

Modelos de Fronteira de Produção Estocástica:
Uma Abordagem Dinâmica Para Múltiplos
Produtos.

por

Luis Alberto Toscano Medrano

Orientador: Hélio S. Migon

Universidade Federal do Rio de Janeiro

Instituto de Matemática

Dezembro de 2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Modelos de Fronteira de Produção Estocástica: Uma Abordagem Dinâmica Para Múltiplos Produtos.

Luis Alberto Toscano Medrano

Orientador: Hélio S. Migon

Tese de Doutorado submetida ao programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do grau de Doutor em Estatística.

Aprovada por:

Presidente Prof. Hélio S. Migon
IM-UFRJ

Prof^a. Alexandra M. Schmidt
IM-UFRJ

Prof. Dani Gamerman
IM-UFRJ

Prof. Geraldo da Silva e Souza
UNB-Embrapa

Prof^a. Roseli Aparecida Leandro
ESALQ-USP

Rio de Janeiro, Dezembro de 2008

Medrano, Luis

Modelos de Fronteira de Produção Estocástica: Uma Abordagem Dinâmica Para Múltiplos Produtos/ Luis Medrano. – Rio de Janeiro: UFRJ/IM, 2008.

xvii, 133 f. : il. ; 31cm.

Tese (Doutorado) – UFRJ/IM. Programa de Pós-Graduação em Estatística, 2008.

Orientador: Hélio S. Migon

Referências bibliográficas: p.119–126.

1. Estatística Matemática - Tese. I. Migon, Hélio. II. Universidade Federal do Rio de Janeiro. Instituto de Matemática. III. Título.

RESUMO

Modelos de Fronteira de Produção Estocástica: Uma Abordagem Dinâmica Para Múltiplos Produtos.

Luis Alberto Toscano Medrano

Orientador: Hélio S. Migon

Resumo da Tese de Doutorado submetida ao programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do grau de Doutor em Estatística.

Nesta primeira parte da tese vamos desenvolver uma análise Bayesiana baseada na priori de Jeffreys para alguns modelos de fronteira estocástica. A proposta é obter a distribuição do erro composto usando o método de Laplace, quando for o caso, e determinar a matriz de informação de Fisher para obter uma priori de Jeffreys para os parâmetros de assimetria e de graus de liberdade.

Na segunda parte da tese uma nova classe de modelos de fronteira estocástica para múltiplos produtos derivados do modelo fatorial é proposta. Um modelo fatorial com transformação de Box-Cox será utilizado para obter um único fator latente, definindo-se no logaritmo deste fator a fronteira estocástica.

Na ultima parte da tese, uma extensão do modelo de fronteira estocástica para múltiplos produtos é proposta. Neste modelo é considerado que as ineficiências técnicas variem no tempo. Uma estrutura dinâmica é especificada para modelar a variação temporal existente entre as ineficiências técnicas;

Estudos simulados são apresentados para testar a aplicabilidade dos modelos. Adicionalmente, uma aplicação com dados de produção em pesquisas agrícolas da Embrapa é mostrada.

Sumário

1	Introdução	1
1.1	Organização da tese	2
2	Seleção de Modelos de Fronteira Estocástica	4
2.1	Modelo básico de fronteira estocástica	5
2.1.1	Inferência Bayesiana	6
2.2	Métodos de MCMC	9
2.2.1	Metropolis Otimizado	9
2.2.2	Slice Sampling	10
2.2.3	Estudo Numérico	10
2.3	Critério de Seleção	16
2.4	Seleção de Modelos baseados em dados artificiais	17
2.4.1	Sensibilidade do DIC ao tamanho da amostra (N) e a elasticidade de substituição (ζ)	18
2.4.2	Seleção da função de produção: Cobb-Douglas e CES	19
3	Modelo de fronteira estocástica com erro assimétrico	21
3.1	Introdução	22
3.2	A fronteira Normal – Normal Truncada	24
3.2.1	O Modelo	24
3.2.2	Procedimento de Inferência	28

3.3	A Fronteira Normal – Gama	31
3.3.1	O Modelo	32
3.3.2	Procedimento de Inferência	34
3.4	Fronteira t-Student – t-Student Truncada	37
3.4.1	O Modelo	37
3.4.2	Procedimento de Inferência	40
3.5	Aplicação com dados artificiais	46
3.5.1	Esquema de MCMC	46
3.5.2	Resultados	49
3.6	Aplicação a dados Reais	50
4	Modelo de Fronteira Estocástica Fatorial	55
4.1	Introdução	56
4.2	Função CET	58
4.3	Uma abordagem de fronteira estocástica para múltiplos produtos .	61
4.3.1	O Modelo	62
4.3.2	Inferência Bayesiana e algoritmo MCMC	63
4.4	Modelo Fatorial Bayesiano	65
4.4.1	Modelo Fatorial Canônico	65
4.4.2	Identificação do modelo	66
4.4.3	Inferência Bayesiana e algoritmo de MCMC	67
4.4.4	Curva de possibilidade de produção	69
4.5	Modelo Fatorial com Transformação de Box-Cox	70
4.5.1	O Modelo	70
4.5.2	Inferência Bayesiana e Algoritmo MCMC	71
4.5.3	Curva de possibilidade de produção	73
4.6	Modelo de fronteira estocástica fatorial com transformação Box-Cox	76
4.6.1	O Modelo	76

4.6.2	Inferência Bayesiana e algoritmo MCMC	77
4.7	Estudo com dados artificiais	79
4.7.1	Exemplo 1: Análise Fatorial com transformação Box-Cox .	80
4.7.2	Exemplo 2: Estudo da $E(f y, L, \Psi, q)$	82
4.7.3	Exemplo 3: Fronteira estocástica fatorial com transformação Box-Cox	85
4.7.4	Exemplo 4: Uma outra alternativa de fronteira estocástica para múltiplos-produtos.	88
4.8	Aplicação a dados reais	94
5	Modelo de Fronteira Estocástica Fatorial Dinâmico	98
5.1	Introdução	98
5.2	Modelo Proposto	100
5.3	Inferência Bayesiana e algoritmo de MCMC	104
5.3.1	Distribuição a priori	104
5.3.2	Inferência a posteriori	105
5.4	Estudo simulado	108
5.5	Aplicação a dados da Embrapa	112
6	Conclusões e extensões	115

Lista de Figuras

2.1	Função de autocorrelação para os valores da cadeia de P e θ para dois esquemas de amostragem e dois diferentes valores de P . A linha cheia são as médias de 100 replicações e as linhas quebradas são os intervalos de credibilidade de 90%.	13
2.2	Box-plots do fator de ineficiência dos métodos de MCMC para parâmetros P , θ , β_2 e σ_v^2 com $P = 0.8$	14
2.3	Box-plots do fator de ineficiência dos métodos de MCMC para parâmetros P , θ , β_2 e σ_v^2 com $P = 2$	14
2.4	Box-plots do fator de ineficiência de u_{min} , u_{med} e u_{max} para os 2 esquemas de amostragem e diferentes valores de P	15
3.1	Densidades normais assimétricas	25
3.2	Curvas de contorno da função de verossimilhança condicional de λ versus α	29
3.3	Curvas de contorno da distribuição a posteriori de λ versus α . . .	31
3.4	Distribuição normal - exponencial	34
3.5	Curva de contorno da função de verossimilhança condicional de θ versus τ_v	35
3.6	Curva de contorno da distribuição a posteriori de θ versus τ_v . . .	36
3.7	Densidade da t-assimétrica	38
3.8	Curvas de contorno da verossimilhança condicional de λ versus α . . .	40

3.9	Curvas de contorno da verossimilhança condicional de η versus α .	41
3.10	Curva de contorno da distribuição a posteriori de λ versus α	43
3.11	Curva de contorno da distribuição a posteriori de η versus α	45
3.12	Histograma para os parâmetros do modelo de fronteira normal – normal truncada utilizando a priori Jeffreys para λ . As retas verticais tracejadas e pontilhadas representam a média e o IC 95% a posteriori, respectivamente.	52
3.13	Histograma para os parâmetros do modelo de fronteira t-student – t-student truncada utilizando a priori Jeffreys para λ . As retas verticais tracejadas e pontilhadas representam a média e o IC 95% a posteriori, respectivamente.	53
4.1	Curvas de possibilidade de produção para dois produtos. Todas as curvas correspondem a $g(y_i) = 1$	60
4.2	Elasticidade de transformação. A linha cheia representa $\varepsilon = 0$, a linha tracejada corresponde $\varepsilon = -1$ e a linha pontilhada representa $\varepsilon \rightarrow -\infty$	61
4.3	Curvas de possibilidade de produção para dois produtos. Todas as curvas correspondem a $E(f_i y, L, \Psi, q) = 1$	75
4.4	Histograma das distribuições a posteriori dos parâmetros - Modelo Fatorial usando a transformação de Box-Cox. As retas verticais tracejada e pontilhada representam a média e o IC 95% respectivamente.	81
4.5	Gráfico de dispersão de $E(f resto)$ versus o valor real de f	81
4.6	Box-plot das amostras de $E(f y, L, \Psi, q)$. A linha cheia horizontal representa o valor verdadeiro de $g(y)$	83
4.7	Box-plot das amostras de $E(f y, L, \Psi, q)$. A linha cheia horizontal representa o valor verdadeiro de $g(y)$	84

4.8	Histograma da distribuições a posteriori dos parâmetros do modelo - Modelo Fronteira Fatorial com transformação Box-Cox. As retas verticais tracejada e pontilhada representam a média e o IC 95% respectivamente.	86
4.9	Gráficos de dispersão do valor esperado do fator versus o valor verdadeiro.	87
4.10	Gráficos de dispersão do valor esperado da eficiência versus o valor verdadeiro.	87
4.11	Histograma dos parâmetros - Modelo de fronteira estocástica para múltiplos-produtos proposto por Fernández et al (2000). As retas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%, respectivamente.	90
4.12	Histograma dos parâmetros - Modelo de fronteira estocástica fatorial com transformação de Box-Cox. As retas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%, respectivamente.	91
4.13	Gráfico de dispersão do valor esperado de ρ versus o seu valor verdadeiro.	92
4.14	Histogramas dos <i>ranking</i> da eficiência para as firmas 10 e 17 segundo o modelo de fronteira estocástica fatorial	93
4.15	Histogramas dos <i>ranking</i> da eficiência para as firmas 10 e 17 segundo o modelo proposto por Fernández <i>et al.</i> (2000)	93
4.16	Histograma das distribuições marginais a posteriori segundo o Modelo de fronteira estocástica fatorial	95
4.17	Gráfico de dispersão do valor esperado de ρ apresentado em Souza <i>et al.</i> (1996) versus o valor esperado de ρ utilizando o modelo de fronteira fatorial.	96

4.18	Gráfico de dispersão do valor esperado de ρ obtido pelo modelo da Fernández <i>et al.</i> (2000) versus o valor esperado de ρ utilizando o modelo de fronteira fatorial.	96
4.19	Histograma dos <i>rankings</i> da eficiência para as unidades de pesquisas da Embrapa <i>UD1</i> e <i>UD26</i>	97
5.1	Histograma dos parâmetros - modelo de fronteira estocástica dinâmico fatorial. As retas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%, respectivamente.	109
5.2	Histograma do rank das ineficiências para $t = 1, 2, 3, 4$	110
5.3	Histograma do rank das ineficiências para $t = 5, 6, 7, 8$	111
5.4	Histograma dos parâmetros segundo o Modelo de fronteira estocástica fatorial dinâmico para dados da Embrapa. As retas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%, respectivamente.	113
5.5	Histograma dos <i>rankings</i> da eficiência para as unidades de pesquisas da Embrapa <i>UD13</i> e <i>UD21</i> no período de 1996 – 1999.	114

Lista de Tabelas

2.1	tempo computacional (em segundos)	12
2.2	Resultado do Experimento 1	19
2.3	Resultado do DIC para os 100 conjuntos de dados artificiais baseado na função de produção CES.	20
3.1	Estimativas a posteriori para o modelo de fronteira normal – exponencial e estimador de máxima verossimilhança.	50
3.2	Tempo Computacional.	50
3.3	Eficiência para as cinco primeiras firmas.	54
1	Estatística \hat{R} de Gelman & Rubin para Seção 3.6	126
2	Estatística \hat{R} de Gelman & Rubin para Seção 4.7.1 - Modelo Fatorial usando a transformação de Box-Cox	127
3	Estatística \hat{R} de Gelman & Rubin para Seção 4.7.3 - Modelo Fatorial usando a transformação de Box-Cox	127
4	Estatística \hat{R} de Gelman & Rubin para Seção 4.8 - Modelo Fatorial usando a transformação de Box-Cox	127
5	Estatística \hat{R} de Gelman & Rubin para Seção 5.4 - modelo de fronteira estocástica dinâmico fatorial.	127

Capítulo 1

Introdução

Os modelos de fronteira de produção têm origem na teoria microeconômica e descrevem o comportamento dos agentes econômicos na busca de otimizar seus desempenhos, produzindo o máximo baseado na menor quantidade possível de insumos. Diversos fatores podem fazer com que produzam abaixo da possibilidade máxima de produção admitida pela tecnologia atual. Por isso surgem as medidas de ineficiência técnica. Os modelos de fronteira de produção estocástica foram usados amplamente para descrever a produtividade e a eficiência da firma. Esses modelos foram introduzidos por Agnier et al. (1977) e Meeusen e Van den Broecker (1977). Um modelo de fronteira de produção estocástica decompõe o produto em dois componentes, o primeiro é um componente determinístico que inclui a função de produção e outras variáveis que afetam a produtividade, o segundo é um termo de erro composto por dois componentes. Um dos componentes do erro, é normalmente distribuído, representando pelo o que não pode ser controlado pelas firmas, ou seja, o distúrbio aleatório. O outro é um componente de erro assimétrico que representa a ineficiência de cada agente (e mede a distância da fronteira). Na literatura encontram-se diferentes propostas para a distribuição da ineficiência: a exponencial (Meeusen; Van Den Broecker, 1977), a

normal truncada em zero (Agnier et al., 1977; Stevenson, 1980), a gama (Greene, 1990) e a lognormal (Migon, 2006).

Inicialmente somente problemas de um único produto eram tratados economicamente pelas fronteiras estocásticas. Para lidar com a situação de múltiplos produtos a abordagem econométrica lança mão de diversos expedientes. Um deles consiste em modelar a função de custo, a qual é obtida pela minimização dos custos dos insumos dado ao nível de produção. Assim, a função de custo é explicada pelos preços dos insumos e pelos vários produtos. Outra alternativa é modelar o fator crítico, isto é, modelar aquele insumo essencial para a produção dos múltiplos produtos. Vale notar que a modelagem determinística (DEA), formalmente desenvolvido por Charnes, Cooper e Rhodes (1978), permite a modelagem de múltiplos produtos. Do ponto de vista econométrico, o único artigo da literatura a lidar com esta questão sob uma abordagem Bayesiana é Fernández et al. (2000).

Relativo à inferência, esses modelos econométricos podem ser utilizados sob uma abordagem clássica ou Bayesiana. Em ambas a natureza do erro composto faz com o que a inferência mereça certos cuidados especiais. Assim, mesmo em modelos de um único produto já nos deparamos com dificuldades na estimação dos parâmetros que caracterizam a parte assimétrica da fronteira estocástica. É claro que, como já mencionado, temos várias outras decisões importantes na modelagem estocástica, como a escolha da distribuição da componente de erro associada a ineficiência e a seleção da função de produção.

1.1 Organização da tese

Nesta tese iremos apresentar respostas a algumas das questões introduzidas neste capítulo. Assim, os demais aspectos estão organizados da seguinte forma. Dois estudos desenvolvidos na fase inicial de nossa pesquisa, e ora materializados nos

artigos "*Bayesian Stochastic Frontier: model selection via deviance based criteria*" e "*Sampling Schemes for Asymmetric Models a comparative study*" serão apresentados no capítulo 2. No primeiro discutimos a seleção de modelos, incluindo a forma funcional da fronteira de produção e a escolha da distribuição do termo de ineficiência, no outro consideramos algoritmos alternativos de simulação estocástica. No capítulo 3 caracterizamos diversas formas alternativas de erros compostos e apresentaremos as prioris de Jeffreys para o parâmetro que caracteriza a assimetria da distribuição do erro composto e o graus de liberdade quando for o caso. A proposta neste capítulo será: obter a distribuição do erro composto usando o método de Laplace quando for o caso, determinar a matriz de informação de Fisher e obter a priori de Jeffreys para os parâmetros de assimetria e de graus de liberdade quando for o caso.

Um modelo de fronteira estocástica para múltiplo produtos seguindo uma abordagem Bayesiana será descrito no capítulo 4. Um modelo fatorial com transformação de Box-Cox será utilizado para obter um único fator latente, definindo-se no logaritmo deste fator a fronteira estocástica. Este modelo possibilita um número maior de combinações entre os produtos transformados para construir um único fator latente, ou seja, similar a elasticidades constantes de transformação (CET). No capítulo 5, uma extensão do modelo de fronteira estocástica para múltiplos produtos serão proposto, considerando que as ineficiências técnicas variam no tempo. Esta generalização permite incorporar nas ineficiências algum tipo de influência temporal ou até mesmo de variáveis exógenas. Uma estrutura dinâmica será especificada para modelar a variação temporal existente entre as ineficiências técnicas; Por último, no Capítulo 6 serão apresentadas as conclusões e possíveis extensões.

Capítulo 2

Seleção de Modelos de Fronteira Estocástica

A escolha da função de produção e da distribuição da componente de erro associada a ineficiência é decisão importante nos modelos de fronteira estocástica. A maioria dos estudos assume a distribuição normal truncada com o parâmetro de locação igual a zero ou a exponencial pela facilidade na estimação dos parâmetros. Estas propostas são pouco flexíveis uma vez que impõem a densidade da ineficiência que esteja concentrada mais perto da origem. Uma alternativa é assumir que a distribuição da ineficiência seja gama ou log-normal, desta forma com uma moda deslocada da origem. Em Medrano e Migon (2008) é apresentado um estudo simulado comparando as distribuições gama e log-normal nos modelos de fronteira estocástica. Na seção 2.4 serão apresentados os principais resultados obtidos neste estudo de simulação.

A inferência Bayesiana apresenta neste modelo algumas questões que precisam ser tratadas cuidadosamente. Alguns parâmetros não possuem a distribuição condicional a posteriori conhecida e, no caso da fronteira normal-gama, esta distribuição condicional não é log-côncava, desta forma é difícil amostrar desta dis-

tribuição. Recentemente, Tsionas (2000) apresentou uma proposta para gerar destas distribuições. O algoritmo proposto parece não produzir estimativas precisas em alguns casos. Uma alternativa para gerar destas distribuições utilizando o algoritmo de Slice Sampling (Neal, 2003) é apresentado em Medrano e Migon (2007). Este algoritmo, juntamente com os principais resultados obtidos num estudo simulado serão apresentados, resumidamente, na seção 2.2.

2.1 Modelo básico de fronteira estocástica

Seja y_i o logaritmo do produto para firma i , para $i = 1, 2, \dots, N$. O modelo de fronteira estocástica é definido por

$$y_i = f(x_i, \beta) + v_i - u_i \quad v_i \sim N(0, \sigma_v^2) \quad (2.1)$$

onde $N(0, \sigma_v^2)$ denota a distribuição normal com média zero e variância σ_v^2 , $f(x_i, \beta)$ que representa a função de produção, com x_i sendo o logaritmo do vetor de insumos e β um vetor de coeficientes. A componente v_i é responsável por capturar o erro de medida, sendo, portanto, simétrica. A componente u_i é não-negativa e responsável pela ineficiência técnica das unidades operacionais.

Duas diferentes distribuições para a componente de ineficiência u_i em (2.1) foram consideradas em Medrano e Migon (2007). Mais especificamente, os u_i 's foram modelados como

$$u_i \sim G(P, \theta) \quad \text{ou} \quad u_i \sim LN[\mu, \psi^2], \quad (2.2)$$

em que $G(P, \theta)$ denota a distribuição gama com parâmetros $P, \theta > 0$ e $LN[\mu, \psi^2]$ que denota a distribuição log-normal com parâmetros $\mu \in \Re$ e $\psi^2 > 0$.

As funções de produção para um único produto Q e dois insumos K, L considerado em Medrano e Migon (2007) foram: a Cobb-Douglas:

$$Q = \gamma K^{\beta_1} L^{\beta_2} \quad \text{para} \quad \beta_1, \beta_2 > 0$$

e a CES (Constant Elasticity of Substitution):

$$Q = \gamma[(1 - \delta)K^{-\rho} + \delta L^{-\rho}]^{-v/\rho}$$

para $0 < \gamma < \infty, 0 < \delta < 1, -\infty < v < \infty$ e $-1 < \rho < \infty$. Note que em ambas especificações $x = (\log K, \log L)$ e $f(x, \beta) = \log Q$.

Apesar da função Cobb-Douglas ser mais fácil de estimar e de se manipular matematicamente, as propriedades impostas pela estrutura de produção exigem retorno constante de escala, $\beta_1 + \beta_2 = 1$. O retorno de escala reflete o grau em que um aumento proporcional em todas as quantidades de insumos, aumenta a quantidade do produto. Outro conceito relacionado a função de produção é a elasticidade de substituição, o qual mede a facilidade com que se pode substituir um insumo por outro. A elasticidade de substituição da função CES é dada por $\varepsilon = 1/(1 + \rho)$.

Vale destacar que será dada uma maior atenção à comparação dos modelos de fronteira de produção estocástica com diferentes distribuições para a ineficiência.

Função de Verossimilhança

Suponha que $y = (y_1, \dots, y_N)'$ são observações i.i.d. do modelo descrito em (2.1). A função de verossimilhança de (β, σ_v^2, u) será

$$L(\beta, \sigma_v^2, u|y) \propto (\sigma_v^2)^{-n/2} \exp \left\{ -\frac{1}{2\sigma_v^2} \sum_{i=1}^N (y_i - f(x_i, \beta) + u_i)^2 \right\} \quad (2.3)$$

2.1.1 Inferência Bayesiana

Nesta seção é apresentado o procedimento de inferência proposto seguindo uma abordagem completamente Bayesiana. Inicialmente, as distribuições a priori para todos os parâmetros são apresentados. Em seguida, as distribuições condicionais

completas a posteriori, necessárias à implementação do algoritmo de MCMC, são apresentadas.

Distribuição a priori

Para realizar a análise Bayesiana necessitamos especificar a distribuição a priori de todos os parâmetros dos modelos. Foi assumindo uma independência a priori para os parâmetros β, σ_v^2, P e θ para a fronteira normal – gama e β, σ_v^2 e Ψ^2 para o caso log-normal. Em todos os modelos, a distribuição a priori atribuído para σ_v^2 foi a distribuição gama inversa com parâmetros n_0, a_0 , denotado por $\sigma_v^2 \sim GI(n_0, a_0)$, enquanto que para os parâmetros relacionados a componente de ineficiência, as distribuições a priori foram: (i) $P \sim G(d_0, \varepsilon_0)$ e $\theta \sim G(\nu_0, \omega_0)$ no caso normal – gama e (ii) $\psi^2 \sim IG(q_0, D_0)$ no caso normal – lognormal. Para os demais dos parâmetros, as distribuições a priori foram:

(I) Cobb-Douglas: $\beta_j \sim N(b_0, H_0)$, $j = 0, \dots, 3$.

(II) CES: $\log \gamma \sim N(b_0, H_0)$, $v \sim N(m_0, s_0)$, $\delta \sim Be(w_0, \varphi_0)$ e $\rho \sim NT_{[-1, \infty]}(r_0, g_0)$ onde $Be(w_0, \varphi_0)$ denota a distribuição beta com parâmetros $w_0 > 0, \varphi_0 > 0$, e $NT_{[-1, \infty]}(r_0, g_0)$ denota a distribuição normal truncada em -1 com parâmetros de posição r_0 e de escala g_0 .

