P. (2001). Maximum likelihood estimation of models with beta-distributed dependent variables. *Political Analysis*, 9:325–346.
- Quintana, J. and West, H. (1988). Time series analysis of compositional data. In *Bayesian Statistics 3*. J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith (Eds.). Oxford University Press.
- R (2009). *A Language and Environment for Statistical Computing*. R Foundation For Statistical Computing, Vienna, Austria, electronic edition. (<http://www.r-project.org>).
- Rue, H., Martino, S., and Chopin, N. (2006). Decomposition of time series models in state-space form. *Journal of the Royal Statistical Society: Series B*, 50(9):2232–2246.
- Smithson, M. and Verkuilen, J. (2005). A better lemon-squeezer? maximum likelihood regression with beta-distributed dependent variables. *Psychological Methods*, 11(1):54–71.
- Tierney, L. and Kadane, J. (1986). Accurate approximations for the posterior moments and marginal densities. *Journal of the American Statistical Association*, 81:82–86.
- West, M. and Harrison, P. (1997). *Bayesian Forecast and Dynamic Models*. Springer Verlag, 2nd edition.
- West, M., Harrison, P., and Migon, H. (1985). Dynamic generalized linear models and bayesian forecasting (with discussion). *Journal of the American Statistical Association*, 80:73–97.
- Williams, B., Nichols, J., and Conroy, M. (2002). *Analysis and Management of Animal Populations*. Academic Press.

Apêndice

A - Teorema de Bayes

Seja $y = y_1, \dots, y_n$ uma amostra aleatória de observações descritas pela função de densidade $f_Y(y_i|\theta)$ $i = 1, \dots, n$. Portanto, a função de verossimilhança é dada por: $l(\theta) = f_Y(y_1, \dots, y_n|\theta) = f_Y(y|\theta)$. A quantidade θ corresponde ao vetor de parâmetros correspondentes a função de densidade.

Dentro da abordagem *bayesiana*, a incerteza sobre o parâmetro θ é caracterizada probabilisticamente por meio de uma densidade $p(\theta)$.

Uma vez estabelecida a distribuição de probabilidade $p(\theta)$, o processo de inferência é baseado na distribuição de θ após observar y . Esta é chamada de *distribuição a posteriori* de θ e é obtida por meio do seguinte **teorema**:

$$p(\theta|y) = \frac{f_Y(y|\theta)p(\theta)}{f_Y(y)}, \quad (13)$$

onde

$$f_Y(y) = \int f_Y(y|\theta)p(\theta)d\theta. \quad (14)$$

Em geral, encontra-se a *distribuição a posteriori* por meio de:

$$p(\theta|y) \propto f_Y(y|\theta)p(\theta) \quad (15)$$

$$\int p(\theta|y)d\theta = 1. \quad (16)$$

B - Propriedades da Distribuição Normal

B.1 - Transformações lineares:

Se $X \sim N_p(\mu, \Sigma)$, A é uma matriz $r \times p$ e b um vetor de dimensão $r \times 1$, então:

$$y = Ax + b \sim N(Ax + b, A\Sigma A') \quad (17)$$

B.2 - Distribuições marginais

Se o vetor x é dividido em dois blocos, com x_1 contendo os primeiros r componentes e x_2 contendo os $p - r$ demais. Procedendo a mesma partição nos parâmetro μ e Σ , tem-se:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{e} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad (18)$$

As marginais são dadas por: $x_i \sim N(\mu_i, \Sigma_{ii})$, $i = 1, 2$.

B.3 - Distribuições condicionais

Considerando, ainda, as mesmas partições x_1 e x_2 de x , μ e Σ , segue que:

$$x_1|x_2 \sim N(\mu_{1.2}, \Sigma_{11.2}) \quad (19)$$

onde

$$\mu_{1.2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) \quad \text{e} \quad \Sigma_{11.2} = \Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}. \quad (20)$$

Lembrando que para estes resultados é necessário que as partições Σ_{11} e Σ_{22} tenham posto máximo, para que suas inversas existam.