Distribuição a posteriori

A distribuição a posteriori conjunta de $\beta, \sigma_v^2, \lambda$ e u é dada por

$$p(\beta, \sigma_v^2, \lambda, u|y) \propto \prod_{i=1}^N p(y_i|\beta, u_i, \sigma_v^2) \prod_{i=1}^N p(u_i|\lambda) p(\beta) p(\sigma_v^2) p(\lambda) \quad (2.4)$$

a qual é analiticamente intratável e, portanto, inferência a posteriori será feita utilizando-se simulação estocástica - MCMC. As distribuições condicionais completas dos parâmetros são obtidas a partir da distribuição conjunta a posteriori (2.4) são

- Para a fronteira normal-gama

$$\begin{aligned}
& - (\theta|y, \beta, \sigma_v, P, u) \sim G\left(NP + v_0, \sum_{i=1}^N u_i + \omega_0\right) \\
& - p(P|y, \beta, \sigma_v, \theta, u) \propto P^{d_0-1} \Gamma(P)^{-N} \exp(QP) I_{(0,\infty)}(P), \quad \text{where } Q = \\
& \quad \sum_{i=1}^N \log u_i - \varepsilon_0 + N \log \theta \\
& - p(u_i|y, \beta, \sigma_v, P, \theta) \propto u_i^{P-1} \exp\left[\frac{-(u_i + e_i)^2}{2\sigma^2} - \theta u_i\right], \quad 1, \dots, N.
\end{aligned}$$

- Para a fronteira normal-lognormal

$$\begin{aligned}
& - (\psi^2|y, \beta, \sigma_v, u) \sim G[d_0 + N, D_0 + \sum_1^N (\log(u_i - \mu)^2) - 1] \\
& - p(u_i|y, \beta, \sigma_v, \sigma_u) \propto N[y_i - x'_i \beta, \sigma^2] LN[0, \psi^2], \quad 1, \dots, N.
\end{aligned}$$

As outras distribuições condicionais a posteriori foram:

- Cobb-Douglas:

$$(\beta|y, \sigma_v, u) \sim N_k \left[(H_0 + \sigma^{-2} X' X)^{-1} (H_0 b_0 + X' (Y + u) \sigma^{-2}), H_0 + \sigma^{-2} X' X \right]$$

- CES:

$$\begin{aligned}
& - (\log \gamma|y, \sigma_v, \delta, v, \rho, u) \sim N \left[(H_0 + \sigma^{-2})^{-1} (H_0 b_0 + \mathbf{1}' (Y + u) \sigma^{-2}), H_0 + \sigma^{-2} \right] \\
& - p(\delta|y, \sigma_v, v, \gamma, \rho, u) \propto \exp \left[-\frac{(y+u-X^*\pi)'(y+u-X^*\pi)}{2\sigma^2} \right] \\
& - p(v|y, \sigma_v, \delta, \gamma, \rho, u) \propto \exp \left[-\frac{(y+u-X^*\pi)'(y+u-X^*\pi)}{2\sigma^2} - \frac{s_0(v-m_0)^2}{2} \right] \\
& - (\rho|y, \sigma_v, \delta, v, \gamma, u) \propto \exp \left[-\frac{(y+u-X^*\pi)'(y+u-X^*\pi)}{2\sigma^2} - \frac{(\rho-\tau_0)^2}{2g_0} \right] I_{(-1,\infty)}(\rho) \\
& \quad \text{where } X^* = \left[\begin{array}{c} 1 \\ \log\{[(1-\delta)\exp(X_1)^{-\rho} + \delta\exp(X_2)^{-\rho}]^{-1/\rho}\} \end{array} \right] \text{ e} \\
& \quad \pi = (\beta \quad v)'
\end{aligned}$$

Em todos os modelos a distribuição condicional a posteriori para σ_v^{-2} foi

$$(\sigma_v^{-2}|y, \beta, u) \sim G \left[\frac{(N + N_0)}{2}, \frac{(y + u - X\beta)'(y + u - X\beta) + a_0}{2} \right].$$

Observamos que algumas dessas distribuições condicionais completas não têm forma conhecida e ainda mais no caso da fronteira normal-gama a distribuição condicional de u_i não é log côncava quando o parâmetro $P < 1$. Neste caso, um método de aceitação e rejeição otimizado (Metropolis otimizado) foi proposto por Tsionas (2000). Uma estratégia alternativa neste caso é utilizar o método Slice Sampling (Neal, 2003). Em Medrano e Migon (2007) é realizado um estudo simulado que compara esse dois métodos no qual o método Slice Sampling teve melhores resultados. Esses métodos são brevemente apresentados na próxima seção, assim como com os principais resultados obtidos no estudo simulado.

2.2 Métodos de MCMC

2.2.1 Metropolis Otimizado

Seja $p(x)$ a densidade que queremos amostrar e seja $g(x; \alpha)$ a distribuição proposta para x , defina-se $R(x, \alpha) = p(x)/g(x, \alpha)$ e $r(x, \alpha) = \log R(x, \alpha)$, em que o parâmetro α é escolhido para maximizar $R(x, \alpha)$, isto é, para encontrar o $\max_x \min_{\alpha} \{r(x, \alpha)\}$.

Dado um valor ótimo de α , o algoritmo segue em dois passos:

1. Amostrar x^{new} de $g(x, \alpha)$.
2. Aceitar x^{new} com probabilidade $\frac{R(x^{new})}{R(x^*)}$ onde x^* é o valor ótimo de x .

Algumas distribuições propostas, usadas em Tsionas (2000), são a exponencial e a gama. Mais detalhes destes procedimentos poderão ser encontrados no artigo citado.

2.2.2 Slice Sampling

Uma outra alternativa quando a distribuição condicional não é log-côncava é utilizar o método de Slice Sampling proposto por Neal (2003). Vamos supor novamente que nós queremos amostrar da densidade $p(x)$. O algoritmo Slice consiste em introduzir uma variável auxiliar z e temos que a distribuição conjunta de x e z é uniforme na região $U = \{(x, z) : 0 < z < p(x)\}$.

Para amostrar de x basta amostrar conjuntamente de (x, z) e ignorar z . Gerar (x, z) conjuntamente amostrando uniformemente de U pode ser muito difícil. Uma alternativa é utilizar o amostrador de Gibbs, onde as condicionais completas são uniformes.

(a) Gerar $z^{(i)}$ de $(z|x) \sim U(0, g(x^{(i-1)}))$

(b) Gerar $x^{(i)}$ de $(x|z) \sim U(S(z^{(i)}))$ onde $S(z^{(i)}) = \{x : p(x) \geq z^{(i)}\}$.

Dentre as vantagens desse método estão: a não necessita de especificar uma densidade proposta como em Metropolis-Hastings e somente utiliza a distribuição uniforme para gerar os valores. O método é detalhado em Neal (2003), em que se propõe uma maneira de solucionar a determinação de $S(z)$.

Na próxima seção vamos apresentar um estudo simulado com o objetivo de avaliar a performance dos métodos de MCMC aqui apresentados.

2.2.3 Estudo Numérico

Este exercício vai mostrar um estudo de Monte Carlo em que várias amostras simuladas são geradas com o objetivo de avaliar a eficiência dos esquemas de MCMC. Os esquemas de amostragem apresentados na seção anterior produzem cadeias com diferentes propriedades de convergência. Características teóricas, como a autocorrelação da cadeia ou a taxa de convergência podem ser usadas para compará-las.

A autocorrelação é importante para determinar a eficiência do estimador de Monte Carlo. A variância do estimador de Monte Carlo, $\bar{h}_M = (1/M) \sum_{i=1}^M h(\theta^{(i)})$ é dada por $Var[\bar{h}_M] = (\sigma_h^2/M)i$, onde σ_h^2 é a variância a posteriori de $h(\theta)$ e i é o fator de ineficiência associado com a cadeia:

$$i = 1 + 2 \sum_{k=1}^{M-1} \frac{(M-k)}{M} \rho_k \quad (2.5)$$

onde ρ_i é a autocorrelação do lag k para a cadeia de valores de $h(\theta)$. Neste estudo vamos calcular o fator de ineficiência i para comparar os esquemas de amostragem. Mais detalhes sobre este critério de eficiência por ser visto em Reis, Salazar e Gamerman (2006).

O estudo de simulação desenvolvido nesta seção envolvem dois cenários do modelo de fronteira estocástica normal-gama, ambos com $N = 100$. Os cenários diferem nos valores dos parâmetros da distribuição gama. No primeiro, $(P, \theta) = (0.8, 1)$, que correspondem ao caso em que a distribuição condicional completa da ineficiência u não é log-côncava, e o segundo, $(P, \theta) = (2, 1)$. Neste caso, as estimativas obtidas por Tsionas (2000) não eram muito precisas, pois seu exemplo é baseado em uma única amostra simulada. O outro parâmetro da distribuição gama é fixado em $\theta = 1$.

Para cada cenário alternativo foram geradas e analisadas 100 replicações através dos dois esquemas amostrais alternativos discutidos na seção anterior. Foram construídos o intervalo de credibilidade para a função de autocorrelação e para o fator de ineficiência.

Um total de 15.000 iterações para cada cadeia foram realizadas descartando-se as primeiras 5.000 iterações como período de aquecimento. Nas 10 primeiras replicações, para cada parâmetro foram geradas duas cadeias paralelas (iniciam por valores diferentes), em que a convergência foi verificada graficamente para todos os parâmetros de interesse. O tempo computacional gasto por cada método serão apresentadas na Tabela 2.1. Note que para qualquer valor de P , o método

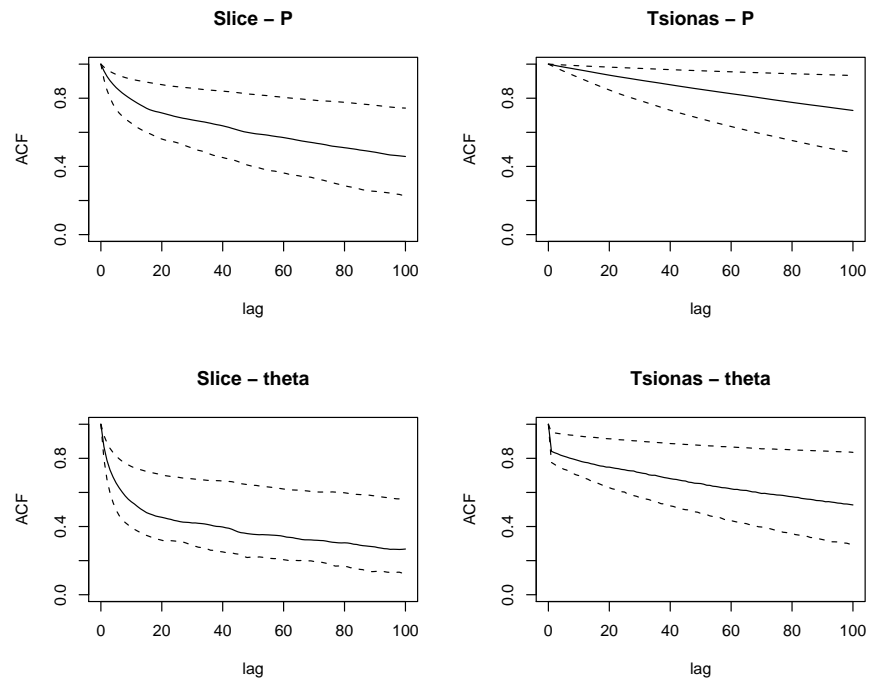
Metropolis otimizado leva cerca de 1,5 vezes mais do que o método Slice Sampling, isto ocorre devido a este método requerer uma otimização a cada iteração.

Tabela 2.1: tempo computacional (em segundos)

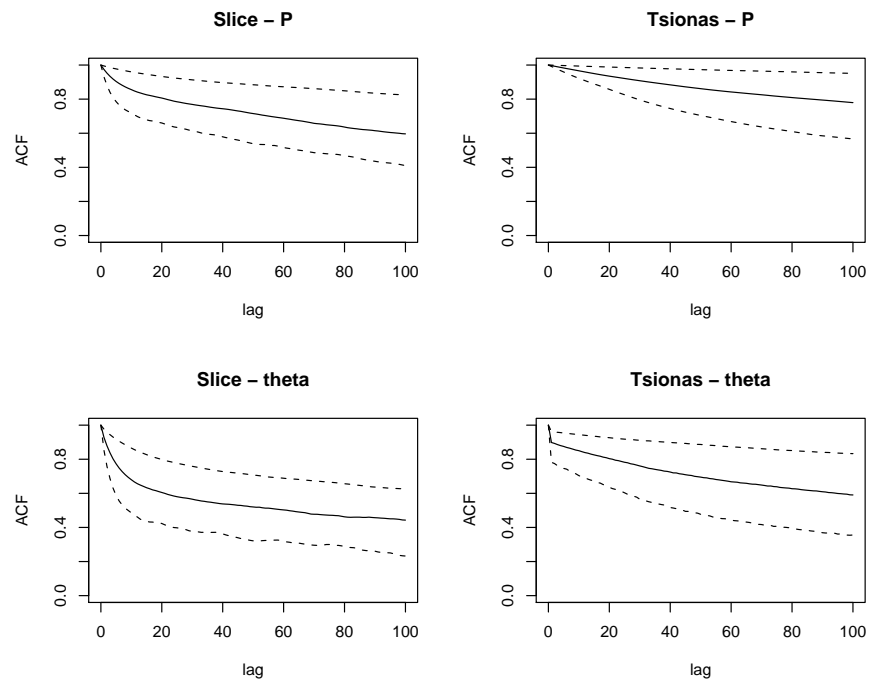
Esquemas de amostragem	$P = 0.8$	$P = 2$
Slice	329	320
Metropolis Otimizado	442	454

A Figura 2.1 mostra a função de autocorrelação para as cadeias de valores de P e θ . Para cada uma das duas combinações dos valores de P são obtidas 100 estimativas da função de autocorrelação, e assim, a média e envelopes podem ser construídos. Para P e θ , o método Metrópolis otimizado mostra uma pequena queda na autocorrelação para as cadeias geradas, refletindo a dependência forte entre os parâmetros previamente mencionados. O método Slice Sampling, em troca, apresenta uma decadência muito rápida nessa autocorrelações. Por exemplo, para cada caso $P = 2$, a autocorrelação para cada cadeia gerada com método Slice Sampling é próximo de 0.50 no lag 20, enquanto o Metropolis otimizado alcança o mesmo valor somente no lag 100.

Estes resultados obtidos podem ser resumidos agregando as autocorrelações obtidas, ou seja, observando a ineficiência amostral dos esquemas de MCMC. Para cada valor de P e cada esquema, 100 valores do fator de ineficiência para θ , P e σ^2 são obtidos, um para cada replicação. Estes podem ser resumidos no Box-plots das Figuras 2.2 e 2.3. Para todos os esquemas considerados, as ineficiências ficam maiores quando P é menor. As figuras mostram que o método Slice Sampling parece ser mais eficiente para P , θ , β_2 e σ_v^2 . Como pode ser visto as caixas centrais não se sobrepõem para quase todas as alternativas simuladas e todos os parâmetros examinados.



(a) $P=0.8$



(b) $P=2$

Figura 2.1: Função de autocorrelação para os valores da cadeia de P e para dois esquemas de amostragem e dois diferentes valores de P . A linha cheia são as médias de 100 replicações e as linhas quebradas são os intervalos de credibilidade de 90%.

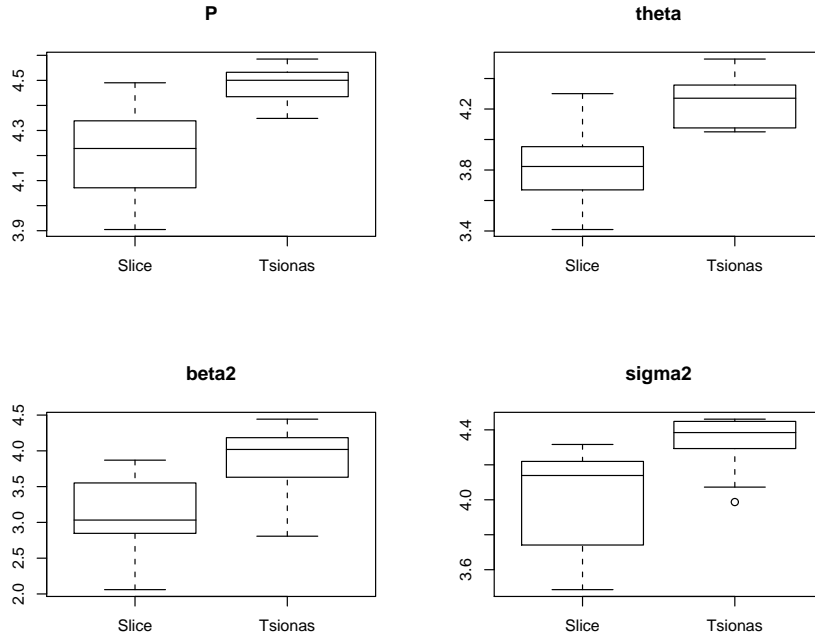


Figura 2.2: Box-plots do fator de ineficiência dos métodos de MCMC para parâmetros

P , θ , β_2 e σ_v^2 com $P = 0.8$.

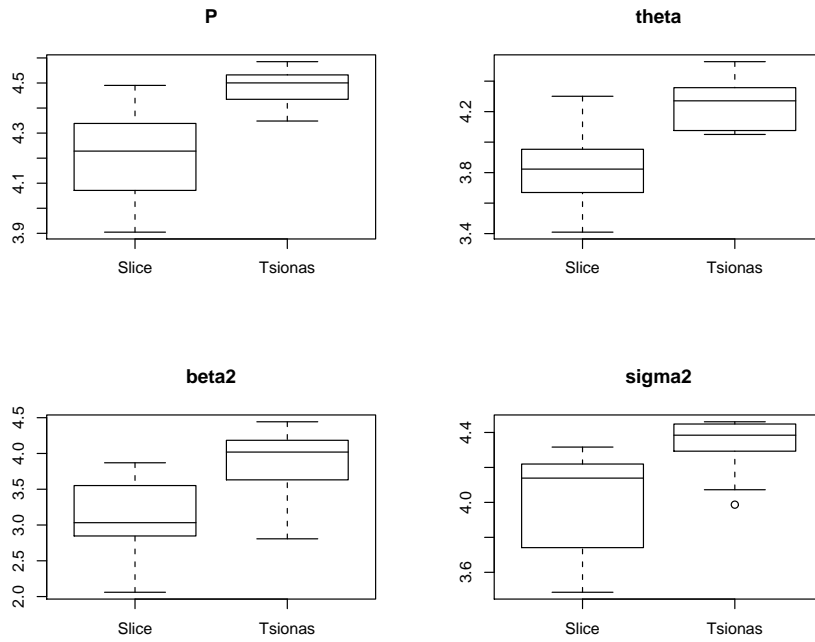
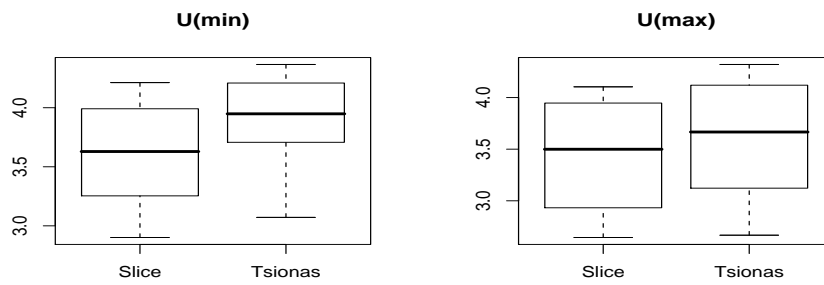


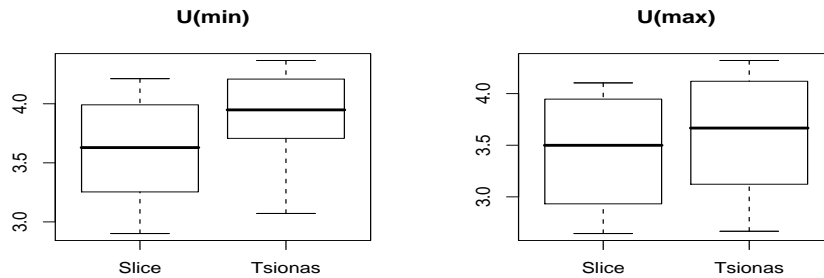
Figura 2.3: Box-plots do fator de ineficiência dos métodos de MCMC para parâmetros

P , θ , β_2 e σ_v^2 com $P = 2$.

Por último, para cada valor de P e cada esquema são obtidos 100 valores do fator de ineficiências de u_{max} e u_{min} , onde que u_{max} e u_{min} representam a ineficiência técnica da pior e da melhor firma, respectivamente. Estes podem ser resumidos no Box-plots da Figura 3.14. Novamente, para quase todas as alternativas de simulação, as medianas das ineficiências obtidas, usando o método de Slice, são ligeiramente melhores.



(a) $P=0.8$



(b) $P=2$

Figura 2.4: Box-plots do fator de ineficiência de u_{min} , u_{med} e u_{max} para os 2 esquemas de amostragem e diferentes valores de P .

Assim, na aplicação apresentada na seção 2.4 será usado somente o método de Slice Sampling para gerar as distribuições condicionais completas, que não têm uma forma conhecida. Na próxima seção serão apresentados os critérios de seleção de modelos que serão usados no estudo simulado.

2.3 Critério de Seleção

O problema de selecionar o melhor de uma coleção de modelos candidatos têm uma longa historia na literatura estatística. Durante anos, pesquisadores que trabalham dentro do paradigma Bayesiano usaram o fator de Bayes frequentemente, critério baseado na comparação da verossimilhança marginal (Kass e Raftery, 1985), para este propósito. Porém, Han e Carlin (2001) apresentam uma revisão de algumas dificuldades computacionais e conceituais usando o fator de Bayes para comparar os modelos hierárquicos complexos. O comumente usado critério de informação Akaike (AIC) (Akaike, 1973) e o critério de informação Bayesiano (BIC) (Schwarz, 1978), requer a especificação do número de parâmetros em cada modelo.

Essas dificuldades conduziram o desenvolvimento de critérios de escolhas de modelos Bayesianos. Estes incluem a análise residual de validação cruzada (Gelfand et. al., 1992) e método de teoria da decisão por minimização da perda preditiva a posteriori (EPD) (Gelfand e Ghosh, 1998).

Mais recentemente, Spiegelhalter et. al. (2002) sugere a generalização do AIC e deriva um critério de informação da deviance (DIC) para comparar modelos hierárquicos complexos, as quais os números de parâmetros não são bem definidos. O DIC é dado por

$$DIC = \overline{D(\alpha)} + P_D \quad (2.6)$$

no qual a complexidade do modelo é capturado através do efetivo número de

parâmetros $P_D = E_{\alpha|y}[D] + D(E_{\alpha|y}(\alpha))$. Note que P_D é frequentemente menor que o total do número de parâmetros do modelo devido ao "empréstimo da força" através dos parâmetros dos níveis individuais dentro dos modelos hierárquicos.

Neste estudo, a estimativa do DIC aplicado nos modelos de fronteira estocástica será calculado através da função de verossimilhança obtida pela verossimilhança dada em 2.3, assim obtemos a seguinte deviance

$$D(\alpha) = -2 \log p(y_i|\alpha) \quad (2.7)$$

$$= \sum_{i=1}^N \left(\frac{(y_i - f(x_i, \beta) + u_i)^2}{\sigma_v^2} + \log \sigma_v^2 + c \right) \quad (2.8)$$

onde $\alpha = (\sigma_v^2, \beta, u)$. Note que a deviance $D(\alpha)$ será o mesmo, independente da distribuição atribuída a componente de ineficiência u .

2.4 Seleção de Modelos baseados em dados artificiais

O objetivo desta seção é verificar se o critério de seleção proposto é capaz de identificar os verdadeiros processos geradores dos dados. Vamos apresentar um exercício de simulação que ilustra a sensibilidade do critério, variando algumas funções dos parâmetros.

Para analisar a sensibilidade do critério de seleção DIC, foram realizados dois experimentos. O primeiro, para o modelo de fronteira estocástica com função Cobb-Douglas com duas distribuições alternativas para a componente de ineficiência. Vamos variar o tamanho da amostra (N) e a proporção da variância do termo de ineficiência (u) sobre o total da variância do modelo, $\zeta = \frac{V(u)}{\sigma^2 + V(u)}$. No segundo experimento, para o modelo de fronteira estocástica com função CES, com as mesmas distribuições alternativas para a componente de ineficiência, consideramos valores diferentes para a elasticidade de substituição $\varepsilon = \frac{1}{1+\rho}$. Em

ambos os experimentos serão gerados 100 conjuntos de dados.

Nestes estudos simulados vamos fixar $\mu = 0$ na distribuição log-normal, obtendo mesmo assim uma moda fora da origem, mas com a vantagem de estimar apenas um parâmetro. Para realizar uma comparação imparcial entre os modelos, vamos determinar que as duas distribuições alternativas (gama e log-normal) tenham média e variância similares. Estas distribuições são semelhantes em relação a média e a variância, se definimos que os parâmetros da distribuição gama sejam: $P = [\exp(\psi^2) - 1]^{-1}$ e $\theta = [\exp(0.5\psi^2)(\exp(\psi^2) - 1)]^{-1}$ onde ψ^2 é o parâmetro de escala da distribuição lognormal. Assim, nos estudos simulados será definido $\psi^2 = 0.5$ e obtemos $p = 1,54$ e $\theta = 1,2$ na distribuição gama.

Os resultados seguintes são baseados em 1.000 iterações de MCMC. Foram geradas 15.000 iterações descartando-se as 5.000 primeiras com período de aquecimento e guardando a cada 10 iterações.

2.4.1 Sensibilidade do DIC ao tamanho da amostra (N) e a elasticidade de substituição (ζ)

Neste experimento nós observaremos a capacidade do critério DIC para selecionar o verdadeiro processo gerador dos dados. A função de produção considerada na análise foi a Cobb-Douglas com duas variáveis regressoras e duas alternativas para a distribuição da ineficiência são usadas. O objetivo deste exercício é determinar a sensibilidade do critério de seleção DIC para valores alternativos de ζ e N .

O experimento difere nos valores de ζ (0,70 ou 0,95) e nos valores de N (50 ou 100). Os valores definidos para o coeficiente de regressão foram $\beta = (1, 1, 1)'$. E a variância da componente de erro simétrico foi $\sigma^2 = \frac{V(u)(1-\zeta)}{\zeta}$.

Na Tabela 2.2 a frequência em que o critério de seleção DIC seleciona o modelo correto dentre as 100 replicações de amostras artificiais.

Observamos que o critério de seleção DIC pode identificar o verdadeiro pro-

Tabela 2.2: Resultado do Experimento 1

	Modelo verdadeiro	N	ζ	Modelo ajustado	
				N-LN	N-Ga
DIC	N-LN	50	0.70	71	29
			0.95	80	20
		100	0.70	56	44
			0.95	77	23
	N-Ga	50	0.70	29	71
			0.95	37	63
		100	0.70	17	83
			0.95	21	79

cesso gerador dos dados em mais de 50% dos conjuntos de dados artificiais. A porcentagem dos resultados bem sucedido aumenta com o valor de ζ no modelo normal-lognormal (N-LN), enquanto que no modelo normal-gama, o comportamento é o oposto, como era de se esperar. Também podemos ver que a porcentagem de resultados bem sucedidos aumenta com o tamanho da amostra quando o modelo normal-gama é ajustado, mas o mesmo não é verdade para o normal-lognormal.

2.4.2 Seleção da função de produção: Cobb-Douglas e CES

Neste segundo experimento, o modelo de fronteira estocástica com a função Cobb-Douglas será comparado com a função CES. Da mesma forma que no experimento anterior duas alternativas de distribuições foram usadas para descrever a componente de ineficiência (lognormal e gama). O objetivo é verificar a capacidade

do critério de seleção DIC discriminar para valores alternativos de ϵ . Os valores considerado para ϵ neste estudo simulado foram: $\epsilon = 0,5$, $\epsilon = 1$ (o qual corresponde ao modelo com a Cobb-Douglas) e $\epsilon = 2$. Foram gerados 100 conjuntos de dados artificiais, cada uma com $N = 100$ observações e $\sigma_v^2 = 0,05$. Os outros parâmetros definidos foram: $(\gamma, v, \delta) = (2, 72; 2; 0, 5)$.

A frequência com que o critério de seleção DIC escolhe o modelo verdadeiro em 100 conjuntos de dados artificiais será apresentado na Tabela 2.3. Observamos que o DIC identificou o modelo verdadeiro em mais de 50% das amostras geradas. Por exemplo, o DIC pode identificar o modelo verdadeiro em 69 das 100 amostras do modelo normal-log-normal que foi gerado usando a função CES com $\epsilon < 1$. Outro resultado acontece quando geramos o modelo usando a função CES com $\epsilon = 1$. Observamos que o DIC escolhe o modelo usando a função Cobb-Douglas, já que a CES com $\epsilon = 1$ se torna a função Cobb-Douglas. Notamos também que o DIC pode identificar bem os modelos que assumem um erro gama ou log-normal no termo de ineficiência.

Tabela 2.3: Resultado do DIC para os 100 conjuntos de dados artificiais baseado na função de produção CES.

	Modelo verdadeiro	ϵ	Modelo ajustado			
			N-LN		N-Ga	
			CD	CES	CD	CES
DIC	N-LN	0.5	0	69	6	25
		1	72	5	17	6
		2	1	68	13	18
	N-Ga	0.5	0	14	17	69
		1	16	1	65	18
		2	0	16	24	60

Capítulo 3

Modelo de fronteira estocástica com erro assimétrico

O desenvolvimento de novos modelos para análise de dados assimétricos vem sendo uma constante na literatura estatística atual. As distribuições assimétricas vêm despertando grande interesse nos últimos anos devido a sua vasta aplicação. A aplicação, abordada nesta tese, são os modelos de fronteira estocástica. Mostra-se neste capítulo 3 que em alguns dos modelos de fronteira estocástica conhecidos na literatura a densidade marginal do erro composto pertence a classe de distribuições assimétricas (Azzalini, 1985). Estudos anteriores mostram que pode existir dificuldades na estimação de alguns parâmetros do modelo de regressão com erro assimétrico.

A proposta aqui consiste na utilização de prioris de referência. Seguindo estas linhas de pesquisas, neste estudo propõe-se a marginalização do erro composto, usando o método de Laplace e a construção da priori de Jeffreys. Duas alternativas de modelagem se apresentam: através da verossimilhança marginal ou via modelagem hierárquica. Em ambos os casos usa-se a priori de referência mencionada.

3.1 Introdução

Uma família de distribuições assimétricas foi introduzida por Azzalini (1985, 1986). Inicialmente definida para a distribuição normal assimétrica, pode ser geralmente aplicada a qualquer distribuição assimétrica contínua. Por exemplo, no caso univariado, temos a distribuição Cauchy assimétrica proposta por Arnold e Beaver (2000), a t-assimétrica proposta por Branco e Dey (2001) e Azzalini e Capitanio (2003) e a distribuição logística assimétrica proposta por Wahed e Ali (2001). A característica principal destas distribuições é a introdução de um novo parâmetro que controla a assimetria.

Os modelos de regressão simples podem ser facilmente generalizados ao considerar que o termo de erro tenha distribuição assimétrica. Como consequência, o modelo de regressão com erro assimétrico têm sido recentemente usado em uma variedade de aplicações. Um contexto muito natural para o uso das distribuições assimétricas aparece nos modelos de fronteira estocástica. Por exemplo, no modelo de fronteira normal – normal truncada, o erro composto segue exatamente a distribuição normal assimétrica. Então, este modelo de fronteira estocástica pode ser representado como um modelo de regressão com erros normais assimétricos. A extensão para erros t-assimétricos pode ser obtido através do modelo de fronteira t-Student – t-Student truncada. No modelo de regressão com erro normal assimétrico, a estimação do parâmetro de assimetria apresenta o problema de não convergência no caso do método de máxima verossimilhança (ver Liseo e Loperfido, 2006), os quais sugerem a utilização de uma priori de referência para solucionar esses problemas. Nesta tese, mostra-se que este problema também pode ser encontrado em outras distribuições assimétricas comumente usadas nos modelos de fronteira estocástica.

Outras distribuições assimétricas podem ser obtidas a partir de outros modelos de fronteira estocástica. Como por exemplo, o modelo de fronteira normal –

gama, a densidade marginal do erro composto não tem uma expressão fechada. A densidade marginal do erro composto pode ser aproximada resolvendo-se a integral por alguma técnica numérica. A alternativa proposta neste estudo é aproximar esta integral através do método de Laplace, obtendo-se a seguir a priori de Jeffreys. Será mostrado, para alguns modelos de fronteira estocástica, que esta prática resolve as dificuldades decorrentes da má informação da função de verossimilhança.

A proposta deste capítulo é, portanto, apresentar uma análise Bayesiana objetiva do modelo assimétrico, com ênfase na aplicação a fronteira de produção estocástica, baseado na priori de Jeffreys. Nesta tese investigamos os seguintes modelos de fronteira estocástica: (i) normal – normal truncada, (ii) t-Student – t-Student truncada e (iii) normal – gama. A inferência Bayesiana poderá ser realizada considerando-se a função de verossimilhança marginal (modelo assimétrico) ou a modelagem hierárquica. A modelagem hierárquica relaciona-se com a modelagem formulada por O’Hagan e Leonard (1976). A inferência sobre as ineficiências é feita a partir dessa componente, ou seja, através de uma variável latente gerada juntamente com os outros parâmetros a cada iteração do algoritmo MCMC.

Este capítulo está organizado da seguinte forma. Nas três primeiras seções apresentamos a relação entre a distribuição assimétrica e o modelo de fronteira estocástica, obtemos a distribuição do erro composto e em seguida a distribuição a priori de Jeffreys que é desenvolvida para cada modelo de fronteira estocástica investigada. Finalizando, na seção 3.5 é apresentada duas aplicações: uma a dados artificiais e outra a dados reais.

3.2 A fronteira Normal – Normal Truncada

Um dos primeiros artigos a modelar a ineficiência técnica foi proposto por Aigner et al. (1977) onde a ineficiência técnica tem distribuição normal truncada em zero com parâmetros μ e σ_u^2 . Esta proposta assume que $\mu = 0$, isto implica em assumir que as ineficiências estão próximas da origem. Várias generalizações sobre a distribuição da componente de ineficiência foram feitas desde então. Stevenson (1980) propôs o uso da distribuição normal truncada em zero com $\mu \neq 0$ para modelar as ineficiências, de forma a permitir que a moda não seja nula.

3.2.1 O Modelo

O modelo de fronteira estocástica normal – normal truncada é expresso como

$$y_i = x_i\beta + e_i, \quad i = 1, \dots, N \quad \text{com} \quad e_i = v_i - |u_i| \quad (3.1)$$

onde $v_i \sim N(0, \sigma_v^2)$ e $u_i \sim N(0, \sigma_u^2)$.

Para obter uma priori de Jeffreys, primeiro necessitamos obter a distribuição marginal do erro composto $e_i = v_i - |u_i|$.

Distribuição Normal Assimétrica

No contexto das distribuições assimétricas, para obter a distribuição marginal do erro composto precisaremos da proposição abaixo, cuja prova pode ser encontrada em Henze (1986).

Proposição 3.1. *Considere as variáveis aleatórias independentes $U, V \sim N(0, 1)$ e defina*

$$Z = \sqrt{1 - \delta^2}V + \delta|U|$$

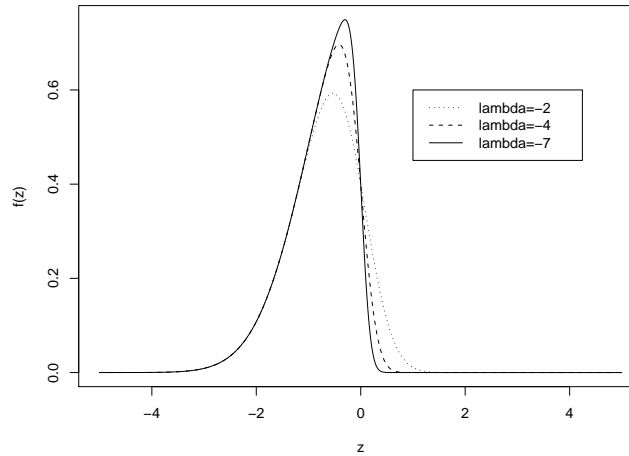
onde $-1 < \delta < 1$. Então a variável aleatória (v.a.) Z tem distribuição normal

assimétrica com parâmetro de assimetria λ e sua função de densidade é dada por

$$f(z|\lambda) = 2\phi(z) \Phi\left(\frac{\delta}{\sqrt{1-\delta^2}}z\right). \quad (3.2)$$

Aqui ϕ e Φ denotam a função de densidade e a função de distribuição da variável aleatória normal padrão, respectivamente. Usaremos a notação $Z \sim SN(\lambda)$ onde $\lambda = \frac{\delta}{\sqrt{1-\delta^2}}$. Pode-se verificar que se $Z \sim SN(\lambda)$ e $Y = \mu + \alpha Z$ então a v.a. Y tem distribuição normal assimétrica com parâmetros de posição μ , de escala α e de assimetria λ ; para simplificar denotaremos por $Y \sim SN(\mu, \alpha^2, \lambda)$. A média e variância de Y são, respectivamente, $E[Y] = \mu + \alpha\delta\frac{2}{\pi}$ e $Var[Y] = \alpha^2[1 - \frac{2}{\pi}\delta^2]$ onde $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. Outras propriedades da distribuição normal assimétrica podem ser encontradas em Genton (2004b). A Figura 3.1 mostra o comportamento desta densidade para diferentes valores do parâmetro de assimetria λ com posição $\mu = 0$ e escala $\alpha = 1$ fixos. Observe que a medida que o parâmetro λ cresce, o mesmo ocorre com a assimetria. Para valores negativos de λ a assimetria fica a esquerda.

Figura 3.1: Densidades normais assimétricas



Para obtermos a densidade marginal do erro e_i , basta definir que $e_i = \alpha z_i$ onde $\alpha = \sqrt{\sigma_u^2 + \sigma_v^2}$ e z_i é a normal assimétrica obtida através da proposição 3.1 com $\delta = -\frac{\sigma_u}{\sqrt{\sigma_u^2 + \sigma_v^2}}$.

Aproximação via Laplace

Uma alternativa para obter a densidade marginal do erro composto no caso da fronteira normal – normal truncada baseia-se na integração sobre $u_i \in \Re$ da densidade conjunta

$$f(u_i, e_i | \sigma_v, \sigma_u) = \frac{1}{2\pi\sigma_v\sigma_u} \exp \left\{ -\frac{(e_i + u_i)^2}{2\sigma_v^2} - \frac{u_i^2}{2\sigma_u^2} \right\} \quad (3.3)$$

Obtendo-se:

$$f(e_i | \sigma_v, \sigma_u) = \frac{1}{2\pi\sigma_v\sigma_u} \int_0^\infty \exp \left\{ -\frac{(e_i + u_i)^2}{2\sigma_v^2} - \frac{u_i^2}{2\sigma_u^2} \right\} du_i \quad (3.4)$$

a qual não pode ser escrita em forma fechada. Uma aproximação via método de Laplace é apresentada no lema abaixo.

Lema 3.1. *Sejam V e U v.a.'s independentes tal que $V \sim N(0, \sigma_v^2)$ e $U \sim N(0, \sigma_u^2)$, e defina a v.a. $Z = V - |U|$. A aproximação de $f(z | \alpha, \lambda) = \int_{\Re} p_V(z + u)p_U(u)du$ pelo método de Laplace resultara na seguinte densidade para Z*

$$f(z | \sigma_v, \sigma_u) = 2 \frac{1}{\alpha} \phi \left(\frac{z}{\alpha} \right) \Phi \left(\frac{\lambda}{\alpha} z \right) \quad (3.5)$$

onde $\alpha = \sqrt{\sigma_v^2 + \sigma_u^2}$, $\lambda = -\sigma_u/\sigma_v$.

Demonstração: Defina $nh(u) = -\frac{(z+u)^2}{2\sigma_v^2} - \frac{u^2}{2\sigma_u^2}$. Então, a aproximação de Laplace para integral definida em (3.4) é

$$\begin{aligned} f(z | \sigma_v, \sigma_u) &= \frac{1}{2\pi\sigma_v\sigma_u} \int_0^\infty \exp\{nh(u)\} du \\ &\simeq \frac{1}{\sqrt{2\pi}\sigma_v\sigma_u} \exp\{nh(\hat{u}_0)\} \Delta^{-1} \Phi(\Delta \hat{u}_0) \\ &= \frac{2}{\sqrt{2\pi}\sigma_v\sigma_u} \exp \left\{ -\frac{\hat{u}_0^2}{2\sigma_u^2} - \frac{(z + \hat{u}_0)^2}{2\sigma_v^2} \right\} \Delta^{-1} \Phi(\Delta \hat{u}_0), \end{aligned}$$

onde

$$\hat{u}_0 = \max h(u) = -\frac{z\sigma_u^2}{(\sigma_u^2 + \sigma_v^2)} \quad \text{e} \quad \Delta = [-nh''(\hat{u}_0)]^{1/2} = \left(\frac{\sigma_u^2 + \sigma_v^2}{\sigma_u^2\sigma_v^2} \right)^{1/2}$$

Após alguns cálculos obtemos a seguinte função de densidade de probabilidade:

$$\begin{aligned} f(z|\sigma_v, \sigma_u) &= \frac{2}{\sqrt{\sigma_v^2 + \sigma_u^2}} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z^2}{\sqrt{\sigma_v^2 + \sigma_u^2}} \right\} \Phi \left(-\frac{\sigma_u}{\sigma_v \sqrt{\sigma_v^2 + \sigma_u^2}} z \right) \\ &= 2 \frac{1}{\alpha} \phi \left(\frac{z}{\alpha} \right) \Phi \left(\frac{\lambda}{\alpha} z \right) \end{aligned}$$

onde $\alpha = \sqrt{\sigma_v^2 + \sigma_u^2}$ e $\lambda = -\sigma_u/\sigma_v$. \square

A densidade dada em (3.5) obtida através da aproximação de Laplace, é exatamente a densidade da distribuição normal assimétrica com parâmetro de escala α e de assimetria λ , ou seja, $Z \sim SN(0, \alpha^2, \lambda)$. Assim, no modelo de fronteira normal – normal truncada dado em (3.1) a densidade marginal do erro e_i pode ser obtida através do lema 3.1. Esse modelo de fronteira estocástica pode ser visto como um modelo de regressão com erro normal assimétrico. Domínguez-Molina et al. (2004) exploram em detalhes a equivalência entre a distribuição normal assimétrica e o modelo de fronteira estocástica normal – normal truncada.

Função de verossimilhança

Assim, considerando um conjunto de dados *cross-section* com N firmas independentes. Assuma o modelo de fronteira normal – normal truncada: $y_i = x_i\beta + e_i$ para $i = 1, \dots, N$ onde $e_i \sim SN(0, \alpha^2, \lambda)$ com densidade dada em (3.5). A função de verossimilhança para esse modelo é

$$L(\beta, \alpha, \lambda; y) = \prod_{i=1}^n 2 \frac{1}{\alpha} \phi \left(\frac{y_i - x_i\beta}{\alpha} \right) \Phi \left(\lambda \frac{y_i - x_i\beta}{\alpha} \right) \quad (3.6)$$

Após obter a função de verossimilhança para o modelo de fronteira normal – normal truncada, podemos obter uma priori de Jeffreys para este modelo. Na seção 3.2.2, iniciamos mostrando os problemas encontrados na estimação do parâmetro de assimetria pelo método de máxima verossimilhança. Em seguida, apresentamos uma priori de Jeffreys para solucionar este problema.

3.2.2 Procedimento de Inferência

Nesta seção abordamos a inferência sobre modelos de fronteira estocástica normal-normal truncada. Como dito anteriormente, problemas são encontrados na estimação do parâmetro de assimetria desse modelo. Os problemas são tanto numéricos como teóricos. Inicialmente, mostramos alguns problemas na obtenção de estimadores de máxima verossimilhança para o modelo de fronteira estocástica normal – normal truncada. A seguir apresentamos algumas questões relevantes no tratamento de dados utilizando este modelo sob o ponto de vista Bayesiano.

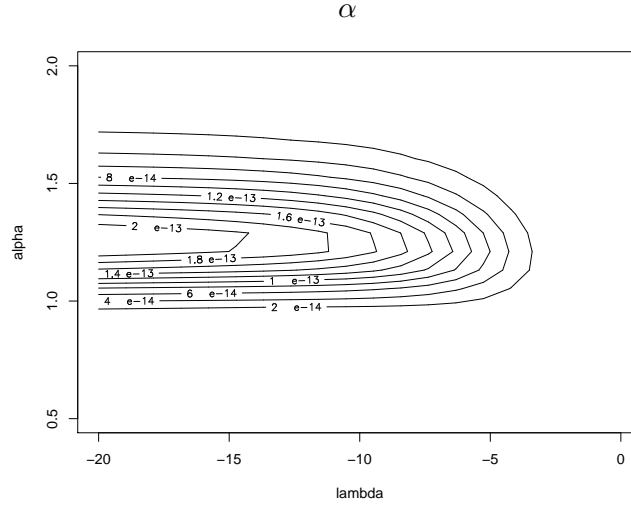
Inferência pelo método de máxima verossimilhança

Apesar das boas propriedades da distribuição normal assimétrica, problemas surgem na etapa de inferência. A função de verossimilhança para o modelo de regressão com erro normal assimétrico tem alguns problemas associados. Um dos parâmetros da normal assimétrica que, no caso do modelo de fronteira, é dado por $-\sigma_u/\sigma_v$ pode causar problemas no procedimento de estimação. Isso ocorre pois, quando σ_v tende para 0, $-\sigma_u/\sigma_v$ tende a $-\infty$. Podemos constatar graficamente que não existe estimador de máxima verossimilhança para o parâmetro $\lambda = -\sigma_u/\sigma_v$. Para isso, foi gerada uma amostra com 30 realizações do modelo normal assimétrico com $\alpha = 1$ e $\lambda = -7$. A Figura 3.2 mostra as curvas de contorno da função de verossimilhança condicional de λ versus α .

Liseo e Loperfido (2006) mostram que há uma probabilidade positiva de o estimador de máxima verossimilhança ser infinito, especialmente para amostras muito pequenas. Para $N = 10$, essa probabilidade chega a aproximadamente 50%. Porém, para $N = 30$ ela cai para 13% para $-\sigma_u/\sigma_v$ igual a -5 (que é um valor bastante extremo). É esperado que para amostras suficientemente grandes essa probabilidade seja aproximadamente nula. Liseo e Loperfido (2006) sugerem a utilização de prioris de referência para solucionar os problemas encontrados na

verossimilhança. A seguir apresentamos uma priori de Jeffreys para o modelo de fronteira estocástica normal – normal truncada.

Figura 3.2: Curvas de contorno da função de verossimilhança condicional de λ versus



Inferência Bayesiana

Sob a abordagem bayesiana, para estimar o vetor de parâmetros $\Theta = (\beta, \alpha, \lambda)$ precisamos definir a distribuição a priori de Θ . Essa distribuição a priori deve refletir a informação disponível a respeito dos parâmetros antes de observarmos o conjunto de dados. Seja $y = (y_1, \dots, y_n)$ uma amostra aleatória (a.a.) de $Y|\Theta$. A inferência Bayesiana será dada pela distribuição a posteriori de Θ , obtida via fórmula de Bayes por $p(\Theta|y) \propto L(\Theta; y)p(\Theta)$.

Uma questão relevante aqui é a escolha da distribuição a priori. Nosso objetivo é calcular a distribuição a priori de Jeffreys para a distribuição normal assimétrica, retemos nossa atenção somente ao parâmetro de assimetria λ com β e α fixos. Em alguns casos, é possível trabalhar com a distribuição a priori de Jeffreys para todos os parâmetros de forma conjunta. Entretanto, tal distribuição a priori conjunta apresenta alguns problemas devido a matriz de informação de Fisher

ser singular para $\lambda \rightarrow 0$ (Azzalini, 1985). Detalhes desta priori e algumas outras discussões podem ser encontradas em Liseo e Loperfido (2006).

Em primeiro lugar, consideramos β, α e λ independentes a priori. A distribuição a priori para β e α serão $\beta \sim N(b, B)$ e $\alpha \sim GI(c/2, d/2)$ com valores para b, B, c e d que garantem que a distribuição a priori seja relativamente vaga. Enquanto que para λ temos a distribuição a priori de Jeffreys para dados independentes e identicamente distribuídos (i.i.d.) da normal assimétrica, como veremos a seguir.

Seja $y = (y_1, \dots, y_N)$ uma a.a. de uma SN com $\mu = 0$ e $\alpha = 1$, o logaritmo da função de verossimilhança é dado por $l(\lambda|y) = N \log(2) + \sum_{i=1}^N \log(\phi(y_i)) + \sum_{i=1}^N \log(\Phi(\lambda y_i))$. Como em Liseo e Loperfido (2004), vamos nos concentrar na seguinte distribuição a priori de Jeffreys

$$p(\lambda) \propto |I(\lambda)|^{1/2} = \left(\int_{-\infty}^{\infty} 2y^2 \phi(y) \frac{\phi^2(\lambda y)}{\Phi(\lambda y)} dy \right)^{1/2}, \quad (3.7)$$

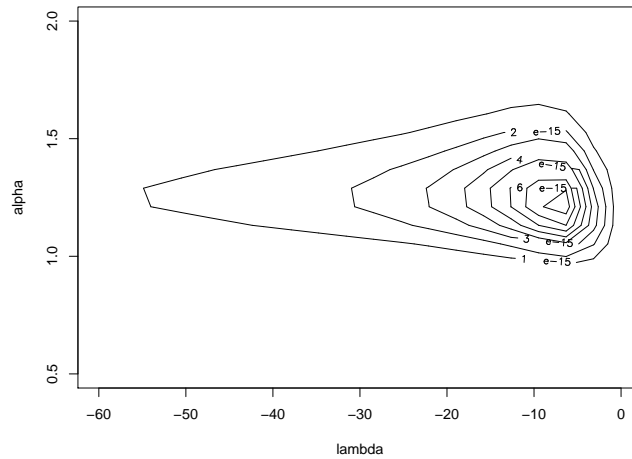
onde $I(\lambda)$ é a matriz de informação de Fisher dada por

$$I(\lambda) = E_{Y|\lambda}(U^2(Y; \lambda)) = E_{Y|\lambda} \left(\frac{\partial}{\partial \lambda} \log(f(y|\lambda)) \right)^2.$$

Liseo e Loperfido (2006) provaram que esta distribuição a priori é simétrica em torno de 0 e possui caudas da ordem $O(\lambda^{-3/2})$, o que implica que é uma distribuição a priori própria. Esta distribuição a priori necessita do cálculo numérico da integral, por exemplo, usando a quadratura Gaussiana, para cada valor de λ . Este cálculo pode ser implementado em algumas linguagens de programação desde que esteja disponível também a função densidade de probabilidade (f.d.p.) e a função distribuição acumulada (f.d.a.) da normal padrão. Em nosso caso, foram usadas a linguagem Fortran e a biblioteca IMSL, foi necessário apenas a implementação da função integrando e chamar uma rotina de quadratura Gaussiana disponível. A Figura 3.3 mostra como o problema encontrado na verossimilhança fica resolvido para o conjunto de dados mostrado anteriormente quando utilizamos a priori

de Jeffreys. Se quisermos nos furtar da responsabilidade da escolha de uma priori subjetiva não seria indicado colocar uma distribuição a priori muito vaga para λ .

Figura 3.3: Curvas de contorno da distribuição a posteriori de λ versus α



O objetivo desta seção era mostrar que a priori de Jeffreys para o caso da fronteira normal – normal truncada é a mesma definida em Liseo e Loperfido (2006) obtida a partir da distribuição normal assimétrica. Esta priori resolve os problemas encontrados na função de verossimilhança. É esperado que estes problemas ocorram com outros modelos de fronteira estocástica em que não se tem uma priori de Jeffreys definida na literatura. Assim, a proposta nas seguintes seções é obter a densidade marginal do erro para as fronteiras normal – gama e t-Student – t-Student truncada e, em seguida, obter uma distribuição a priori de Jeffreys para cada uma destas fronteiras.

3.3 A Fronteira Normal – Gama

Da mesma maneira que no modelo de fronteira normal – normal truncada, Greene (1990) generaliza a fronteira normal – exponencial propondo que a componente de ineficiência u tenha distribuição gama.