Se $x_1|x_2 \sim N(\mu_1 + B(x_2 - \mu_2), B_2)$ e $x_2 \sim N(\mu_2, \Sigma_{22})$, então:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N(\mu, \Sigma) \quad (21)$$

com

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{e} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

onde

$$\Sigma_{11} = B_2 + B_1 \Sigma_{22} B_1' \quad \text{e} \quad \Sigma_{21}' = \Sigma_{12} = B_1 \Sigma_{22}. \quad (22)$$

C - Séries de Tendências Polinomiais

Séries de tendências polinomiais são úteis para descrever tendências que se desenvolvem, em geral, de maneira suave ao decorrer do tempo. No caso de séries curtas, as tendências costumam ser bem descritas por funções polinomiais de ordem 1 ou 2.

Os modelos lineares dinâmicos polinomiais formam um subconjunto da classe de séries temporais em modelos dinâmicos. Qualquer modelo pertencente a este subconjunto tem função da forma:

$$f_t(k) = a_{t0} + a_{t1}k + \dots + a_{t,n-1}k^{n-1} \quad (23)$$

que define um modelo linear dinâmico *polinomial de ordem n*.

Algumas características descrevem os modelos dinâmicos polinomiais:

- um modelo linear dinâmico é polinomial de ordem n se e somente se ele for caracterizado por $\{F_t, G_t, V_t, W_t\} = \{E_n, J_{(n)}(1), V_t, W_t\}$ onde:

$$E_n = (1, 0, \dots, 0)' \quad \text{e} \quad J_{(n)}(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & & 0 \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{bmatrix}$$

- qualquer modelo linear dinâmico equivalente ao modelo constante caracterizado por $\{E_n, J_{(n)}(1), V, W\}$ é um modelo dinâmico polinomial constante de ordem n .

D - Simulação de Séries de Tempo com Distribuição Beta

Uma vez gerados os valores dos parâmetros θ_t das equações de evolução e estabelecido o parâmetro de escala ϕ_t , obtém-se μ_t por:

$$\mu_t = \frac{\exp(\theta_t)}{1 + \exp(\theta_t)} \quad (24)$$

Em seguida, são gerados os dados Y_t em que: $Y \sim \text{Beta}(\phi_t\mu, \phi[1 - \mu_t])$. Segue abaixo a programação em R:

```
#####  
# Modelo Beta Dinâmico  
#  
#####  
set.seed(123456)  
  
# Gerando o modelo beta dinâmico de primeira ordem  
# -----  
W<-0.5  
  
phi<-2  
n<-100  
x0<-0.00  
w<-rnorm(n,0,sqrt(W)) #gera os erros de y_t  
x<-rep(0,n)  
mu<-x  
y<-x  
x[1]<-x0 + w[1]  
mu[1]<-exp(x[1])/(1+exp(x[1]))  
y[1]<-rbeta(1,phi*mu[1],phi*(1-mu[1]))  
for (t in 2:n){
```

```

x[t]<-x[t-1] + w[t] #gera beta_t
mu[t]<-exp(x[t])/(1+exp(x[t])) #gera mu_t
y[t]<-rbeta(1,phi*mu[t],phi*(1-mu[t]))
}
y1<-y
x1<-x
mu1<-mu

# Gerando o modelo beta dinâmico de segunda ordem
# -----

W1<-0.001
W2<-0.001
phi<-2
n<-100
x0<-c(0.00,0.00)
w<-cbind(rnorm(n,0,sqrt(W1)),rnorm(n,0,sqrt(W2)))
### simulando a distrib. conjunta      ####
###          dos erros de beta_t      ####

y<-rep(0,n)
x<-matrix(0,n,2)
mu<-y
x[1,2]<-x0[2] + w[1,2] #gera beta2t
x[1,1]<-x0[1] + x[1,2] + w[1,1] #gera beta1t
mu[1]<-exp(x[1,1])/(1+exp(x[1,1]))
y[1]<-rbeta(1,phi*mu[1],phi*(1-mu[1])) #gera y_t
for (t in 2:n){
  x[t,2]<-x[t-1,2] + w[t,2]
  x[t,1]<-x[t-1,1] + x[t,2] + w[t,1]
  mu[t]<-exp(x[t,1])/(1+exp(x[t,1]))
  y[t]<-rbeta(1,phi*mu[t],phi*(1-mu[t]))
}

```

}

y2<-y

x2<-x

mu2<-mu

E - Estimação Via Linear Bayes

Sejam ϵ_t e ω_t os erros aleatórios das equações de observação e evolução respectivamente, estabelecidos de forma:

$$\epsilon_t \sim (0, V_t); \quad \omega_t \sim (0, W_t).$$

Ou seja, as distribuições de probabilidade não são conhecidas e estão definidos somente o 1º e 2º momentos das variáveis.