3.3.1 O Modelo

O modelo de fronteira estocástica normal – gama é expresso como

$$y_i = x_i\beta + e_i, \quad i = 1, \dots, N \quad \text{com} \quad e_i = v_i - u_i \quad (3.8)$$

onde $v_i \sim N(0, \sigma_v^2)$ e $u_i \sim G(P, \theta)$.

Para obter uma priori de Jeffreys, primeiro necessitamos obter a distribuição marginal do erro composto $e_i = v_i - u_i$. A densidade marginal do erro composto baseia-se na integração sobre $u_i \in \Re$ da densidade conjunta

$$f(u_i, e_i | P, \theta, \sigma_v) = \frac{\theta^P u_i^{P-1}}{\Gamma(P) \sqrt{2\pi} \sigma_v} \exp \left\{ -\frac{(e_i + u_i)^2}{2\sigma_v^2} - \theta u_i \right\} \quad (3.9)$$

Obtendo-se:

$$f(e_i | P, \theta, \sigma_v) = \frac{\theta^P}{\Gamma(P) \sqrt{2\pi} \sigma_v} \int_0^\infty u_i^{P-1} \exp \left\{ -\frac{(e_i + u_i)^2}{2\sigma_v^2} - \theta u_i \right\} du_i \quad (3.10)$$

a qual não pode ser escrita em forma fechada. Greene (1990) reescreve a densidade dada em (3.10) da seguinte forma

$$p(e_i | P, \theta, \sigma_v) = \frac{\theta^P}{\Gamma(P)} \exp(\theta e_i + \sigma_v^2 \theta^2 / 2) h[P - 1, e_i] \Phi(-(e_i + \theta \sigma_v^2))$$

onde $h(r, e_i) = E[Q^r | Q > 0, e_i]$ e $Q | e_i \sim N[-(e_i + \theta \sigma_v^2), \sigma_v^2]$.

Aproximação via Laplace

Uma alternativa é aproximar a densidade do erro composto através do método de Laplace, que será apresentado no lema a seguir.

Lema 3.2. *Sejam V e U v.a.'s independentes tal que $V \sim N(0, \sigma_v^2)$ e $U \sim G(P, \theta)$, e defina a v.a. $Z = V - U$. A aproximação de $f(z | P, \theta, \sigma_v) = \int_{\Re} p_v(z + u) p_u(u) du$ pelo método de Laplace resultará na seguinte densidade para Z*

$$f(z | P, \theta, \sigma_v) = g(z) \Phi \left(\sqrt{\hat{u}_0^2 + \sigma_v^2(P - 1)} \right) \quad (3.11)$$

onde

$$\begin{aligned} g(z) &= \frac{\theta^P \hat{u}_0^P(z)}{\Gamma(P) \sqrt{\hat{u}_0^2(z) + \sigma_v^2(P-1)}} \exp \left\{ -\frac{(z + \hat{u}_0(z))^2}{2\sigma_v^2} - \theta \hat{u}_0(z) \right\} \quad e \\ \hat{u}_0(z) &= -\frac{(z + \theta\sigma_v^2)}{2} - \frac{\sqrt{(z + \theta\sigma_v^2)^2 - 4\sigma_v^2(P-1)}}{2}. \end{aligned}$$

Demonstração: Defina $nh(u) = -\frac{(z+u)^2}{2\sigma_v^2} - \theta u + (P-1)\log(u)$. Então, a aproximação de Laplace para integral definida em (3.10) é

$$\begin{aligned} f(z|P, \theta, \sigma_v) &= \frac{\theta^P}{\Gamma(P) \sqrt{2\pi\sigma_v}} \int_0^\infty \exp\{nh(u)\} du \\ &\simeq \frac{\theta^P}{\Gamma(P)\sigma_v} \exp\{nh(\hat{u}_0)\} \Delta^{-1} \Phi(\Delta \hat{u}_0), \\ &= \frac{\theta^P \hat{u}_0^P}{\Gamma(P) \sqrt{\hat{u}_0^2 + \sigma_v^2(P-1)}} \exp \left\{ -\frac{(z + \hat{u}_0)^2}{2\sigma_v^2} - \theta \hat{u}_0 \right\} \Phi(\Delta \hat{u}_0) \end{aligned}$$

onde

$$\hat{u}_0(z) = -\frac{(z + \theta\sigma_v^2)}{2} - \frac{\sqrt{(z + \theta\sigma_v^2)^2 - 4\sigma_v^2(P-1)}}{2} \quad e \quad \Delta = [-nh''(\hat{u}_0)]^{1/2} = \frac{1}{\sigma_v^2}.$$

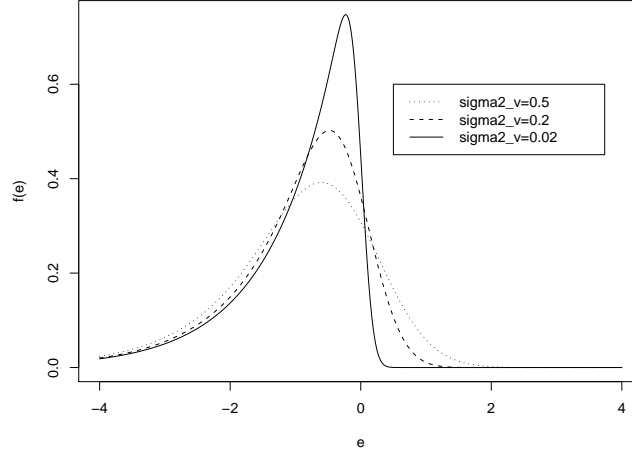
Após alguns cálculos obtemos a seguinte densidade:

$$f(z|P, \theta, \sigma_v) = g(z) \Phi \left(\sqrt{\hat{u}_0(z)^2 + \sigma_v^2(P-1)} \right). \quad (3.12)$$

□

A densidade dada em (3.12), obtida através da aproximação de Laplace, é parecida com a classe de distribuições assimétricas proposta por Azzalini (1985). A figura 3.4 mostra o comportamento desta densidade para diferentes valores de σ_v^2 e com $\theta = 1$ fixo.

Figura 3.4: Distribuição normal - exponencial



Função de verossimilhança

Assim, considerando um conjunto de dados *cross-section*, com N firmas independentes. Assuma o modelo de fronteira normal – gama: $y_i = x_i\beta + e_i$ para $i = 1, \dots, N$ onde e_i tem densidade dada em (3.12). A função de verossimilhança para esse modelo é

$$L(\sigma_v, \sigma_u, \theta; y) = \prod_{i=1}^n 2g(y_i - x_i\beta) \Phi \left(\sqrt{\hat{u}_0(y_i - x_i\beta)^2 + \sigma_v^2(P-1)} \right) \quad (3.13)$$

3.3.2 Procedimento de Inferência

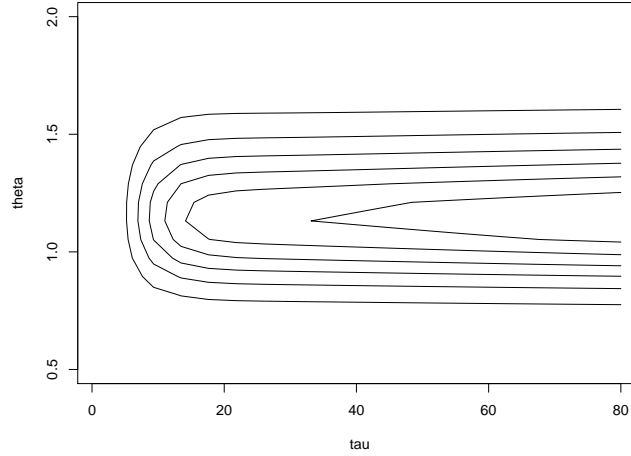
Nesta seção abordamos a inferência sobre modelos de fronteira estocástica normal – gama. Inicialmente, mostra-se o problema na obtenção do estimador de máxima verossimilhança para a fronteira estocástica normal – gama. Em seguida, apresentamos uma distribuição a priori de Jeffreys para solucionar este problema.

Inferência pelo método de máxima verossimilhança

Iniciaremos mostrando graficamente de que pode não existir o estimador de máxima verossimilhança para o parâmetro $\tau_v = 1/\sigma_v$ da componente normal.

A Figura 3.5 mostra os gráficos de contorno da função de verossimilhança condicional de τ_v versus θ para uma amostra com $\theta = 1$, $\sigma_v = 0.14$, $\beta = (1 \ 1)$ e $N = 30$. Como se observa, a curva é ilimitada na direção de θ .

Figura 3.5: Curva de contorno da função de verossimilhança condicional de θ versus τ_v



Neste contexto, um procedimento usual de maximização da verossimilhança não levará a resultados adequados. Essas características da verossimilhança sugerem que devemos procurar uma maneira de calibrar a informação obtida dos dados. A seguir vamos definir a distribuição a priori de Jeffreys para o parâmetro σ_v para o modelo de fronteira normal – gama.

Inferência Bayesiana

Sob a abordagem Bayesiana, estimaremos o vetor de parâmetros $\Omega = (\beta, \sigma_v, P, \theta)$. Mas para isso precisamos definir a distribuição a priori de Ω . Consideramos σ_v , β e θ independentes a priori. A distribuição a priori para β , P e θ serão $\beta \sim N(b, B)$, $P \sim GI(d_0, c_0)$ e $\theta \sim GI(g/2, h/2)$ com valores para b, B, d_0, c_0, g e h que garantem que a distribuição a priori seja relativamente vaga. Enquanto que para σ_v obteremos a distribuição a priori de Jeffreys para dados i.i.d. do modelo de fronteira normal – gama, como veremos a seguir.

Seja $y = (y_1, \dots, y_n)$ uma amostra aleatória do modelo de fronteira normal – gama com $\beta = 0$ e $P = 1$. O logaritmo da função de verossimilhança é dado por $l(\sigma_v, \theta|y) = N \log(2) + \sum_{i=1}^N \log(g(y_i)) + \sum_{i=1}^N \log \Phi \left(-\frac{y_i}{\sigma_v} - \theta \sigma_v \right)$. A distribuição a priori de Jeffreys para o parâmetro σ_v dado θ é dada por

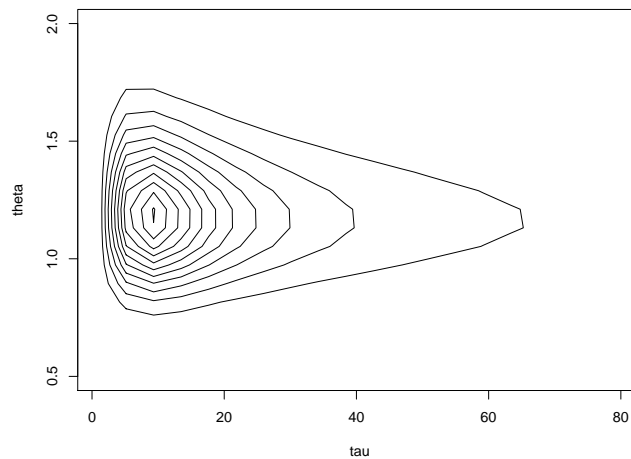
$$p(\sigma_v|\theta) \propto |I(\sigma_v)|^{1/2} = \left(\theta^2 \sigma_v + \int_{-\infty}^{\infty} 2 \left(\frac{y}{\sigma_v^2} - \theta \right) g(y) \frac{\phi^2 \left(-\frac{y}{\sigma_v} - \theta \sigma_v \right)}{\Phi \left(-\frac{y}{\sigma_v} - \theta \sigma_v \right)} dy \right)^{1/2}$$

onde $I(\sigma_v)$ é a matriz de informação de Fisher dada por

$$I(\sigma_v) = E_{Y|\sigma_v}(U^2(Y; \sigma_v)) = E_{Y|\sigma_v} \left(\frac{\partial}{\partial \sigma_v} \log(f(y|\sigma_v, \theta)) \right)^2.$$

Esse cálculo pode ser implementado em algumas linguagens de programação desde que estejam disponíveis também a f.d.a da normal padrão. Em nosso caso, foi usado a rotina de quadratura Gaussiana na linguagem Fortran. A priori de Jeffreys para $\tau_v = 1/\sigma_v$ é $p(\tau_v) \propto I^{1/2}(\sigma_v) |\partial \sigma_v / \partial \tau_v|$. A Figura 3.6 mostra como o problema encontrado na verossimilhança fica resolvido para o conjunto de dados mostrado anteriormente quando utilizamos a priori de Jeffreys.

Figura 3.6: Curva de contorno da distribuição a posteriori de θ versus τ_v



3.4 Fronteira t-Student – t-Student Truncada

Um problema pouco abordado nos modelos de fronteira estocástica é considerar uma fronteira que pode levar em conta a presença de caudas pesadas no erro composto. Por exemplo, se a distribuição do erro simétrico é normal, a análise do modelo pode se tornar sensível a presença de observações discrepantes. O uso da distribuição t-student na componente de erro acomoda a presença dessas observações. Tancredi (2003) propõe um modelo de fronteira estocástica onde o erro de medida tem distribuição t-student e a ineficiência técnica tem distribuição t-Student truncada.

3.4.1 O Modelo

O modelo de fronteira estocástica t-Student – t-Student truncado é dado por

$$y_i = x_i\beta + e_i, \quad i = 1, \dots, N \quad \text{com} \quad e_i = v_i - |u_i| \quad (3.14)$$

onde $v_i \sim St(0, \sigma_v^2, \eta)$ e $u_i \sim St(0, \sigma_u^2, \eta)$.

Para obter uma priori de Jeffreys, primeiro necessitamos obter a distribuição marginal do erro composto $e_i = v_i - u_i$.

Distribuição t-Assimétrica

No contexto das distribuições assimétricas, para obter a distribuição marginal do erro composto precisaremos da proposição que veremos a seguir, cuja demonstração pode ser encontrada em Azzalini e Capitanio (2003).

Proposição 3.2. *Considere que o vetor aleatório (U, V) segue uma distribuição t-Student bivariada padrão com parâmetro de forma η , e defina*

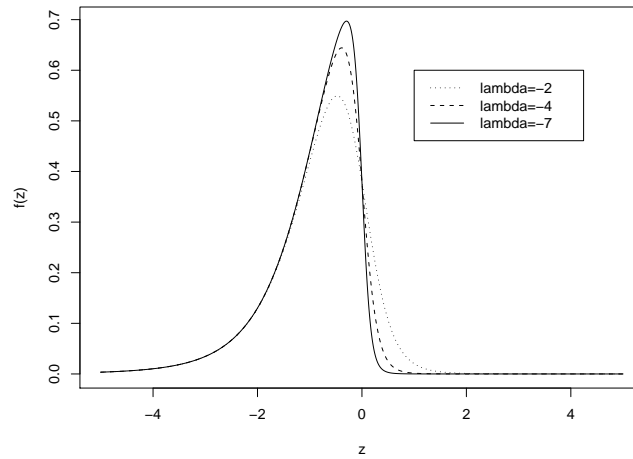
$$Z = \sqrt{1 - \delta^2}V + \delta|U|.$$

Então, a variável aleatória Z tem distribuição t-assimétrica com parâmetros de assimetria λ e de graus de liberdade η , e sua função de densidade é dada por

$$f(z) = 2t_\eta(z) T_{\eta+1} \left[\frac{\delta}{\sqrt{1-\delta^2}} z \left(\frac{\eta+1}{z^2+\eta} \right)^{1/2} \right]. \quad (3.15)$$

Lembre que t_η e $T_{\eta+1}$ denotam a função de densidade e a função de distribuição de variáveis aleatórias t-Student padrão com η e $\eta+1$ graus de liberdade, respectivamente. Usaremos a notação $Z \sim St_\eta(\lambda)$ onde $\lambda = \frac{\delta}{\sqrt{1-\delta^2}}$. É fácil verificar que se $Z \sim St_\eta(\lambda)$ e $Y = \mu + \alpha Z$ então a v.a. Y tem distribuição t-assimétrica com parâmetros de posição μ , de escala α , de assimetria λ e de graus de liberdade η ; para simplificar, denotaremos esta distribuição por $Y \sim St_\eta(\mu, \alpha^2, \lambda)$. A Figura 3.7 mostra o comportamento da distribuição t-assimétrica para diferentes valores do parâmetro de assimetria λ com posição $\mu = 0$, escala $\alpha = 1$ e graus de liberdade $\eta = 3$, fixos. Observe que a medida que o parâmetro λ cresce, o mesmo ocorre com a assimetria.

Figura 3.7: Densidade da t-assimétrica



Para obtermos a densidade marginal do erro composto e_i , basta definir que $e_i = \alpha z_i$ onde $\alpha = \sqrt{\sigma_u^2 + \sigma_v^2}$ e z_i é a t-assimétrica obtida através da proposição 3.2, com $\delta = -\frac{\sigma_u}{\sqrt{\sigma_u^2 + \sigma_v^2}}$. Logo, a densidade do erro composto é exatamente a

densidade da distribuição t-assimétrica com parâmetro de escala α e de assimetria λ , ou seja, $e_i \sim St_\eta(0, \alpha^2, \lambda)$. O modelo de fronteira estocástica pode ser visto como um modelo de regressão com erro t-assimétrico.

Uma característica importante da distribuição t-assimétrica é o fato dela poder ser representada como uma mistura no inverso da escala de uma distribuição normal assimétrica com uma distribuição gama, isto é, $Z = \frac{Z_0}{\sqrt{c}}$, onde $Z_0 \sim SN(0, 1, \lambda)$, e $c \sim G(\eta/2, \eta/2)$, independentes. Para isso precisaremos dos Lemas a seguir, cujas demonstrações podem ser encontradas em Azzalini e Capitanio (2003). Assim, a distribuição t-assimétrica, como pode ser visto, é uma extensão natural da normal assimétrica.

Lema 3.3. *Se $V \sim Gama(\alpha, \beta)$, então para qualquer $a, b \in \Re$*

$$E(\Phi(a\sqrt{V} + b)) = Pr(T \leq a\sqrt{\alpha/\beta}),$$

onde T tem distribuição t-Student não centrada com 2α graus de liberdade e parâmetro de posição $-b$.

Lema 3.4. *Se $X|V = v \sim SN(0, v^{-1}, \lambda)$ e $V \sim Gama(v/2, v/2)$, então X tem f.d.p dada por*

$$f(z) = 2\frac{1}{\alpha}t_\eta\left(\frac{z-\mu}{\alpha}\right)T_{\eta+1}\left[\lambda\frac{z-\mu}{\alpha}\left(\frac{\eta+1}{\left(\frac{z-\mu}{\alpha}\right)^2+\eta}\right)^{1/2}\right].$$

Função de verossimilhança

Assim, considerando um conjunto de dados *cross-section*, com N firmas independentes, assuma o modelo de fronteira t-Student – t-Student truncada: $y_i = x_i\beta + e_i$ para $i = 1, \dots, N$ onde $e_i \sim St_\eta(0, \alpha^2, \lambda)$. A função de verossimilhança para esse modelo é

$$L(\alpha, \mu, \lambda, \eta; y) = \prod_{i=1}^n \frac{1}{\alpha}t_\eta\left(\frac{y_i-\mu}{\alpha}\right)T_{\eta+1}\left[\lambda\frac{y_i-\mu}{\alpha}\left(\frac{\eta+1}{\left(\frac{y_i-\mu}{\alpha}\right)^2+\eta}\right)^{1/2}\right]$$

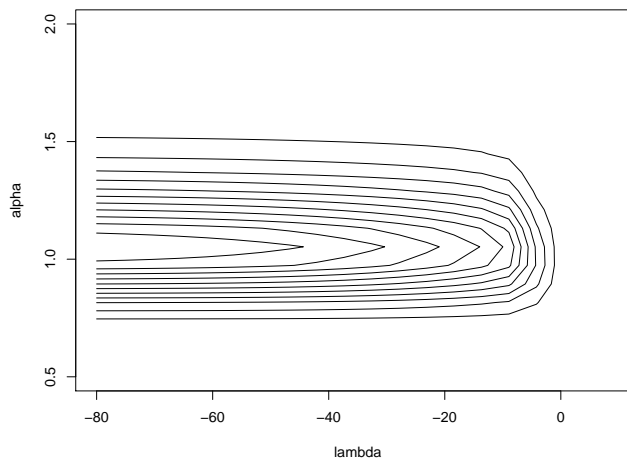
3.4.2 Procedimento de Inferência

Nesta seção abordamos a inferência sobre modelos de fronteira estocástica t-Student – t-Student truncada. Nesta seção, iniciamos mostrando os problemas encontrados na estimação dos parâmetros de assimetria e de grau de liberdade pelo método de máxima verossimilhança. Em seguida, desenvolve-se a priori de Jeffreys para o caso mais simples, o caso $St_\eta(0, 1, \lambda)$.

Inferência pelo método de máxima verossimilhança

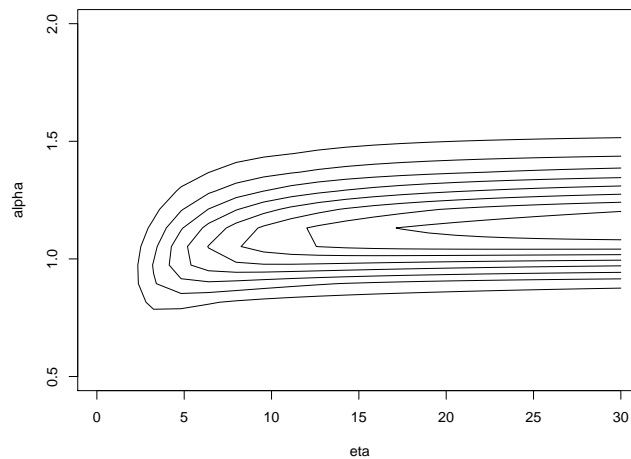
Observamos que a distribuição t-assimétrica é uma extensão natural do caso normal assimétrica. Como a distribuição normal assimétrica apresenta alguns problemas que surgem na etapa da inferência, então espera-se o mesmo na função de verossimilhança para o modelo de regressão com erro t-assimétrico, principalmente no parâmetro de assimetria, $\lambda = -\sigma_u/\sigma_v$. A Figura 3.8 mostra os gráficos de contorno da verossimilhança condicional de λ versus α para uma amostra com $\alpha = 0, 7$, $\lambda = -7$, $\eta = 4$ e $n = 30$. Para este particular conjunto de dados o valor máximo da função de verossimilhança é alcançado quando $\lambda \rightarrow -\infty$.

Figura 3.8: Curvas de contorno da verossimilhança condicional de λ versus α .



Observe que a distribuição t-assimétrica tem como caso particular ($\lambda = 0$) a distribuição t-Student com η graus de liberdade. Como a distribuição t-assimétrica apresenta as mesmas características em relação ao grau de liberdade do seu caso particular, então espera-se que a função de verossimilhança do grau de liberdade condicional aos outros parâmetros, também possa ir para uma assíntota a medida que o grau de liberdade vai para infinito, podendo assim não existir o estimador de máxima verossimilhança. A Figura 3.9 ilustra o problema para o mesmo conjunto de dados gerados, onde é apresentado a função de verossimilhança de η para uma amostra de tamanho 30, com os demais parâmetros fixos: $\mu = 0$, $\alpha = 0,7$ e $\lambda = -7$.

Figura 3.9: Curvas de contorno da verossimilhança condicional de η versus α



Para corrigir a dificuldade mencionada para o modelo de regressão com erro t-Student, Fonseca et al. (2006) utilizaram a priori de Jeffreys, cujos resultados mostraram ser eficientes utilizando a perda absoluta, isto é, tomando-se a mediana a posteriori como estimador pontual para η . A seguir vamos obter a distribuição a priori de Jeffreys de (λ, η) para o modelo de regressão com erro t-assimétrico.

Inferência Bayesiana

Sob a abordagem bayesiana, estimaremos o vetor de parâmetros $(\beta, \alpha, \lambda, \eta)$, mas para isso precisamos definir a distribuição a priori de $(\beta, \alpha, \lambda, \eta)$. Consideramos (λ, η) , β e α independentes a priori. Para η temos a distribuição a priori de Jeffreys para dados independentes e identicamente distribuídos da t-assimétrica, dada em (3.21). Enquanto que $\beta \sim N(b, B)$ e $\alpha \sim GI(c/2, d/2)$ com valores para b, B, c e d que garantem que a distribuição a priori seja relativamente vaga. Para $(\lambda|\eta)$ temos a distribuição a priori de Jeffreys para dados independentes e identicamente distribuídos da t-assimétrica, como veremos a seguir.

Seja $y = (y_1, \dots, y_n)$ uma amostra aleatória de uma t-assimétrica com $\mu = 0$ e $\alpha = 1$. O logaritmo da função de verossimilhança é dado por $l(\lambda, \eta|y) = N \log(2) + \sum_{i=1}^N \log(t_\eta(y_i)) + \sum_{i=1}^N \log \left(T_{\eta+1} \left[\lambda y_i \left(\frac{\eta+1}{y_i^2+\eta} \right)^{1/2} \right] \right)$. A distribuição a priori de Jeffreys para o parâmetro λ condicional a η é dado por:

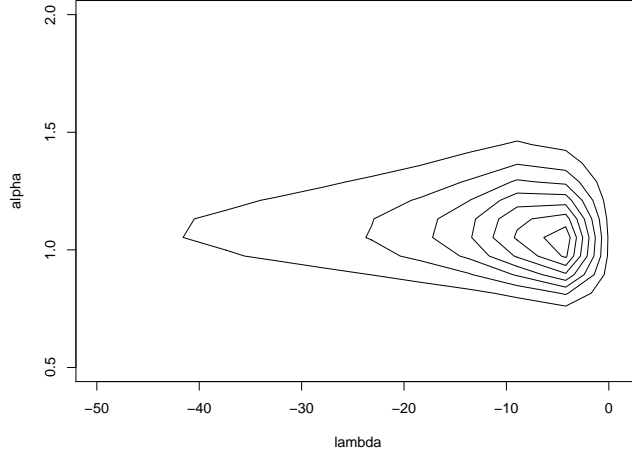
$$p(\lambda|\eta) \propto |I(\lambda)|^{1/2} = \left(\int_{-\infty}^{\infty} 2y^2 \left(\frac{\eta+1}{\eta+y^2} \right) t_\eta(y) \frac{t_{\eta+1}^2 \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right]}{T_{\eta+1} \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right]} dy \right)^{1/2} \quad (3.16)$$

onde $I(\lambda)$ é a matriz de informação de Fisher dada por

$$I(\lambda) = E_{Y|\lambda}(U^2(Y; \lambda)) = E_{Y|\lambda} \left(\frac{\partial}{\partial \lambda} \log(f(y|\lambda, \eta)) \right)^2.$$

Resta a demonstração formal que a distribuição a priori seja própria, logo ainda não é garantindo que esta irá gerar uma distribuição posteriori própria. Graficamente podemos constatar como o uso da priori de Jeffreys realmente corrige o problema encontrado na verossimilhança, ver Figura 3.10, pelo menos para o conjunto de dados mostrado anteriormente.

Figura 3.10: Curva de contorno da distribuição a posteriori de λ versus α .



Agora, precisamos definir a distribuição a priori para o grau de liberdade da t-assimétrica, que é dada por

$$\begin{aligned} p(\eta) \propto |I(\eta)|^{1/2} &= E_{Y|\eta} \left(\frac{\partial}{\partial \eta} \log(f(y|\lambda, \eta)) \right)^2 \\ &= E_{Y|\eta} \left(\frac{\partial \log t_\eta(y)}{\partial \eta} + \frac{\partial \log T_{\eta+1} \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right]}{\partial \eta} \right)^2 \end{aligned}$$

Observe que não é possível determinar uma expressão para

$$\frac{\partial \log T_{\eta+1} \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right]}{\partial \eta} \quad (3.17)$$

$$\text{onde } T_{\eta+1} \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right] = \int_{-\infty}^{\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2}} \frac{\Gamma(\frac{\eta+2}{2})}{\Gamma(\frac{\eta+1}{2}) \sqrt{\pi(\eta+1)}} \left[1 + \frac{z^2}{\eta+1} \right]^{-\frac{\eta+2}{2}} dz$$

Assim, a expressão (3.17) parece ser intratável e deve ser avaliada numericamente. Vamos utilizar a aproximação de Laplace para obter uma expressão para a derivada da função de log-verossimilhança.

Lema 3.5. *Seja $Y \sim ST_\eta(0, 1, \lambda)$ com função de densidade de probabilidade dada por*

$$\begin{aligned} f(y|\lambda, \eta) &= 2t_\eta(y) T_{\eta+1} \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right] \\ &= 2t_\eta(y) \int_{-\infty}^{\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2}} \frac{\Gamma(\frac{\eta+2}{2}) \left[1 + \frac{z^2}{\eta+1} \right]^{-\frac{\eta+2}{2}}}{\Gamma(\frac{\eta+1}{2}) \sqrt{\pi} (\eta+1)^{1/2}} dz \end{aligned} \quad (3.18)$$

Então, aproximando a densidade dada em (3.18) via método de Laplace temos

$$f(y|\lambda, \eta) = 2g(y) \Phi \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right], \quad (3.19)$$

onde $g(y) = \frac{\Gamma(\frac{\eta+2}{2}) t_\eta(y) \sqrt{2}}{\Gamma(\frac{\eta+1}{2}) (\eta+2)^{1/2}}.$

Prova: Defina $nh(z) = -\frac{\eta+2}{2} \log \left[1 + \frac{z^2}{\eta+1} \right]$. Então, a aproximação de Laplace para integral definida em (3.18) é

$$\begin{aligned} f(y|\lambda, \eta) &= \frac{2\Gamma(\frac{\eta+2}{2}) t_\eta(y)}{\Gamma(\frac{\eta+1}{2}) (\eta+1)^{1/2}} \int_{-\infty}^{\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2}} \exp\{nh(z)\} dz \\ &\simeq \frac{2\Gamma(\frac{\eta+2}{2}) t_\eta(y)}{\Gamma(\frac{\eta+1}{2}) (\eta+1)^{1/2}} \exp\{nh(\hat{z}_0)\} (2\pi)^{1/2} \Delta^{-1} \Phi[\Delta(b - \hat{z}_0)], \\ &= \frac{2\Gamma(\frac{\eta+2}{2}) t_\eta\left(\frac{e}{\alpha}\right) \sqrt{2}}{\Gamma(\frac{\eta+1}{2}) (\eta+1)^{1/2}} \left[1 + \frac{\hat{z}_0^2}{\eta+1} \right]^{-\frac{\eta+2}{2}} \Delta^{-1} \Phi[\Delta(b - \hat{z}_0)] \end{aligned}$$

onde

$$b = \lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2}, \quad \hat{z}_0 = \max h(z) = 0 \quad \text{e} \quad \Delta = [-nh''(\hat{z}_0)]^{1/2} = \frac{\sqrt{\eta+2}}{\sqrt{\eta+1}}$$

Após alguns cálculos obtemos a seguinte densidade:

$$f(y|\lambda, \eta) = 2g(y) \Phi \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right]. \quad (3.20)$$

□

Agora, a função log-verossimilhança será $l(\lambda, \eta|y) = N \log(2) + \sum_{i=1}^N \log(g(y_i)) + \sum_{i=1}^N \log \left(\Phi \left[\lambda y_i \left(\frac{\eta+1}{y_i^2+\eta} \right)^{1/2} \right] \right)$ e a distribuição a priori de Jeffreys para o parâmetro de graus de liberdade será

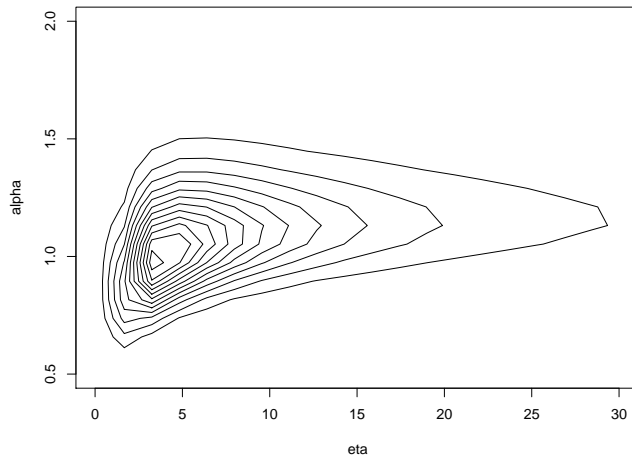
$$p(\eta) \propto E_{Y|\eta} \left(\frac{\partial \log g(y)}{\partial \eta} + \frac{\partial \log \Phi \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right]}{\partial \eta} \right)^2 \quad (3.21)$$

onde

$$\begin{aligned} \frac{\partial \log g(y)}{\partial \eta} &= \frac{d \log \Gamma(\frac{\eta+2}{2})}{d\eta} - \frac{d \log \Gamma(\frac{\eta+1}{2})}{d\eta} - \frac{1}{2(\eta+2)} + \\ &+ \frac{1}{2} \left(\frac{d \log \Gamma(\frac{\eta+1}{2})}{d\eta} - \frac{d \log \Gamma(\frac{\eta}{2})}{d\eta} - \frac{1}{\eta} + \frac{(\eta+1)y^2}{\eta^2(1+y^2/\eta)} - \log(1+y^2/\eta) \right) \\ \frac{\partial \log \Phi \left[\lambda y \left(\frac{\eta+1}{y^2+\eta} \right)^{1/2} \right]}{\partial \eta} &= \frac{\phi \left[\lambda y \left(\frac{\eta+2}{y^2+\eta} \right)^{1/2} \right]}{\Phi \left[\lambda y \left(\frac{\eta+2}{y^2+\eta} \right)^{1/2} \right]} \left(\frac{\lambda y (y^2 - 2)}{2(\eta+2)^{1/2} (y^2 + \eta)^{3/2}} \right). \end{aligned}$$

Resta a demonstração formal que a distribuição a priori seja própria. A Figura 3.11 mostra como o problema encontrado na verossimilhança fica resolvido para o conjunto de dados mostrado anteriormente utilizando a distribuição a priori de Jeffreys para o parâmetro η .

Figura 3.11: Curva de contorno da distribuição a posteriori de η versus α .



3.5 Aplicação com dados artificiais

Nesta seção, são apresentados os resultados de uma aplicação a dados artificiais utilizando a distribuição a priori de Jeffreys desenvolvida na seção anterior. Nossa aplicação são os dados artificiais do modelo de fronteira normal – gama com $P = 1$, ou seja, de fronteira normal – exponencial. Um conjunto de dados artificiais de tamanho 30 (N) foi gerada do modelo de fronteira normal – exponencial com $\theta = 1$, $\beta = (1, 1)$ e para diferentes valores para a precisão $\tau_v = 1/\sigma_v$ ($\tau_v = 3, 5, 7, 9$). Foram utilizadas duas covariáveis: uma assumindo o valor 1 e outra uma geração aleatória da $N(0, 1)$.

3.5.1 Esquema de MCMC

A inferência Bayesiana usando a priori referência será feita considerando duas alternativas de modelagem: a primeira, via modelagem hierárquica (Henze, 1986) e outra via verossimilhança marginal (modelo assimétrico). Em ambos os casos usamos a priori de Jeffreys para σ_v apresentada na seção 3.3.2. As distribuições a priori definidas para os outros parâmetros foram: $\beta \sim N(\mu, \Sigma)$ para $\mu = \underline{0}$ e $\Sigma = \text{diag}(0.001, 0.001)$ e $\theta \sim G(v_0, \omega_0)$ para $v_0 = 0.01$ e $\omega_0 = 0.01$.

A modelagem hierárquica segue a seguinte estrutura:

$$(y_i|u_i, \beta, \sigma_v^2) \sim N(x_i\beta - u_i, \sigma_v^2) \quad (3.22)$$

$$(u_i|\theta) \sim \text{Exp}(\theta) \quad (3.23)$$

A inferência sobre β, θ, σ_v e $u = (u_1, \dots, u_N)$ foi obtida através do algoritmo de MCMC. A partir das equações (3.22) e (3.23) e junto com a distribuição a priori definida anteriormente, obtemos as seguintes distribuições condicionais completas envolvidas no algoritmo de MCMC:

- Distribuição condicional completa de β

$$(\beta|y, \sigma_v, \theta, u) \sim N(b_*, H_*^{-1}),$$

onde $b_* = H_*^{-1}\{H_0 b_0 + \sigma_v^{-2} X'(y + u)\}$ e $H_* = H_0 + \sigma_v^{-2}(X'X)^{-1}$.

- Distribuição condicional completa de θ

$$(\theta|y, \beta, \sigma_v, u) \sim G\left(N + v_0, \sum_{i=1}^N u_i + \omega_0\right).$$

- Distribuição condicional completa de u_i para $i = 1, \dots, N$,

$$(u_i|y, \beta, \sigma_v, \theta) \sim NT^0(m, R),$$

onde $m = x'_i \beta - y_i - \sigma_v^2 \theta$ e $R = \sigma_v^2$.

- Distribuição condicional completa de σ_v

$$p(\sigma_v|\beta, \theta, u) \propto \frac{1}{\sigma_v^N} \exp\left[-\frac{1}{2\sigma_v^2} \sum_{i=1}^N (y_i - x'_i \beta + u_i)^2\right] \pi(\sigma_v),$$

onde $\pi(\sigma_v)$ é a priori de Jeffrey para σ_v .

Note que a distribuição condicional completa de σ_v não tem uma forma conhecida e, portanto, será utilizado o algoritmo de Slice Sampling para a amostragem desta distribuição. O principal interesse na modelagem de fronteira estocástica é a estimação das ineficiências u_i para cada firma i . Na modelagem hierárquica, u_i será gerada a partir da distribuição condicional completa de u_i a cada interação do algoritmo de MCMC. Esta modelagem hierárquica facilita a implementação Bayesiana.

Uma outra alternativa para realizar a inferência Bayesiana é utilizando a aproximação da função de verossimilhança marginal

$$L(y|\beta, \sigma_v, \theta) = \theta \exp\left\{\frac{\theta^2 \sigma_v^2}{2} + \theta(y - x\beta)\right\} \Phi\left[-\frac{(y - x\beta)}{\sigma_v} - \theta\sigma_v\right] \quad (3.24)$$

Neste caso a inferência será realizada apenas para os parâmetros β , θ e σ_v . A partir da função de verossimilhança marginal definida na Eq (3.24) e junto com a distribuição a priori definida anteriormente, obtém-se as seguintes distribuições condicionais completas envolvidas no algoritmo de MCMC:

- Distribuição condicional completa de β

$$p(\beta|y, \sigma_v, \theta) \propto \exp \left\{ \theta(y - x\beta) - \frac{1}{2}\beta'\Sigma^{-1}\beta \right\} \Phi \left[-\frac{(y - x\beta)}{\sigma_v} - \theta\sigma_v \right].$$

- Distribuição condicional completa de θ

$$p(\theta|y, \sigma_v, \beta) \propto \theta^{\nu_0} \exp \left\{ \frac{\theta^2 \sigma_v^2}{2} + \theta(y - x\beta - \omega_0) \right\} \Phi \left[-\frac{(y - x\beta)}{\sigma_v} - \theta\sigma_v \right].$$

- Distribuição condicional completa de σ_v

$$p(\sigma_v|y, \theta, \beta) \propto \theta \exp \left\{ \frac{\theta^2 \sigma_v^2}{2} \right\} \Phi \left[-\frac{(y - x\beta)}{\sigma_v} - \theta\sigma_v \right] \pi(\sigma_v),$$

onde $\pi(\sigma_v)$ é a priori de Jeffrey para σ_v .

Nesta abordagem as distribuições condicionais a posteriori não possuem uma forma analítica fechada devido a função de verossimilhança marginal não ter uma forma conhecida, por isso é necessário utilizar algum método de amostragem.

A estimação das ineficiências u'_i s é feita utilizando a media da distribuição de u_i condicional ao erro composto e_i , $E(u_i|e_i)$

$$E(u_i|e_i) = \sigma_v \left[\frac{\phi(A)}{\Phi(-A)} - A \right] \quad (3.25)$$

onde $A = (e_i + \sigma_v^2 \theta)/\sigma_v$. Observe que $E(u_i|e_i)$ é uma função dos parâmetros do modelo.

Todas as rotinas e os códigos em Fortran envolvido no algoritmo de MCMC neste estudo simulado estão disponíveis sob requisição.

3.5.2 Resultados

A seguir vamos apresentar os resultados da simulação utilizando a priori de Jeffreys para σ_v . Duas alternativas de modelagem foram utilizadas: via verossimilhança marginal ou via modelagem hierárquica.

O esquema MCMC proposto foi utilizado considerando 20.000 iterações onde as primeiras 10.000 foram descartadas. Amostras foram guardadas a cada 5 iterações e, portanto, os resultados são baseados em cadeias de tamanho 2.000. Para verificar a convergência das cadeias foi utilizada estatística \hat{R} (ver Gelman & Rubin, 1992, para mais detalhes) onde valores próximos de 1 indicam convergência das cadeias. As estatísticas \hat{R} ficaram próximas de 1 o que sugere que as cadeias convergiram (ver estatísticas no Apêndice).

Para cada cenário foram calculados, a média a posteriori ($E[\tau_v|y]$) e o desvio padrão a posteriori ($SD[\tau_v|y]$). Além disso, obteve-se o estimador de máxima verossimilhança ($\hat{\tau}_v$) através do pacote Frontier desenvolvido por Coelli (1996). A Tabela 3.1 contém o sumário dessas informações. Observa-se que a média a posteriori está sempre bem próxima do valor verdadeiro de σ_v . Observamos que para $N = 30$, o EMV obtido para σ_v é infinito para $\sigma_v = 5, 7, 9$ o que sugere que há uma probabilidade positiva do estimador de máxima verossimilhança ser infinito, que depende de σ_v e do tamanho do conjunto de dados.

O tempo computacional dos esquemas são dados na Tabela 3.2. Observe que o tempo computacional da modelagem via verossimilhança marginal requereu um tempo mais prolongado. Porém, quando N for relativamente grande, a modelagem hierárquica pode levar certa desvantagem, pois a cada iteração será necessário gerar $(N + 3)$ parâmetros, e isso pode levar um custo computacional maior.

Tabela 3.1: Estimativas a posteriori para o modelo de fronteira normal – exponencial e estimador de máxima verossimilhança.

	verossimilhança marginal		modelagem hierárquica		
Priori	Jeffreys		Jeffreys		
Valor	$E[\tau_v y]$	$SD[\tau_v y]$	$E[\tau_v y]$	$SD[\tau_v y]$	EMV
$\tau_v = 3$	2,678	0,625	3,782	0,767	2,752
$\tau_v = 5$	4,605	1,417	5,079	1,336	4047,850
$\tau_v = 7$	6,445	1,774	7,134	1,453	9933,494
$\tau_v = 9$	9,301	1,348	9,782	1,583	8952,294

Tabela 3.2: Tempo Computacional.

Modelagem	Tempo
verossimilhança marginal e priori de Jeffreys	510 segundos
hierárquica e priori de Jeffreys	330 segundos

3.6 Aplicação a dados Reais

Nesta seção, são apresentados os resultados de uma aplicação com dados reais utilizando as prioris de Jeffreys obtidas. Foram considerados os dados coletados por Christensen e Greene (1976) para 123 companhias de serviço elétrico nos Estados Unidos em 1970. Estes dados também foram analisados por Van den Broeck et al. (1994) e Tancredi (2003).

Christensen e Greene (1976) e Greene (1990) ajustaram uma função de custo sugerido por Nerlove (1963) baseando-se na função de produção Cobb-Douglas, porém generalizaram incluindo o termo quadrático (log do output Q), passando a permitir que o retorno de escala varie com Q . Nestes dados temos três fatores

de produção: trabalho, capital e combustível com os respectivos preços P_l , P_k e P_f ; e com a especificação da função de custo dada por

$$\begin{aligned} y_i = & - \beta_0 - \beta_1 \ln Q_i - \beta_2 \ln^2 Q_i - \beta_3 \ln(P_l/P_f) \\ & - \beta_4 \ln(P_k/P_f) + v_i - u_i, \end{aligned} \quad (3.26)$$

onde $y_i = -\ln(\text{custo da firma } i/P_f)$.

Nesta aplicação serão ajustados apenas os seguintes modelos:

- modelo de fronteira t-student – t-student truncada considerando a priori de Jeffreys para λ e η obtida na seção 3.4.2;
- modelo de fronteira normal – normal truncada considerando a priori de Jeffreys para λ obtida na seção 3.2.2.

Para cada parâmetro foram geradas duas cadeias paralelas (iniciando de valores diferentes) de tamanho 20.000 descartando-se as primeiras 10.000 iterações como período de aquecimento e guardando a cada 5 iterações. Portanto, o tamanho final das amostras é 2.000. A convergência foi verificada através da estatística \hat{R} de Gelman & Rubin em que ficaram próximas de 1, o que sugere que as cadeias convergiram (ver estatísticas no Apêndice). Para se ter uma idéia visual da distribuição da amostra a posteriori, as Figuras 3.12 e 3.13 ilustram os histogramas para os parâmetros do modelo de fronteira normal – normal truncada e t-Student – t-Student truncada, respectivamente. As retas verticais tracejadas e pontilhadas representam a média e o IC 95% a posteriori, respectivamente. Os resultados obtidos são similares àqueles encontrados em Tancredi (2003), em que são ajustados os mesmos modelos, considerando uma abordagem clássica. Observa-se que a densidade a posteriori de η está concentrada em torno de 4,5 indicando que os outliers são bastante prováveis.

Figura 3.12: Histograma para os parâmetros do modelo de fronteira normal – normal truncada utilizando a priori Jeffreys para λ . As retas verticais tracejadas e pontilhadas representam a média e o IC 95% a posteriori, respectivamente.

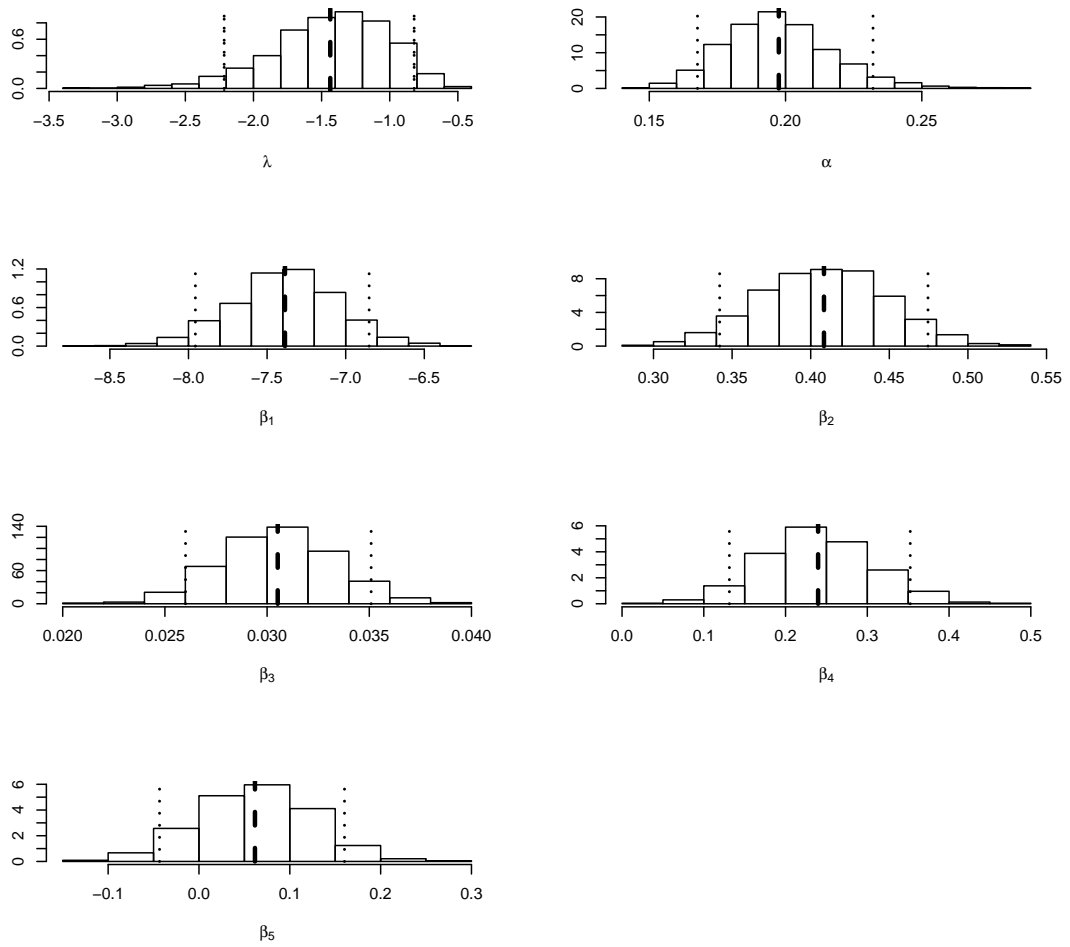
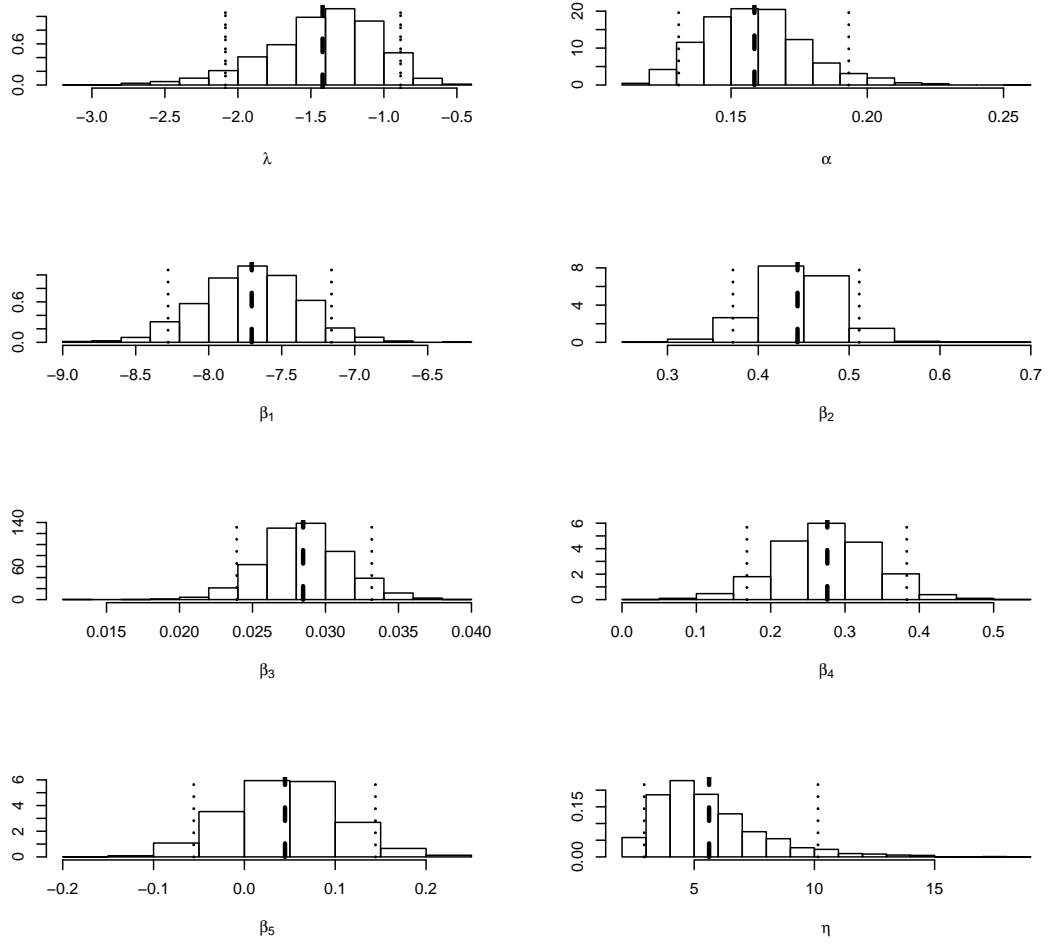


Figura 3.13: Histograma para os parâmetros do modelo de fronteira t-student – t-student truncada utilizando a priori Jeffreys para λ . As retas verticais tracejadas e pontilhadas representam a média e o IC 95% a posteriori, respectivamente.