Seja d uma estimativa de θ e $L(d, \theta)$ a medida de *erro* da estimativa.

$d = m = m(y)$ é ótima ao minimizar $r(d) = E[L(d, \theta)]$, definido como o risco a posteriori. Utilizando a função *quadrática* de risco, tem-se:

$$L(d, \theta) = (\theta - d)'(\theta - d)$$

O risco a posteriori é minimizado em $m = E(\theta|Y)$ (média a posteriori) e o risco mínimo é o traço da variância a posteriori, $r(m) = V[\theta|Y]$.

$$\begin{pmatrix} \theta \\ Y \end{pmatrix} \sim \left[\begin{pmatrix} a \\ f \end{pmatrix}, \begin{pmatrix} R & AQ \\ AQ' & R \end{pmatrix} \right]$$

Estimação via *Linear Bayes* proporciona uma estimativa alternativa que pode ser vista como uma aproximação ótima condicional a Y (mais informações sobre este estimador podem ser encontradas em Kartigan, 1969, Goldstein, 1976 e West e Harrison, 1997, capítulo 4).

$$r(d) = \text{tr}\{E[(\theta - m)(\theta - m)']\}.$$

O estimador $d = d(Y)$ pode ser reescrito como uma função linear de Y :

$$d(Y) = h + HY$$

Consequentemente, a função $E(\theta|Y)$ pode ser aproximada de um modelo linear.

A estimativa via Linear Bayes de θ é a forma linear que otimiza em termos de minimizar a função *risco*. Sejam m e C definidos por:

$$m = a + A(Y - f)$$

$$C = R - AQA'$$

$E[(\theta - m)(\theta - m)'] = C$ e o risco mínimo $r(m) = tr(C)$. Os resultados são demonstrados a seguir:

Para $d = h + HY$ e $R(d) = E[(\theta - d)(\theta - d)']$ segue que:

$$\begin{aligned} R(d) &= E(\theta\theta') - E(\theta d') - E(d\theta') + E(dd') \\ &= E(\theta\theta') - E(\theta h' + \theta(HY)') - E(h\theta' + HY\theta') \\ &\quad + E(hh' + h(HY)' + HYh' + HY(HY)') \\ &= R' + aa' - ah' - AQH' + a(Hf)' - HQA' + (Hf)a' + hh' \\ &\quad + h(Hf)' + (Hf)h' + HQH + Hf(Hf)' \\ &= R + HQH - AQH' - HQA' + (a - h - Hf)(a - h - Hf)'. \end{aligned}$$

Como: $(H - A)Q(H - A)' = AQA' + HQH - AQH' - HQA'$,

$R(d)$ pode ser reescrito como:

$$R - AQA' + (H - A)Q(H - A)' + (a - h - Hf)(a - h - Hf)'.$$

$r(d) = tr[R(d)]$ é composto pela soma de 3 termos:

- $tr[(R - AQA)']$, que independe de d ;
- $tr[(H - A)Q(H - A)']$, que é mínimo ($= 0$) para $H = A$;
- $tr[(a - h - Hf)(a - h - Hf)']$, que é mínimo ($= 0$) para $h + Hf = a$.

Desta forma, $r(d)$ é mínimo com $H = A$ e $h = a - Af$:

$$d(Y) = a + A(Y - f) = m$$

$$E[(\theta - m)(\theta - m)'] = R - AQA' = C.$$

Portanto, independente da distribuição dos erros, a média linear a posteriori é que minimiza o *risco quadrático*.