Finalmente, vamos apresentar os resultados para as eficiências de 5 firmas dentro da amostra. A Tabela 5 compara as estimativas das eficiências $\rho_i = \exp(-u_i)$ para os dois modelos de fronteira analisados. Observamos que no modelo de fronteira t-Student – t-Student truncada as estimativas para ρ_i são um pouco maior que o modelo de fronteira normal – normal truncada, com exceção da primeira firma.

Tabela 3.3: Eficiência para as cinco primeiras firmas.

	Fronteira normal – normal truncada	Fronteira t-student – t-student truncada
ρ_1	0,742	0,691
ρ_2	0,962	0,963
ρ_3	0,909	0,927
ρ_4	0,892	0,898
ρ_5	0,947	0,951

Capítulo 4

Modelo de Fronteira Estocástica Fatorial

Em muitas áreas da economia há o interesse em medir a eficiência técnica de uma determinada firma. Quando temos caso de múltiplos produtos, surge o problema de calcular a eficiência de uma determinada firma. Neste capítulo é formulado um modelo de fronteira estocástica para múltiplos produtos seguindo a abordagem bayesiana. A proposta é reduzir os múltiplos produtos para um único output agregado através da técnica de análise fatorial. Dado este produto agregado, o problema básico de estimar uma fronteira estocástica é essencialmente o mesmo que no caso de um único produto. Neste modelo proposto será utilizado a técnica análise fatorial juntamente com a transformação de Box-Cox (Box e Cox, 1964), onde os produtos transformados podem ser explicados por um fator latente também transformado. A utilização da transformação na família Box-Cox, no contexto do modelo de análise fatorial, permitiu agregar os produtos que são economicamente possíveis. Essa transformação também pode ser particularmente útil quando uma variável resposta não cumpre com os pressupostos de normalidade e/ou homoscedasticidade. A inferência neste modelo proposto será

feita utilizando-se uma abordagem Bayesiana. Finalmente, exemplos com dados artificiais são apresentados para mostrar a aplicabilidade do modelo.

4.1 Introdução

Na teoria neoclássica da firma, a tecnologia de produção é, normalmente, representada por uma função de produção. Esta função sugere uma relação técnica que indica a quantidade máxima do produto que se consegue obter com diferentes combinações dos insumos.

Se a firma produzir diferentes produtos, o processo de produção, que transforma um vetor de insumos x num vetor de produtos y , pode ser representado pela função de transformação dada por

$$h(y, x) = 0, \quad (4.1)$$

onde os produtos são representados pelo vetor $y = (y_1, \dots, y_p)$ e os insumos pelo vetor $x = (x_1, \dots, x_k)$. Se a função de transformação (4.1) é separável, então podemos reescrevê-la da seguinte forma:

$$g(y) = f(x). \quad (4.2)$$

A função $g(y)$ é um agregado dos produtos que unem o "produto agregado" a uma função de produção, $f(x)$. A estimação econométrica do modelo usando uma função de transformação, separável ou não, pode ser difícil. Talvez por esse motivo não existam muitos trabalhos que estimem diretamente a função de transformação. A maioria dos investigadores quando se defrontam com firmas que possuem vários produtos, as funções de custo e de demanda são as mais utilizadas, pelo motivo de facilidades matemáticas, mas isso requer uma quantidade maior de informações sobre os dados. Além das quantidades de insumos e produtos, também são necessários seus preços e custos. Quando existem apenas informações

sobre os insumos e os produtos há a necessidade de se trabalhar com a função de transformação.

Existe uma extensa literatura que implicitamente calcula a função de transformação a fim de avaliar a produtividade de uma firma específica usando uma abordagem não-econométrica. Nesta abordagem é assumida uma função de transformação determinística (sem medida de erro nos dados) e são usadas técnicas de programação linear. Para muitos casos, tal abordagem é indubitavelmente razoável. Porém, em alguns casos poderia ser preferível assumir um erro de medida no modelo e adotar uma abordagem econométrica (ver Koop et al. (1997, 1999) para uma discussão mais detalhada destes assuntos). Outras abordagens econométricas que se encontram na literatura são as de Adams et al.(1999) e de Lothgren (1997), em que assumem a função de transformação separável.

O trabalho mais recente encontrado na literatura utilizando uma abordagem Bayesiana é de Fernández et al. (2000), que usa métodos econométricos assumindo uma função de transformação separável. A forma assumida para $g(y)$ é uma modificação da função CET (*constant elasticity of transformation*). Por analogia com o caso de um único produto, $h(x)$ define a produção máxima que pode ser alcançada com os insumos x , e é chamada de fronteira de produção, e o desvio da fronteira seria interpretada como a ineficiência técnica da firma.

O método mais conhecido para investigar a dependência de um conjunto de variáveis em relação a um número menor de variáveis latentes é chamado análise fatorial. A análise fatorial é uma das técnicas mais usuais do que se convencionou chamar de análise multivariada. A contribuição deste capítulo consiste na utilização da análise fatorial combinada com a transformação de Box-Cox, permitindo obter uma fronteira estocástica quando temos vários produtos. O modelo proposto neste capítulo é uma extensão do modelo de fronteira estocástica para múltiplos produtos proposto por Medrano (2003), em que é utilizada a análise

fatorial sem transformação. A análise fatorial com transformação de Box-Cox possibilita a construção de um fator por meio de combinação entre os produtos transformados e que este fator seja modelado através de uma fronteira estocástica.

Este capítulo está organizado da seguinte forma. Na Seção 4.2 apresentamos a especificação da função de transformação (CET) mais usada na literatura para o caso de múltiplos produtos. Na Seção 4.3 será apresentado o modelo descrito em Fernández et al. (2000), juntamente com o algoritmo MCMC para fazer inferência a posteriori de todos os parâmetros do modelo. Na Seção 4.4, o modelo fatorial Bayesiano será descrito, incluindo as distribuições a priori para os efeitos e seus hiperparâmetros e, também, as distribuições condicionais completas a posteriori. Na Seção 4.5 será incorporado ao modelo fatorial com transformação de Box-Cox, gerando uma nova classe de modelos. Na Seção 4.6 será apresentada a fronteira estocástica fatorial com transformação de Box-Cox. As distribuições a priori, as condicionais completas a posteriori e propostas para o MCMC para todos os parâmetros e hiperparâmetros serão discutidas. Finalizando o capítulo, um estudo de simulação será conduzido comparando as formas alternativas de modelagem.

4.2 Função CET

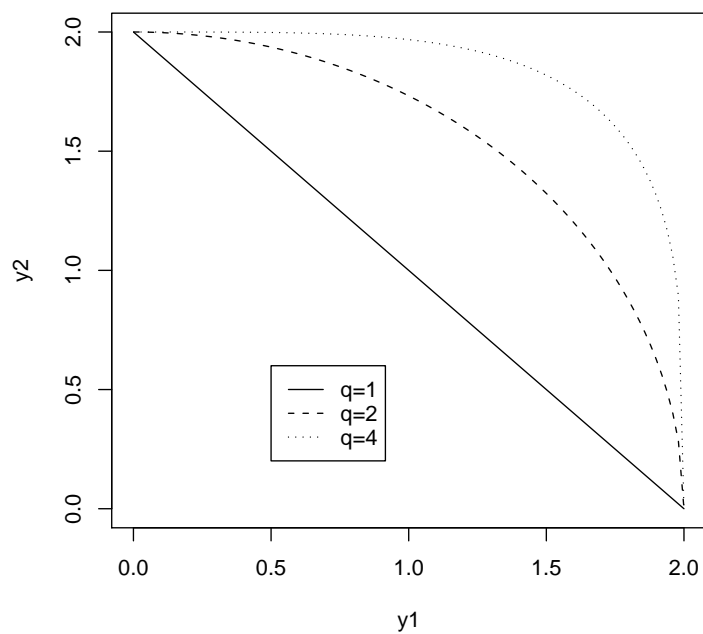
Quando uma firma possui múltiplos produtos, frequentemente esses produtos têm unidade de medida diferentes, o que impossibilita a análise conjunta desses produtos. Para a análise dos dados ser possível, é necessário criar um índice, através de uma transformação ponderada, que garanta uma escala comum para esses produtos. Uma transformação é a função CET (*constant elasticity of transformation*) proposta por Powell e Gruen (1968), a qual é

$$g(y) = \left(\sum_{j=1}^p \alpha_j y_j^q \right)^{1/q}, \quad (4.3)$$

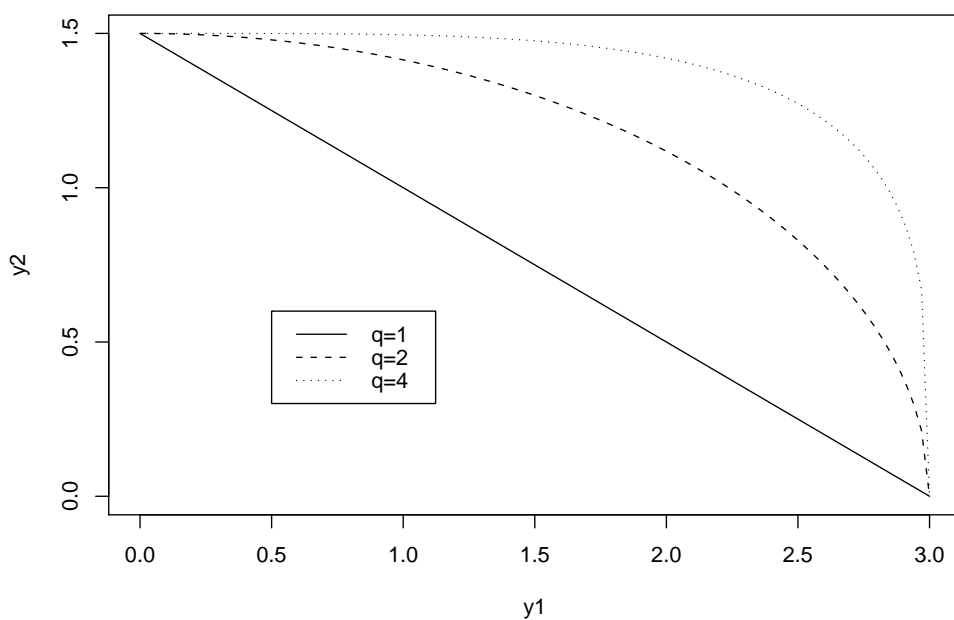
onde $\alpha_j \in (0, 1)$ para $\forall j = 1, \dots, p$ e tal que $\sum_{j=1}^p \alpha_j = 1$ e com $q > 1$. Para valores fixos de $\alpha = (\alpha_1, \dots, \alpha_p)'$, q e $g(y)$, (4.3) define uma superfície de dimensão $(p - 1)$ em \mathbb{R}^p correspondendo todos os vetores dos produtos de dimensão p que são tecnologicamente equivalentes. Em outras palavras, (4.3) representa o conjunto de possibilidades de produção que são tecnologicamente viáveis. Como ilustração, considere $p = 2$ (dois produtos); a Figura 4.1 representa as curvas de possibilidade de produção que são tecnologicamente viáveis para diferentes valores de α_1 e q : (a) $\alpha_1 = 1/2$ e $q = 1, 2, 4$ e (b) $\alpha_1 = 1/3$ e $q = 1, 2, 4$. Note que essas curvas podem ser interpretadas da mesma forma como as curvas isoquantas. Lembre que a curva isoquanta representa as várias combinações dos insumos que podem ser usados para produzir um determinado nível do produto. A curva de possibilidade de produção, por outro lado, representa as várias combinações dos produtos que podem ser produzidas usando um determinado nível do insumo. Na Figura 4.1, observa-se que fixando um valor para y_1 , o valor da contribuição de y_2 para obter $g(y) = 1$ será maior quando o parâmetro q aumenta.

Um elemento importante associado à função de transformação é a elasticidade de transformação. A elasticidade de transformação mede a taxa de mudança na taxa marginal de substituição, ou seja, é uma medida da facilidade com que um produto pode ser substituído por outro. A elasticidade de transformação para a função CET de quaisquer dois produtos é dada por $\varepsilon = 1/(1 - q)$. Impondo-se que $q > 1$ assegura-se a negatividade da elasticidade de transformação. Observe que quando $q \rightarrow 1$ implica que $\varepsilon \rightarrow -\infty$ neste caso os produtos são substitutos perfeitos. Já quando $q \rightarrow \infty$ implica que $\varepsilon \rightarrow 0$ então a substituição não é possível. Assim, quanto menor for o valor de ε melhor será a substituição de um produto pelo outro. Exemplos de curvas de possibilidade de produção para diferentes elasticidades de transformação são apresentados na Figura 4.2.

Figura 4.1: Curvas de possibilidade de produção para dois produtos. Todas as curvas correspondem a $g(y_i) = 1$



(a) $\alpha_1 = \alpha_2 = 1/2$



(b) $\alpha_1 = 1/2, \alpha_2 = 2/3$

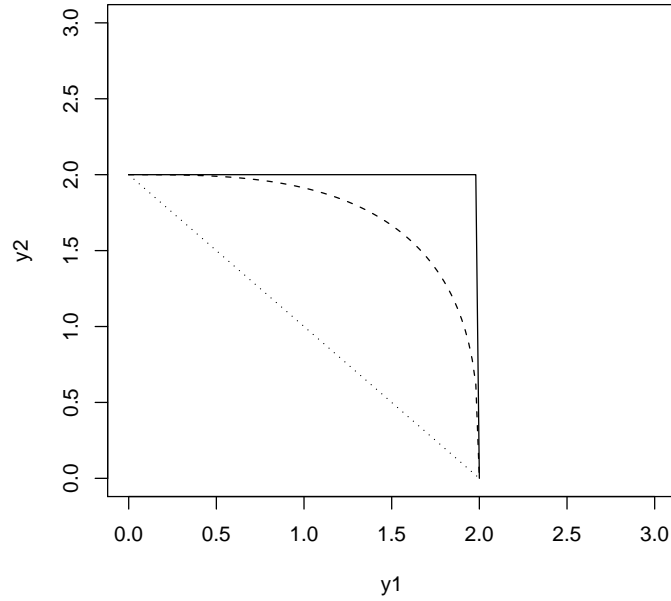


Figura 4.2: Elasticidade de transformação. A linha cheia representa $\varepsilon = 0$, a linha tracejada corresponde $\varepsilon = -1$ e a linha pontilhada representa $\varepsilon \rightarrow -\infty$

4.3 Uma abordagem de fronteira estocástica para múltiplos produtos

Esta seção descreve resumidamente o modelo de fronteira estocástica proposto por Fernández et al. (2000) para o caso de múltiplos produtos. Estes assumem uma função de transformação separável, $g(y) = h(x)$. Em $g(y)$ define o produto agregado e em $h(x)$, uma fronteira de produção. O produto agregado $g(y)$ é obtido por uma função com elasticidade de transformação constante (CET). O objetivo principal desta seção é observar as curvas de possibilidade de produção obtidas através desta função. Assim, nas próximas seções vamos estudar se as possíveis combinações entre os produtos, obtidas pelo modelo aqui proposto, são similares aos obtidos através da função CET, ou seja, se está de acordo com a teoria econômica.

4.3.1 O Modelo

Considere um conjunto de N observações correspondendo aos produtos de N firmas diferentes. O produto da firma i ($i = 1, \dots, N$) é um vetor p -dimensional $y_i = (y_{i,1}, \dots, y_{i,p})' \in \mathbb{R}_+^p$. A função de transformação CET considerada por Fernández *et al.* (2000) é dada por:

$$g(y_i) = \left(\sum_{j=1}^p \alpha_j^q y_{ij}^q \right)^{1/q}, \quad (4.4)$$

onde $\alpha_j \in (0, 1)$ $j = 1, \dots, p$, $\sum_{j=1}^p \alpha_j = 1$ e $q > 1$.

Esta função é uma modificação da função de transformação CET, em que a função dada em (4.4) tem os parâmetros α 's transformados, ou seja, α^q . Esta função define a transformação do vetor de produtos multivariado y_i para a quantidade univariada $g(y_i)$, que pode ser interpretada como um “produto agregado”. Assim, o modelo de fronteira estocástica será definido pelo logaritmo do produto agregado, ou seja, $\log g(y_i)$, tal que

$$\log g(y_i) = x_i' \beta - u_i + v_i \quad v_i \sim N(0, \sigma_v^2), \quad (4.5)$$

onde x_i denota o vetor m -dimensional de insumos e $\beta \in \mathbb{R}^m$ é o vetor de coeficientes de regressão. Finalmente, u_i é o termo de ineficiência que assume uma distribuição exponencial com média $1/\lambda$.

Função de verossimilhança

O modelo dado em (4.5) não é suficiente para definir uma densidade amostral para as observações $y_{i,j}$, pois a função de verossimilhança refere-se apenas ao agregado $\log g(y_i)$, e não aos $y_{i,j}$. Para completar o modelo e usar a inferência Bayesiana, introduzem-se as $(p - 1)$ dimensões restantes considerando a distribuição dos

produtos dentro de cada superfície de equivalência de produção:

$$\eta_{i,j} = \frac{\alpha_j^q y_{i,j}^q}{\sum_{l=1}^p \alpha_l^q y_{i,l}^q}, \quad j = 1, \dots, p, \quad \text{e} \quad \eta_i = (\eta_{i1}, \dots, \eta_{ip})', \quad (4.6)$$

Assumindo independência amostral em (4.6) pode-se definir a função de densidade de probabilidade:

$$\eta_i \sim D_p(s), \quad (4.7)$$

sendo $D_p(s)$ a distribuição Dirichlet com dimensão $(p - 1)$ e parâmetro $s = (s_1, \dots, s_p)' \in \mathbb{R}_+^p$. Nota-se que $0 \leq \eta_{i,j} \leq 1$ e $\sum_{j=1}^p \eta_{i,j} = 1$.

Logo, as equações (4.5)-(4.7) definem a f.d.p de $y = (y_1, \dots, y_N)'$:

$$P(y|\beta, \sigma_v^2, \alpha, q, s, u) = \prod_i p(\log g(y_i)|\beta, u_i, \sigma_v^2) p(\eta_i|s) \prod_{i,j} q^{1-(1/p)} \frac{\eta_{i,j}}{y_{i,j}}, \quad (4.8)$$

sendo $u = (u_1, \dots, u_N)'$, e o último fator da equação (4.8) refere-se ao Jacobiano da transformação.

4.3.2 Inferência Bayesiana e algoritmo MCMC

Nesta seção será apresentado o procedimento de inferência proposto por Fernández *et al.* (2000), baseado no paradigma Bayesiano. Inicialmente, distribuições a priori para todos os hiperparâmetros são apresentadas e, em seguida, é definido o algoritmo MCMC utilizado para obter amostras a posteriori de todos os parâmetros de interesse.

Distribuições a priori

Para a especificação do modelo, nós requeremos as distribuições a prioris para os parâmetros do modelo. Fernández *et al.* (2000) assumem que os parâmetros $(\alpha, q, s, \beta, \lambda, \sigma)$ são independentes a priori e considera as seguintes distribuições a priori: $\alpha \sim D(a)$, $s_j \sim G(b_j, c_j)$, $q \sim \text{Exp}(d)I_{(1,\infty)}(q)$, $\beta \sim N(b_0, H_0^{-1})$, $\sigma^2 \sim GI(n_0/2, a_0/2)$ e $\lambda \sim G(v_0, \omega_0)$.

Inferência a posteriori

A distribuição a posteriori conjunta de $(\beta, \sigma_v^2, \alpha, q, s, \lambda, u)$ é dada por

$$\begin{aligned} p(\beta, \sigma_v^2, \alpha, q, s, \lambda, u|y) &\propto \prod_{i=1}^N p(\log g(y_i)|\beta, u_i, \sigma_v^2) \prod_{i=1}^N p(u_i|\lambda) \prod_{i=1}^N p(\eta_i|s) \\ &\times \prod_{i,j} q^{1-(1/p)} \frac{\eta_{ij}}{y_{ij}} p(\alpha) p(q) p(s) p(\beta) p(\lambda) p(\sigma_v^2) \end{aligned} \quad (4.9)$$

a qual é analiticamente intratável e, portanto, o algoritmo MCMC será utilizado para amostrar os parâmetros de interesse. Neste caso, as distribuições condicionais completas dos parâmetros são obtidas a partir da distribuição conjunta a posteriori (4.9). Neste caso, as ineficiências são amostradas separadamente de uma distribuição normal truncada. Para os parâmetros relacionados à fronteira estocástica, β, σ_v^2 e λ , as distribuições condicionais completas são, respectivamente: distribuições normais (univariadas ou multivariadas) ou distribuições Gama inversa, as amostras são geradas utilizando o amostrador de Gibbs. A amostragem dos parâmetros α, q e s será feita utilizando algum método de amostragem. Detalhes do algoritmo MCMC podem ser vistos em Fernández *et al.* (2000). Todas as distribuições condicionais completas são listadas a seguir

- $p(\alpha|y, \beta, \sigma_v^2, q, s, \lambda, u) \propto \prod_j \alpha_j^{a_j + s_j q^{N-1}} \prod_i \left(\sum_j \alpha_j^q y_{i,j}^q \right)^{-\sum_j s_j} \exp\{-B(\alpha)\}$,
onde $B(\alpha) = \left\{ \frac{1}{2\sigma^2} (\log g(y) - X\beta + u)' (\log g(y) - X\beta + u) \right\}$;
- $p(q|y, \beta, \sigma_v^2, \alpha, s, \lambda, u) \propto q^{N(p-1)} \exp(-dq) \exp\{-A(q)\} I_1(1, \infty)(q)$,
onde $A(q) = \frac{1}{2\sigma^2} (\log g(y) - X\beta + u)' (\log g(y) - X\beta + u) + \sum_{i,j} s_j \log \left(\frac{\sum_{l=1}^p \alpha_l^q y_{i,l}^q}{\alpha_j^q y_{i,j}^q} \right)$;
- Para $j = 1, \dots, p$:
$$p(s_j|y, \beta, \sigma_v^2, \alpha, q, \lambda, u) \propto \frac{\Gamma(\sum_l s_l)^N}{\Gamma(s_j)^N} s_j^{b_j-1} \exp \left[-s_j \left\{ c_j + \sum_i \log \left(\frac{\sum_l \alpha_l^q y_{i,l}^q}{\alpha_j^q y_{i,j}^q} \right) \right\} \right];$$
- $(\beta|y, \sigma_v^2, \alpha, q, s, \lambda, u) \sim N[H_*^{-1}(H_0 b_0 + \sigma_v^{-2} X'(\log g(y) + u)), (H_0 + \sigma^{-2}(X'X)^{-1})^{-1}]$;

- $(\sigma_v^2|y, \beta, \alpha, q, s, \lambda, u) \sim GI\left(\frac{n_0 + N}{2}, \frac{a_0 + (\log g(y) - X\beta + u)'(\log g(y) - x\beta + u)}{2}\right);$
- $(\lambda|y, \beta, \sigma_v^2, \alpha, q, s, u) \sim G\left(N + v_0, \sum_{i=1}^N u_i + \omega_0\right);$
- Para $i = 1, \dots, N$:
 $(u_i|y, \beta, \sigma_v^2, \alpha, q, s) \sim NT(x'_i\beta - \log g(y_i) - \sigma_v^2\lambda, \sigma^2).$

A seguir será descrito o modelo de fronteira estocástica proposto nesta tese para o caso de múltiplos produtos.

4.4 Modelo Fatorial Bayesiano

Nesta seção, o objetivo principal será a estimação da tecnologia de produção e a medição da eficiência relativa a essa tecnologia para o caso de múltiplos produtos. A idéia é propor um modelo alternativo utilizando a técnica de análise fatorial. Inicialmente, vamos descrever o modelo de análise fatorial e, em seguida, generalizar este modelo assumindo que as variáveis transformadas podem ser explicadas por um fator latente também transformado. Será fixado apenas um fator para obter curvas de possibilidade de produção similares a função CET. Assim, no logaritmo desse fator é possível estimar a fronteira estocástica da mesma forma que no caso de um único produto.

4.4.1 Modelo Fatorial Canônico

A análise fatorial tem como objetivo principal investigar a dependência de um vetor aleatório y em termos de um número menor de k variáveis aleatórias não observáveis chamados de fatores. Estes fatores estão relacionados com o vetor y através de um modelo linear. Neste modelo, parte da variabilidade de y é atribuída aos fatores comuns, sendo que o restante da variabilidade é atribuída

às variáveis que não foram incluídas no modelo, ou seja, a uma componente de erro aleatório.

Para definir o modelo, considere y_i um vetor aleatório p -dimensional com média 0_p e matriz de covariância Δ . Uma amostra aleatória de tamanho N é denotada por $\{y_i = (y_{i1}, \dots, y_{ip})', i = 1, \dots, N\}$. Para algum número inteiro positivo $k \ll p$, o modelo k -fatorial é definido por

$$y_i = Lf_i + \epsilon_i \quad \epsilon_i \sim N(0, \Psi), \quad (4.10)$$

onde os fatores f_i , k -dimensionais são independentes com $f_i \sim N(\mathbf{0}, \mathbf{I}_k)$, L é a matriz $(p \times k)$, de carga do fator, $\Psi = \text{diag}(\psi_1, \psi_2, \dots, \psi_p)$ e $i = 1, \dots, N$. Neste modelo, a estrutura de covariância dos dados é dada por $V(y_i|L, \Psi) = LL' + \Psi$, ou seja, a variância de y_i é dividida em duas partes: i) variância explicada pelos fatores comuns (comunalidade) e ii) variância única ou específica.

Função de verossimilhança

Dado f_i , para $i = 1, \dots, N$, o modelo (4.10) pode ser reescrito seguindo uma notação matricial como $y = FL' + \epsilon$, onde $y = (y_1, \dots, y_N)'$ e $F = (f_1, \dots, f_N)'$ são matrizes de dimensão $N \times p$ e $N \times k$, respectivamente. A matriz de erro, ϵ , tem dimensão $N \times p$ e segue uma distribuição normal matriz-variada denotada por $\epsilon \sim N(0, I_N, \Psi)$, assim a função de verossimilhança de (F, L, Ψ) é dada por

$$p(y|\Theta, F, L) = (2\pi)^{-Np/2} |\Psi|^{-N/2} \text{etr} \{ -0.5 \Psi^{-1} (y - FL')' (y - FL') \} \quad (4.11)$$

onde $\text{etr}(X) = \exp(\text{traço}(X))$ para alguma matriz quadrada X .

4.4.2 Identificação do modelo

O modelo k -fatorial deve definir um único modelo livre de problemas de identificação. Nesse sentido, tem-se que o modelo é invariante sob transformações da

forma $f_i^* = Pf_i$ e $L^* = LP'$ onde P é qualquer matriz ortogonal de dimensão $(k \times k)$ (Press 1982, capítulo 10). Na literatura existem muitas formas de identificação do modelo. Uma alternativa é restringir a matriz L tal que seja uma matriz triangular inferior com 1's na diagonal, isto é,

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{2,1} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ l_{k,1} & l_{k,2} & l_{k,3} & \cdots & 1 \\ l_{k+1,1} & l_{k+1,2} & l_{k+1,3} & \cdots & l_{k+1,k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ l_{p,1} & l_{p,2} & l_{p,3} & \cdots & l_{p,k} \end{pmatrix}.$$

Desta forma assegura-se que L seja de posto completo k , e que a matriz de cargas tenha $d = pk - k(k+1)/2$ parâmetros livres. Para mais detalhes e outras soluções, ver Geweke e Zhou (1996), Lopes e West (2004) e Aguilar e West (2001).

4.4.3 Inferência Bayesiana e algoritmo de MCMC

Nesta seção um procedimento de inferência baseado no paradigma Bayesiano é apresentado. Inicialmente, distribuições a priori para todos os hiperparâmetros são apresentadas. Em seguida, o algoritmo MCMC é utilizado para obter amostras a posteriori de todos os parâmetros de interesse.

Distribuições a priori

Para a especificação do modelo, nós requeremos as distribuições a prioris para os parâmetros do modelo (4.10). As distribuições a priori definidas para os parâmetros L e Ψ são: $l_{k,j} \sim N(\mu_0, C_0)$ para $k \neq j$ e $\phi_j \sim GI(v_0/2, s_0/2)$ para $j = 1, \dots, p$.

Inferência a posteriori

A distribuição a posteriori conjunta de (F, L, Ψ) é proporcional a

$$p(F, L, \Psi|y) \propto \prod_{i=1}^N p(y_i|f_i, L, \Psi) \prod_{i=1}^N p(f_i)p(L)p(\Psi), \quad (4.12)$$

na qual é analiticamente intratável e, portanto, o algoritmo MCMC será utilizado para amostrar os parâmetros de interesse. Neste caso, todas as distribuições condicionais completas a posteriori são conhecidas e disponíveis para amostragem. Assim, o algoritmo de simulação utilizado foi o amostrador de Gibbs. As distribuições condicionais completas envolvidas no algoritmo MCMC foram:

- Distribuição condicional completa de L .

- Para as linhas $j = 2, \dots, p$, seja $L_j = (l_{j,1}, \dots, l_{j,j-1})'$, então

$$(L_j|y, F, \Psi) \sim N(\mu_j, C_j),$$

onde $\mu_j = C_j^{-1}(C_0^{-1}\mu_0 1_j + \psi_j^{-2}F'_j y_{1j})$, $C_j = C_0^{-1}1_j + \psi_j^{-2}F'_j F_j$, F_j é a matriz $N \times (j-1)$ contendo as primeiras $(j-1)$ colunas de F e y_{1j} a coluna j de $y = (y_1, \dots, y_N)$.

- Para $j = k+1, \dots, p$, seja $L_j = (l_{j,1}, \dots, l_{j,k})'$, então

$$(L_j|y, F, \Psi) \sim N(\mu_j, C_j),$$

onde $\mu_j = C_j^{-1}(C_0^{-1}\mu_0 1_k + \psi_j^{-2}F' y_{1j})$ e $C_j = C_0^{-1}1_k + \psi_j^{-2}F' F$.

- Distribuição condicional completa de ψ_i . Para $j = 1, \dots, p$,

$$(\psi_j|y, L, f) \sim GI\left(\frac{(v+N)}{2}, \frac{[s + (y_{1j} - FL_j)'(y_{1j} - FL_j)]}{2}\right).$$

- Distribuição condicional completa de f_i . Para $i = 1, \dots, N$,

$$(f_i|y, L, \Psi) \sim N[(I_k + L'\Psi^{-1}L)^{-1}L'\Psi^{-1}y_i, (I_k + L'\Psi^{-1}L)^{-1}].$$

No contexto de fronteira estocástica para múltiplos produtos, considerando apenas um fator, a curva de possibilidade de produção pode ser relacionada a $E(f_i|y, L, \Psi)$ do modelo fatorial. A próxima seção mostrará que a esperança condicional do fator é um caso particular da função CET.

4.4.4 Curva de possibilidade de produção

Com o intuito de ilustrar o comportamento das curvas de possibilidade de produção, suponha que a inferência sobre os fatores possa ser feita com base na média a posteriori de f_i , $E(f_i|y, L, \Psi)$. Como já afirmamos para o problema de fronteira estocástica para múltiplos produtos, vamos nos limitar a um fator ($k = 1$). Assim, a média a posteriori de f_i para um fator será:

$$\begin{aligned} E(f_i|y_i, L, \Psi) &= (1 + L'\Psi^{-1}L)^{-1}L'\Psi^{-1}y_i \\ &= \sum_{j=1}^p \alpha_j y_{ij}, \end{aligned} \quad (4.13)$$

onde $\alpha_1 = \frac{\psi_1^{-1}}{1 + \psi_1^{-1} + \dots + l_p^2 \psi_p^{-1}}$ e $\alpha_j = \frac{l_j \psi_j^{-1}}{1 + \psi_1^{-1} + \dots + l_p^2 \psi_p^{-1}}$ para $j = 2, \dots, p$.

A equação (4.13) mostra a transformação do vetor de produtos p -variado \mathbf{y}_i para um produto agregado de dimensão 1. Observe que a função dada em (4.13) corresponde a um caso particular da função CET, com $q = 1$, ou seja, corresponde ao caso de uma substituição perfeita entre quaisquer dois produtos.

Na próxima seção será apresentada uma extensão do modelo fatorial, onde a transformação de Box-Cox será aplicada tanto sobre o vetor aleatório y quanto ao fator comum f , possibilitando, desta forma, uma maior variedade de combinações entre os elementos do vetor y para obter um determinado nível do fator.

4.5 Modelo Fatorial com Transformação de Box-Cox

Os métodos de simulação iterativos MCMC permitiram um tratamento fácil no modelo fatorial básico. No entanto, este modelo está fundamentado na suposição de normalidade multivariada das observações, y . Uma alternativa para modelar quando não há normalidade é usar dados transformados. As transformações mais utilizadas são as transformações da família de Box-Cox (Box e Cox 1964) onde para uma variável Z , a transformação é dada por:

$$z^{(q)} = \begin{cases} (z^q - 1)/q, & \text{se } q \neq 0, \\ \log(z), & \text{se } q = 0. \end{cases} \quad (4.14)$$

Sob o enfoque Bayesiano, alguns trabalhos recentemente utilizam a transformação de Box-Cox, como por exemplo, em modelos espaciais (ver De Oliveira *et al.*, 1997 e De Oliveira, 2003).

4.5.1 O Modelo

O modelo k -fatorial com transformação de Box-Cox é definido da seguinte forma:

$$y_i^{(q_1)} = L f_i^{(q_2)} + \epsilon, \quad \epsilon \sim N(0, \Psi) \quad (4.15)$$

onde $f_i^{(q_2)} \sim N(m_0, \Delta_k)$, $L = (1 \quad l_2 \quad \dots \quad l_p)'$, $\Psi = \text{diag}(\psi_1, \psi_2, \dots, \psi_p)$. Tal que $y_i^{(q_1)} = (y_{i1}^{(q_1)}, \dots, y_{ip}^{(q_1)})$ e $f_i^{(q_2)} = (f_{i1}^{(q_2)}, \dots, f_{ik}^{(q_2)})$ denotam as transformações de Box-Cox de y_i e f_i respectivamente.

Função de verossimilhança

Dado f_i , para $i = 1, \dots, N$, o modelo (4.15) pode ser reescrito seguindo uma notação matricial como $y^{q_1} = F^{q_2} L' + \epsilon$, onde $y^{q_1} = (y_1^{(q_1)}, \dots, y_N^{(q_1)})'$ e $F^{q_2} = (f_1^{(q_2)}, \dots, f_N^{(q_2)})'$ são matrizes de dimensão $(N \times p)$ e $(N \times k)$, respectivamente.

A matriz de erro ϵ tem dimensão $N \times p$ e segue uma distribuição normal matriz-variada denotada por $\epsilon \sim N(0, I_N, \Psi)$, assim a função de verossimilhança para as observações não transformadas será

$$p(y|f, L, \Psi, q_1, q_2) = (2\pi)^{-Np/2} |\Psi|^{-N/2} \text{etr} \{ -0.5 \Psi^{-1} (y^{q_1} - F^{q_2} L')' (y^{q_1} - F^{q_2} L') \} J_q$$

onde $J_q = \prod_{i,j} y_{ij}^{q_1-1}$ é o jacobiano da transformação.

4.5.2 Inferência Bayesiana e Algoritmo MCMC

Nesta seção é apresentado o procedimento de inferência proposto seguindo uma abordagem completamente Bayesiana. Inicialmente, as distribuições a priori para todos os parâmetros são apresentadas e, em seguida, o algoritmo MCMC será utilizado para obter amostras a posteriori de todos os parâmetros do modelo.

Distribuições a priori

A escolha das distribuições a priori para este modelo requer alguns cuidados, por causa da interpretação dos parâmetros L e Ψ irá depender do valor atribuído ao parâmetro de transformação de Box-Cox.

Considerando o modelo de regressão linear simples com transformação de Box-Cox: $y^{(q)} = x\beta + e$ onde $e \sim N(0, \sigma^2)$, para cada valor de q mudará a escala e locação dos dados transformados. Assim, assumir que β, σ e q sejam independentes a priori pode dar resultados absurdos (Box e Cox 1964). Box e Cox (1964) sugerem a seguinte priori dependente, $p(\beta, \sigma|q) \propto 1/(\sigma J_q^{k/n})$. onde k é a dimensão de β e J_q é o jacobiano da transformação.

Da mesma forma que em Box e Cox (1964), vamos assumir uma priori dependente, assim a distribuição conjunta de (L, Ψ, q_1, q_2) será

$$\begin{aligned} p(L, \Psi, q_1, q_2) &\propto p(L, \Psi|q_1, q_2) p(q_1) p(q_2) \\ &\propto \frac{p(L) p(\Psi)}{J_q^{k/n}} p(q_1) p(q_2). \end{aligned} \quad (4.16)$$

onde as distribuições a priori para os parâmetros L e Ψ são as mesmas definidas para o modelo fatorial apresentado na Seção 4.4.3. Finalmente, vamos assumir uma priori uniforme em (a, b) para cada um dos parâmetros de transformação q_1 e q_2 .

Inferência a posteriori

A distribuição a posteriori conjunta de (F, L, Ψ, q_1, q_2) é proporcional a

$$\begin{aligned} p(F, L, \Psi, q_1, q_2 | y) &\propto \prod_{i=1}^N p(y_i^{(q_1)} | f_i, L, \Psi, q_1, q_2) \prod_{ij} y_{ij}^{(q_1-1)(1+\frac{k}{n})} \\ &\times \prod_{i=1}^N p(f_i^{(q_2)} | m_0, \Delta_k) \prod_i f_i^{q_2-1} p(L) p(\Psi) p(q_1) p(q_2), \end{aligned}$$

a qual é analiticamente intratável e, portanto, o algoritmo MCMC será utilizado para amostrar os parâmetros de interesse. Neste caso, as distribuições condicionais completas a posteriori para parâmetros L e Ψ são conhecidas e disponíveis para amostragem. Apenas as distribuições condicionais completas dos parâmetros q_1 , q_2 e f_i não são conhecidos. Portanto, será utilizado o algoritmo de Slice Sampling para a amostragem destas distribuições. O algoritmo MCMC será desenvolvido a partir das seguintes distribuições condicionais completas:

- Distribuição condicional completa de L

– Para as linhas $j = 2, \dots, p$, seja $L_j = (l_{j,1}, \dots, l_{j,j-1})'$, então

$$(L_j | y, F, \Psi, q) \sim N(\mu_j, C_j),$$

onde $\mu_j = C_j^{-1}(C_0^{-1}\mu_0 1_j + \psi_j^{-2} P_j' z_{1j})$ e $C_j = C_0^{-1} 1_j + \psi_j^{-2} P_j' P_j$. Denota-se por P_j a matriz $N \times (j-1)$ contendo as primeiras $(j-1)$ colunas de $F^{(q_2)} = (f_1^{(q_2)}, \dots, f_N^{(q_2)})$ e z_{1j} é a coluna j de $y^{(q_1)} = (y_1^{(q_1)}, \dots, y_N^{(q_1)})$.

– Para $j = k + 1, \dots, p$, seja $L_j = (l_{j,1}, \dots, l_{j,k})'$, então

$$(L_j|y, F, \Psi, q) \sim N(\mu_j, C_j),$$

onde $\mu_j = C_j^{-1}(C_0^{-1}\mu_0\mathbf{1}_k + \psi_j^{-2}F^{(q_2)'}z_{1j})$ e $C_j = C_0^{-1}\mathbf{1}_k + \psi_j^{-2}F^{(q_2)'}F^{(q_2)}$.

- Distribuição condicional completa de ψ_j . Para $j = 1, \dots, p$,

$$(\psi_j|y, \mathbf{L}, F, q) \sim GI\left(\frac{(v_0 + N)}{2}, \frac{(s_0 + (z_{1j} - F^{(q)}L_j)'(z_{1j} - F^{(q)}L_j))}{2}\right).$$

- Distribuição condicional completa de q_1

$$p(q_1|Y, X, L, f, \Psi) \propto \prod_{i=1}^N \exp\left(\frac{1}{2}(y_i^{(q_1)} - Lf_i^{(q_2)})'\Psi^{-1}(y_i^{(q_1)} - Lf_i^{(q_2)})\right) \prod_{ij} y_{ij}^{(q_1-1)(1+\frac{k}{n})} / (b-a).$$

- Distribuição condicional completa de q_2

$$p(q_2|Y, X, L, f, \Psi) \propto \prod_{i=1}^N \exp\left(\frac{1}{2}(y_i^{(q_1)} - Lf_i^{(q_2)})'\Psi^{-1}(y_i^{(q_1)} - Lf_i^{(q_2)})\right) \prod_i f_i^{q_2-1} / (b-a).$$

- Distribuição condicional completa de f_i . Para $i = 1, \dots, N$,

$$p(f_i|y, L, \Psi, q) \propto \exp\left[\frac{1}{2}(f_i^{(q_2)} - \mu)'\Phi^{-1}(f_i^{(q_2)} - \mu)\right] f_i^{q_2-1},$$

onde $\mu = (\Delta_k^{-1} + L'\Psi^{-1}L)^{-1}(m_0\Delta_k^{-1} + L'\Psi^{-1}y_i^{(q_2)})$ e $\Phi = (\Delta_k^{-1} + L'\Psi^{-1}L)^{-1}$.

4.5.3 Curva de possibilidade de produção

Novamente, com o intuito de ilustrar o comportamento das curvas de possibilidade de produção, vamos examinar a média a posteriori do fator latente. Considerando apenas um único fator ($k = 1$) e $f_i^{(q_2)} \sim N(0, 1)$, a média a posteriori de f_i será:

$$\begin{aligned} E(f_i^{(q_2)}|y, L, \Psi, q_1, q_2) &= (1 + L'\Psi^{-1}L)^{-1}L'\Psi^{-1}y_i^{(q_1)} \\ &= \sum_{j=1}^p \alpha_j y_{ij}^{(q_1)}, \end{aligned}$$

onde $\alpha_1 = \frac{\psi_1^{-1}}{1 + \psi_1^{-1} + \dots + l_p^2 \psi_p^{-1}}$ e $\alpha_j = \frac{l_j \psi_j^{-1}}{1 + \psi_1^{-1} + \dots + l_p^2 \psi_p^{-1}}$ para $j = 2, \dots, p$.

Se $q = q_1 = q_2$ e $\sum_{j=1}^p \alpha_j = 1$ temos que

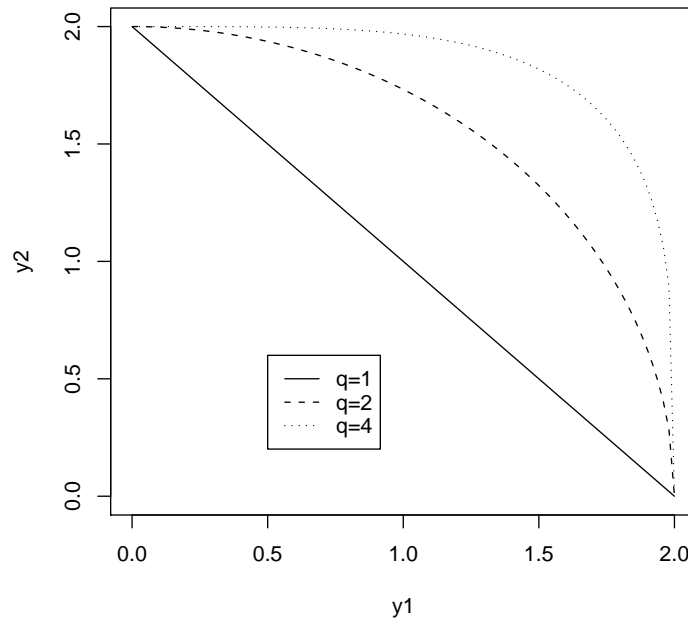
$$\begin{aligned} E(f_i^{(q)}|y, L, \Psi, q) &= \sum_{j=1}^p \alpha_j y_i^{(q)}, \\ E\left(\frac{f_i^q - 1}{q}\right) &= \sum_{j=1}^p \alpha_j \left(\frac{y_i^q - 1}{q}\right), \\ \frac{E(f_i^q) - 1}{q} &= \sum_{j=1}^p \frac{\alpha_j y_i^q}{q} - \sum_{j=1}^p \frac{\alpha_j}{q}, \\ E(f_i^q) &= \sum_{j=1}^p \alpha_j y_i^q. \end{aligned}$$

Então, fazendo $z_i = f_i^q$ e, através da desigualdade de Jensen, temos que

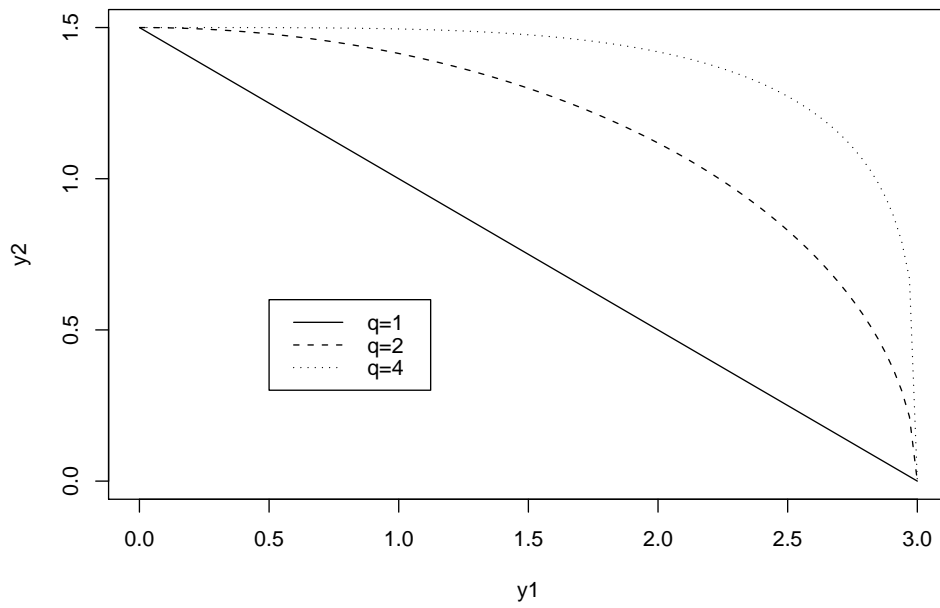
$$E(f_i|y, L, \Psi, q) = E(z_i^{1/q}) \geq (E(z_i))^{1/q} \geq \left(\sum_{j=1}^p \alpha_j y_{i,j}^q\right)^{1/q}.$$

Neste caso, o parâmetro de transformação é mesmo para os produtos e o fator latente, assim a esperança condicional do fator poderá ser similar à função de transformação CET. Considerando $p = 2$, a Figura 4.3 ilustra as curvas de possibilidade de produção para diferentes valores de L , Ψ e fixando $E(f_i|y, L, \Psi, q) = 1$. Observe que a Figura 4.3 apresentam curvas de possibilidade de produção semelhantes às obtidas com a função CET apresentadas na Figura 4.1.

Figura 4.3: Curvas de possibilidade de produção para dois produtos. Todas as curvas correspondem a $E(f_i|y, L, \Psi, q) = 1$.



- (a) $(q = 1) - L = (1 \ 0.5)'$ e $\Psi = \text{diag}(0.5, 0.25)$,
 $(q = 2) - L = (1 \ 2.9)'$ e $\Psi = \text{diag}(0.1, 0.29)$,
 $(q = 4) - L = (1 \ 14.9)'$ e $\Psi = \text{diag}(0.1, 1.49)$.



- (b) $(q = 1) - L = (1 \ 0.8)'$ e $\Psi = \text{diag}(0.4, 0.16)$,
 $(q = 2) - L = (1 \ 1.8)'$ e $\Psi = \text{diag}(0.8, 0.36)$,
 $(q = 4) - L = (1 \ 4.9)'$ e $\Psi = \text{diag}(1.6, 0.49)$.

4.6 Modelo de fronteira estocástica fatorial com transformação Box-Cox

Nesta seção será apresentado um modelo alternativo para obter uma fronteira estocástica para o caso de múltiplos produtos. Uma maneira de tentar solucionar esse problema é utilizar a análise fatorial descrita na Seção 4.5 como uma ferramenta exploratória para a redução do número de variáveis.

4.6.1 O Modelo

Seja N o número de firmas, cada firma com p produtos diferentes. O produto da firma i é dado pelo vetor p -dimensional $y_i = (y_{i,1}, \dots, y_{i,p})' \in \mathbb{R}_+^p$, $i = 1, \dots, N$. O modelo fatorial com transformação de Box-Cox com apenas um fator ($k = 1$) proposto é

$$y_i^{(q)} = Lf_i^{(q)} + \epsilon, \quad \epsilon \sim N(0, \Psi), \quad (4.17)$$

onde $L = (1 \quad l_2 \quad \dots \quad l_p)'$, $\Psi = \text{diag}(\psi_1, \psi_2, \dots, \psi_p)$ e $y_i^{(q)} = (y_{i1}^{(q)}, \dots, y_{ip}^{(q)})$.

Desta forma, a equação (4.17) define a transformação de um vetor de produtos multivariado y_i em uma quantidade univariada f_i . Se interpretarmos o valor de $f_i^{(q)}$ como um produto agregado, então é razoável definir a fronteira estocástica para $\log f_i^{(q)}$, isto é,

$$\log f_i^{(q)} = x_i' \beta - u_i + v_i. \quad (4.18)$$

Como usual, denotamos por x_i o vetor de insumos de dimensão m e por $\beta \in \mathbb{R}^m$ o vetor de coeficientes de regressão. A ineficiência tecnológica é capturada pelo fato das firmas poderem estar operando abaixo da fronteira, obtendo-se assim um vetor de ineficiência u_i , em que esta tem distribuição exponencial com média $1/\lambda$ e v_i tem distribuição normal com média 0 e variância σ_v^2 . A eficiência correspondente à firma i será definida por $\exp(-u_i)$.