F - Simulação de Dados Normal Matriz-Variada

A simulação de dados com distribuição *Normal Matriz-Variada* pode ser feita partindo da propriedade de “vetorização” dos dados:

- Se $X_{n \times q} \sim N_{n \times q}(\mu, \Omega, \Sigma)$ então, a “vetorização” de X segue uma distribuição normal multivariada dada por:

$$\text{vec}(X)_{(nq \times 1)} \sim N_{nq}(\text{vec}(\mu), \Omega \otimes \Sigma)$$

Com isso, são gerados, primeiramente, dados *Normais Multivariados* Z :

$$Z \sim N_{nq}(0, \Omega \otimes \Sigma)$$

Em seguida, X é obtido a partir da relação:

$$Z = \text{vec}(X)$$

Segue abaixo a programação em R:

```
#####  
### Gerando dados normal matriz variada ###  
#####  
  
## definindo as matrizes de variância e covariância  
## a esquerda e a direita respectivamente  
  
omega<-matrix(c(1,.8,.8,1),2,2)  
omega  
      [,1] [,2]  
[1,]  1.0  0.8  
[2,]  0.8  1.0  
  
sigma<-matrix(c(1,.3,.5,.3,1,.2,.5,.2,1),3,3)  
sigma  
      [,1] [,2] [,3]
```

```

[1,] 1.0 0.3 0.5
[2,] 0.3 1.0 0.2
[3,] 0.5 0.2 1.0

# produto de kronecker

sig<-omega%x%sigma

## gerando a normal multivariada (6) da variável VEC(X)

z.aux<-mvrnorm(100,rep(0,6),sig, tol = 1e-6, empirical = FALSE)

## Resgatando X de VEC(X)

x<-array(0,c(100,2,3))

for(i in 1:100){
x[i,,]<-matrix(z.aux[i,],2,3,byrow=T)}

## checagem das correlações entre linhas e colunas
## da normal matriz variada X

cor(x[,1,1],x[,2,1])
#correlação entre linha 1 e linha 2 na coluna 1 (fixa)
[1] 0.8365494

cor(x[,1,2],x[,2,2])

```

```
#correlação entre linha 1 e linha 2 na coluna 2 (fixa)
[1] 0.746508

cor(x[,1,3],x[,2,3])
#correlação entre linha 1 e linha 2 na coluna 3 (fixa)
[1] 0.8270757

cor(x[,1,1],x[,1,2])
#correlação entre coluna 1 e coluna 2 na linha 1 (fixa)
[1] 0.3737405

cor(x[,2,1],x[,2,2])
#correlação entre coluna 1 e coluna 2 na linha 2 (fixa)
[1] 0.3570263

cor(x[,1,1],x[,1,3])
#correlação entre coluna 1 e coluna 3 na linha 1 (fixa)
[1] 0.510031

cor(x[,2,1],x[,2,3])
#correlação entre coluna 1 e coluna 3 na linha 2 (fixa)
[1] 0.5732864

cor(x[,1,2],x[,1,3])
#correlação entre coluna 2 e coluna 3 na linha 1 (fixa)
[1] 0.2654559

cor(x[,2,2],x[,2,3])
#correlação entre coluna 2 e coluna 3 na linha 2 (fixa)
[1] 0.3453939
```

G - Análise Retrospectiva

Em séries de tempo, além do intuito de previsão, olhar para trás pode também auxiliar a enxergar o verdadeiro comportamento da série inteira. Para dados variando nos tempos de 1 a T , a análise retrospectiva utiliza a distribuição suavizada $p(\theta_t|D_T)$, para $t = 1, \dots, T$.

Para um Modelo Dinâmico Linear com $1 < h < t$, temos:

$$\begin{aligned}(\theta_{h-t}|D_t) &\sim [a_t(-h), R_t(-h)], \quad \text{onde} \\ a_t(-h) &= m_{t-h} - B_{t-h}[a_{t-h+1} - a_t(-h+1)], \\ R_t(-h) &= R_{t-h} - B_{t-h}[R_{t-h+1} - R_t(-h+1)]B'_{t-h}, \\ B_t &= C_t G'_{t+1} R_{t+1}^{-1}, \\ a_t(0) &= m_t \quad \text{e} \quad R_t(0) = C_t.\end{aligned}$$

Com $h = 1$, tem-se

$$\begin{aligned}a_t(-1) &= m_{t-1} - B_{t-1}[m_t - a_t], \\ R_t(-1) &= C_{t-1} - B_{t-1}[R_t - C_t]B'_{t-1},\end{aligned}$$

e B_{t-1} como definido acima. Mais detalhes de análise retrospectiva (ou *suavização*) podem ser encontrados no Capítulo 4 de West e Harrison (1997).