Função de verossimilhança

Sejam as observações i.i.d. (y_1, \dots, y_N) e o fator (f_1, \dots, f_N) definido para as N firmas. A função de verossimilhança para $\Omega = (F, L, \Psi, q)$ é dada por

$$p(y|\Omega) = \frac{|\Psi|^{-N/2}}{(2\pi)^{Np/2}} \exp \left\{ -0.5 \sum_{i=1}^N (y_i^{(q)} - f_i^{(q)} L')' \Psi^{-1} (y_i^{(q)} - f_i^{(q)} L') \right\} \prod_{i,j} y_{ij}^{q-1}$$

4.6.2 Inferência Bayesiana e algoritmo MCMC

Nesta seção será apresentado o procedimento de inferência proposto seguindo uma abordagem completamente Bayesiana. Inicialmente, as distribuições a priori para todos os parâmetros são apresentadas. Em seguida, o algoritmo MCMC é utilizado para obter amostras a posteriori de todos os parâmetros do modelo.

Distribuição a priori

Usando a mesma distribuição a priori de (L, Ψ, q) definida para o modelo fatorial com transformação de Box - Cox e considerando uma independência a priori com os parâmetros relacionados a fronteira $(\beta, \sigma_v^2, \lambda)$, obtemos a seguinte distribuição a priori conjunta

$$p(L, \Psi, \beta, \sigma_v^2, \lambda) = \frac{p(L)p(\Psi)p(q)}{J_q^{k/N}} p(\beta)p(\sigma_v^2)p(\lambda), \quad (4.19)$$

onde $L \sim N(\mu_0, C_0)$, $\psi_j \sim GI(v_0/2, s_0/2)$ para $j = 1, \dots, p$, $\beta \sim N(\mu, \Sigma)$, $\sigma_v^2 \sim GI(n_0/2, a_0/2)$, $\lambda \sim G(d_0, \varepsilon_0)$ e $q \sim U(a, b)$. Os hiperparâmetros são escolhidos de forma que as distribuições a priori sejam vagas.

Inferência a posteriori

A distribuição a posteriori conjunta de $(F, L, \Psi, q, u, \beta, \sigma_v^2, \lambda)$ é dada por

$$p(F, L, \Psi, q, u, \beta, \sigma_v^2, \lambda|y) \propto \prod_{i=1}^N p(y_i^{(q)} | f_i, L, \Psi, q_1, q_2) \prod_{i=1}^N \prod_{j=1}^p y_{ij}^{(q-1)(1+\frac{k}{n})}$$

$$\begin{aligned}
& \times p(L)p(\Psi)p(q) \prod_{i=1}^N p(\log f_{it}^{(q)} | u_i, \beta, \sigma_v^2) \prod_{i=1}^N \frac{f_{it}^{q-1}}{f_{it}^{(q)}} \\
& \times \prod_{i=1}^N p(u_i | \lambda) p(\beta) p(\sigma_v) p(\lambda) \tag{4.20}
\end{aligned}$$

a qual é analiticamente intratável e, portanto, o algoritmo MCMC será utilizado para amostrar os parâmetros de interesse. Neste caso, as distribuições condicionais completas a posteriori para parâmetros $L, \Psi, \beta, \sigma_v^2, \lambda$ e u_i são conhecidas e disponíveis para amostragem. Apenas as distribuições condicionais completas dos parâmetros q e f_i não são conhecidas. Portanto, será utilizado o algoritmo de Slice Sampling para a amostragem destas distribuições. As distribuições condicionais de L e Ψ são as mesmas apresentadas na subseção 4.5.2. Para os demais parâmetros, $\beta, \sigma_v^2, \lambda, q, f_i$ e u_i , as distribuições condicionais completas são:

- Distribuição condicional completa de β

$$(\beta | \sigma_v^2, \lambda, u) \sim N(b_*, H_*^{-1}),$$

onde $b_* = H_*^{-1}\{H_0 b_0 + \sigma_v^{-2} X'(\delta + u)\}$ e $H_* = H_0 + \sigma_v^{-2}(X'X)^{-1}$.

- Distribuição condicional completa de σ^2

$$(\sigma_v^2 | \beta, \lambda, u) \sim GI(c_*, d_*),$$

onde $c_* = \frac{(N + N_0)}{2}$ e $d_* = \frac{[(\delta + u - X\beta)'(\delta + u - X\beta) + a_0]}{2}$.

- Distribuição condicional completa de λ

$$(\lambda | \beta, \sigma_v^2, u) \sim G\left(N + v_0, \sum_{i=1}^N u_i + \omega_0\right).$$

- Distribuição condicional completa de q

$$p(q | Y, L, f, \Psi, \beta, \sigma_v^2, u) \propto \prod_{i=1}^N \exp\left(\frac{1}{2}(y_i^{(q)} - Lf_i^{(q)})' \Psi^{-1}(y_i^{(q)} - Lf_i^{(q)})\right)$$

$$\begin{aligned} & \times \exp \left[-\frac{1}{2\sigma^2} (\log f_i^{(q)} - x_i' \beta + u_i)^2 \right] / f_i^{(q)} \\ & \times \prod_{i=1}^N f_i^{q-1} \prod_j^p y_{ij}^{(q-1)(1+\frac{k}{n})} \end{aligned}$$

- Distribuição condicional completa de u_i

$$(u_i | \beta, \sigma_v^2, \lambda) \sim NT(m, R^2),$$

isto é, uma distribuição normal truncada com densidade $2\phi\left(\frac{u_i - m}{R}\right) I\{u_i \geq 0\}$ onde $m = x_i' \beta - \delta_i - \sigma_v^2 \lambda$ e $R^2 = \sigma_v^2$.

- Distribuição condicional completa de f_i . Para $i = 1, \dots, N$

$$\begin{aligned} p(f_i | y, L, \Psi, q, \beta, \sigma_v^2, \lambda, u) & \propto \exp \left(\frac{1}{2} (y_i^{(q)} - L f_i^{(q)})' \Psi^{-1} (y_i^{(q)} - L f_i^{(q)}) \right) f_i^{q-1} \\ & \times \exp \left[-\frac{1}{2\sigma_v^2} (\log f_i^{(q)} - x_i' \beta + u_i)^2 \right] / f_i^{(q)} \end{aligned}$$

A seguir serão apresentados três exemplos com dados simulados e uma aplicação com dados de produção em pesquisas agrícolas da Embrapa.

4.7 Estudo com dados artificiais

Nesta seção são apresentados quatro exemplos com dados artificiais. No primeiro, o modelo fatorial com transformação de Box-Cox será aplicado a dados simulados. A segunda simulação tem por objetivo comparar a esperança a posteriori dos fatores com a função de transformação CET. No terceiro exemplo será realizado um estudo simulado para mostrar a eficiência do modelo de fronteira estocástica fatorial. Na última simulação serão ajustados o modelo de fronteira estocástica fatorial proposto nesta tese e o modelo proposto por Fernández *et al.* (2000), a um conjunto de dados gerado a partir do modelo proposto por Fernández *et al.* (2000). Todas as rotinas e os códigos em Fortran envolvido no algoritmo de MCMC estão disponíveis sob requisição.

4.7.1 Exemplo 1: Análise Fatorial com transformação Box-Cox

Este primeiro estudo simulado considera um modelo fatorial junto a transformação de Box-Cox, com apenas $k = 1$ fator, para um problema com $p = 3$ variáveis com $N = 30$ observações cada uma. Em cada uma das variáveis simuladas, as N observações foram amostradas de um modelo k -fatorial com transformação de Box-Cox com os seguintes parâmetros: $\Psi = \text{diag}(0.05, 0.05, 0.05)$, $L = (1, 0.9, 0.9)'$ e $q = 2.15$.

A inferência será feita com base na amostra da distribuição a posteriori gerada usando o algoritmo MCMC. Foi gerada uma cadeia de tamanho 20.000. A amostra a posteriori foi obtida utilizando um *burn-in* de 10.000. Amostras foram guardadas a cada 5 iterações e, portanto, os resultados estão baseados em cadeias de tamanho 2.000. Foram usadas as seguintes distribuições a priori: $l_{j,1} \sim N(0, 1)$, $\phi_j \sim GI(0.01, 0.01)$ para $j = 1, \dots, 3$ e $q \sim U(0, 10)$.

A verificação da convergência foi verificada através do critério de convergência de Gelman & Rubin, que ficaram próximas a 1 o que sugere que as cadeias convergiram (ver estatísticas no Apêndice). A Figura 4.4 apresenta os histogramas da amostra da distribuição a posteriori de L , ψ e q . Note que as inferências foram satisfatórias, ou seja, o valor estimado está próximo do valor utilizado para gerar os dados.

A Figura 4.5 mostra o gráfico de dispersão da média a posteriori do fator latente versus o valor verdadeiro do fator, f . Nota-se que os valores estimados estão próximos dos valores reais. Em conclusão, o procedimento de inferência se mostrou eficiente na estimação dos parâmetros do modelo e dos fatores latentes.

Figura 4.4: Histograma das distribuições a posteriori dos parâmetros - Modelo Fatorial usando a transformação de Box-Cox. As retas verticais tracejada e pontilhada representam a média e o IC 95% respectivamente.

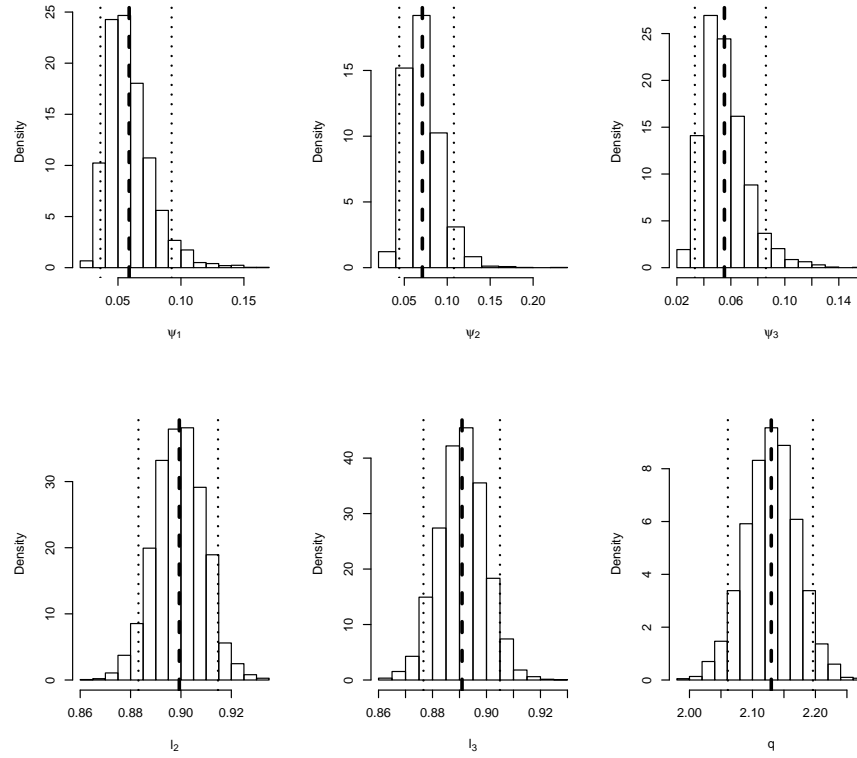
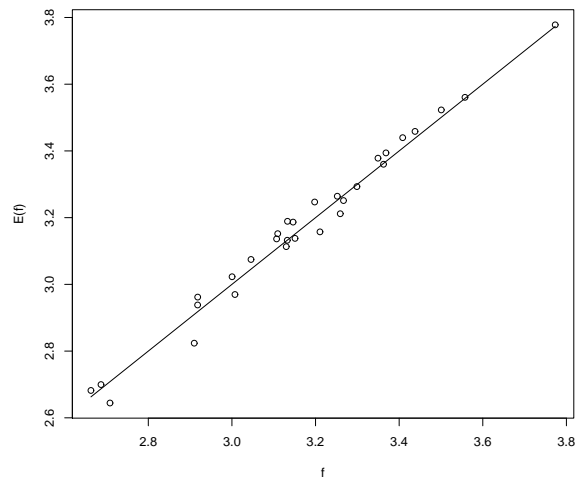


Figura 4.5: Gráfico de dispersão de $E(f|resto)$ versus o valor real de f



4.7.2 Exemplo 2: Estudo da $E(f|y, L, \Psi, q)$

A segunda simulação tem como objetivo comparar a esperança a posteriori dos fatores f_i com a função de transformação CET, sem levar em consideração a fronteira estocástica. Foram gerados dois produtos através da função CET, seguindo a proposta da Fernández *et al.* (2000), no qual o peso de cada produto é definido da seguinte forma

$$\eta_{i,j} = \frac{\alpha_j^q y_{i,j}^q}{\sum_{l=1}^2 \alpha_l^q y_{i,l}^q} = \frac{\alpha_j^q y_{i,j}^q}{g(y_{i,j})^q}, \quad j = 1, 2 \quad e \quad i = 1, \dots, 100. \quad (4.21)$$

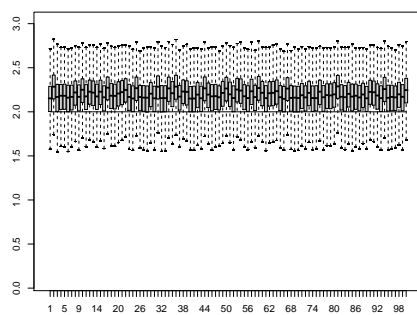
Da equação (4.21) obtemos cada produto da seguinte forma

$$y_{i,j} = \frac{\eta_{i,j}^{1/q} g(y_{i,j})}{\alpha_j}, \quad j = 1, 2 \quad e \quad i = 1, \dots, 100. \quad (4.22)$$

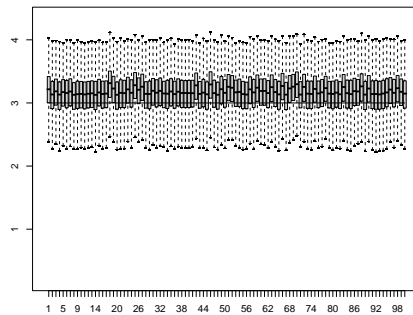
Foram simulados dados com 4 possíveis combinações entre os produtos: $g(y) = 2, 3, 4, 5$. E para cada $g(y)$ foram considerados 4 possíveis transformações de Box-Cox com $q = 2, 3, 4, 5$. Em cada um destes 16 casos foram geradas $N = 100$ valores de $\eta_i = (\eta_{i1}, \eta_{i2})$ por meio da equação (4.22) foram obtidas as observações y_{i1} e y_{i2} para $i = 1, \dots, 100$. Em todos os casos foi definido que $\alpha_1 = \alpha_2 = 1/2$

Em cada amostra simulada foi ajustado o modelo fatorial com transformação de Box-Cox gerando uma cadeia de tamanho 10.000. Nas Figuras 4.6 e 4.7 são apresentados os Box-plots das amostras de $E(f|y, L, \Psi, q)$. Nota-se que as médias a posteriori de f_i ficaram muito próximas dos valores simulados de $g(y)$, obtido através da função CET. Observa-se também uma maior variabilidade a medida que $g(y)$ aumenta por exemplo quando $g(y) = 5$. Este comportamento se reflete para todos os valores de q . Apesar de uma grande variabilidade de $g(y)$ em alguns casos, podemos ter uma idéia que o modelo fatorial com transformação de Box-Cox está combinando os produtos de uma forma similar a função CET.

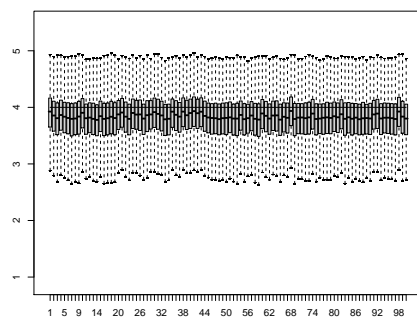
Figura 4.6: Box-plot das amostras de $E(f|y, L, \Psi, q)$. A linha cheia horizontal representa o valor verdadeiro de $g(y)$.



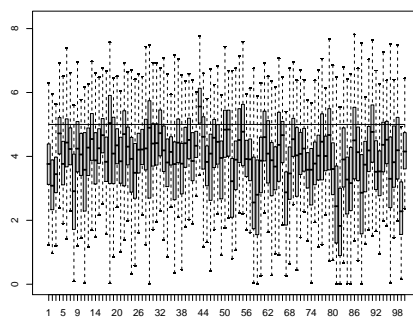
(a) $q = 2$ e $g(y) = 2$



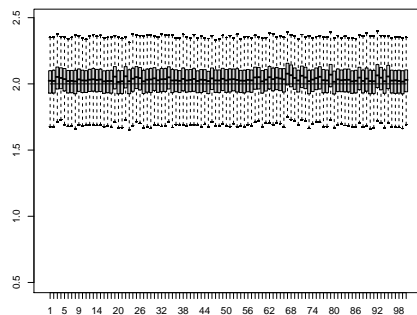
(b) $q = 2$ e $g(y) = 3$



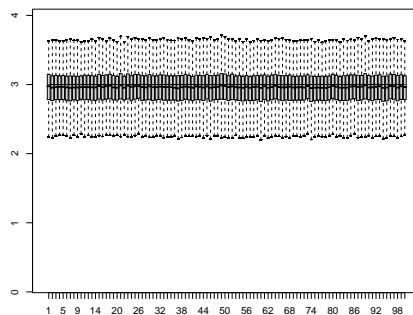
(c) $q = 2$ e $g(y) = 4$



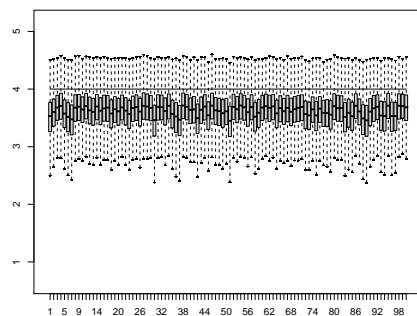
(d) $q = 2$ e $g(y) = 5$



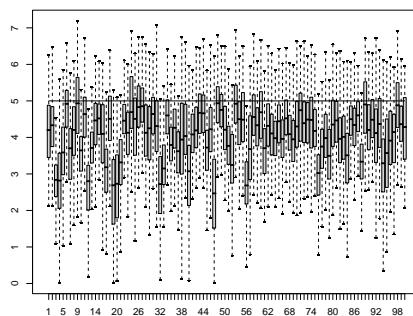
(e) $q = 3$ e $g(y) = 2$



(f) $q = 3$ e $g(y) = 3$

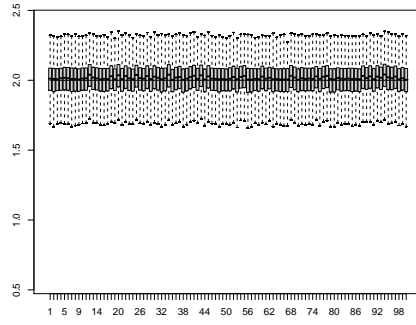


(g) $q = 3$ e $g(y) = 4$

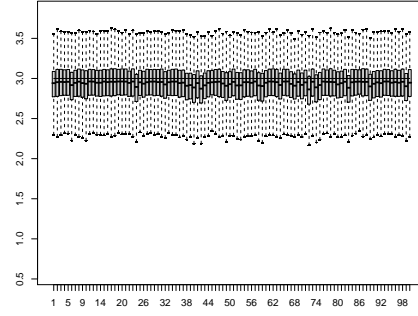


(h) $q = 3$ e $g(y) = 5$

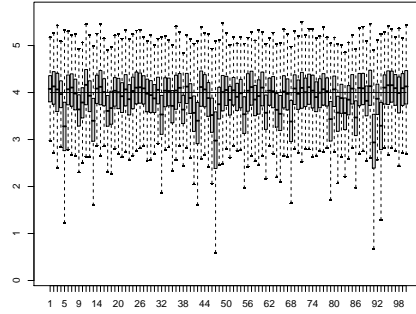
Figura 4.7: Box-plot das amostras de $E(f|y, L, \Psi, q)$. A linha cheia horizontal representa o valor verdadeiro de $g(y)$.



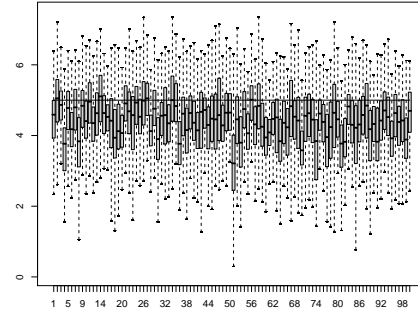
(a) $q = 4$ e $g(y) = 2$



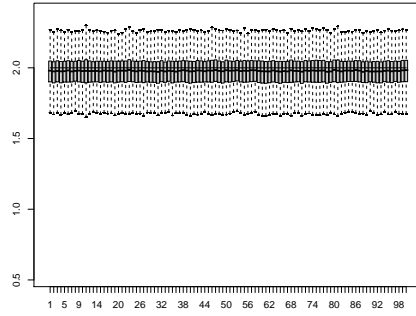
(b) $q = 4$ e $g(y) = 3$



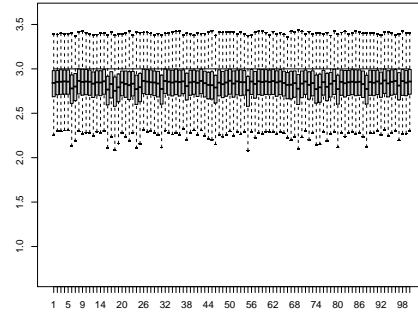
(c) $q = 4$ e $g(y) = 4$



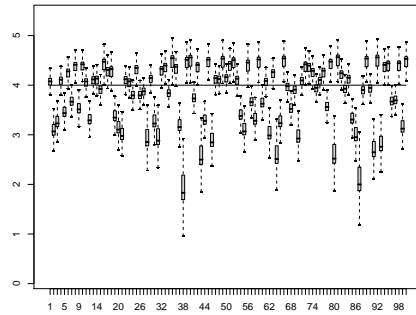
(d) $q = 4$ e $g(y) = 5$



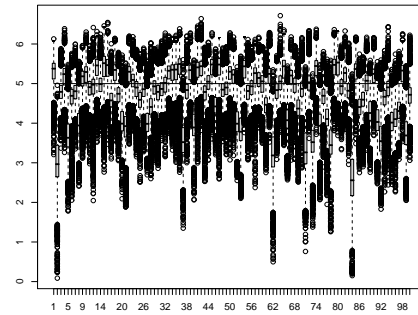
(e) $q = 5$ e $g(y) = 2$



(f) $q = 5$ e $g(y) = 3$



(g) $q = 5$ e $g(y) = 4$



(h) $q = 5$ e $g(y) = 5$

4.7.3 Exemplo 3: Fronteira estocástica fatorial com transformação Box-Cox

O terceiro exemplo segue o padrão da primeira simulação, usando o modelo de fronteira estocástica fatorial com transformação de Box-Cox para $p = 3$ produtos e $N = 30$ firmas. O verdadeiro modelo neste caso tem os seguintes parâmetros: $\Psi = \text{diag}(0.05, 0.05, 0.05)$, $L = (1, 0.9, 0.9)'$, $q = 2.2$, $\beta = (0.5, 0.5)$, $\sigma = 0.5$ e $\lambda = 5$.

O esquema MCMC proposto foi utilizado considerando 20.000, onde as primeiras 10.000 foram descartadas. Amostras foram guardadas a cada 5 iterações e, portanto, os resultados estão baseados em cadeias de tamanho 2.000. A verificação da convergência foi verificada através do critério de convergência de Gelman & Rubin, que ficaram próximas a 1 o que sugere que as cadeias convergiram (ver estatísticas no Apêndice). Foram usadas as seguintes distribuições a priori: $l_{j,1} \sim N(0; 1)$, $\phi_j \sim GI(0.01, 0.01)$ $j = 1, \dots, 3$, $q \sim U(0, 10)$, $\beta \sim NM(\mathbf{0}, \text{diag}(0.001, 0.001))$, $\sigma^2 \sim GI(0.01, 0.01)$.

Na Figura 4.8 são apresentados os histogramas das distribuições a posteriori e os intervalos de credibilidade de 95% de todos os parâmetros do modelo. Claramente, observa-se que todas as densidades estão concentradas em torno dos valores usados na geração dos dados.

Nas Figuras 4.9 e 4.10 são apresentados os gráficos de dispersão da média a posteriori do fator comum f e da medida de eficiência $\rho = \exp(-u)$ versus seus valores verdadeiros, respectivamente, obtido na geração dos dados. Observa-se uma relação linear na Figura 4.9, ou seja, as estimativas do fator ficaram próximas aos valores verdadeiros. Já no gráfico de dispersão da eficiência, nota-se também uma relação linear porém, observa-se que as menores estimativas a posteriori são maiores do que seu valor verdadeiro e, por outro lado, as maiores estimativas da eficiência são menores do que seus respectivos valores verdadeiros.

Note que colocando os ρ_i em ordem para $i = 1, \dots, N$, o rank associado da j^{th} firma é a posição ocupada por ρ_j na ordenação. Assim, analisando através da ordenação das firmas, observe que apesar da relação obtida entre a $E(\rho|resto)$ e o valor verdadeiro de ρ , a ordenação das firmas obtida através das estimativas a posteriori é similar à ordenação obtida na geração dos dados.

Figura 4.8: Histograma da distribuições a posteriori dos parâmetros do modelo - Modelo Fronteira Fatorial com transformação Box-Cox. As retas verticais tracejada e pontilhada representam a média e o IC 95% respectivamente.

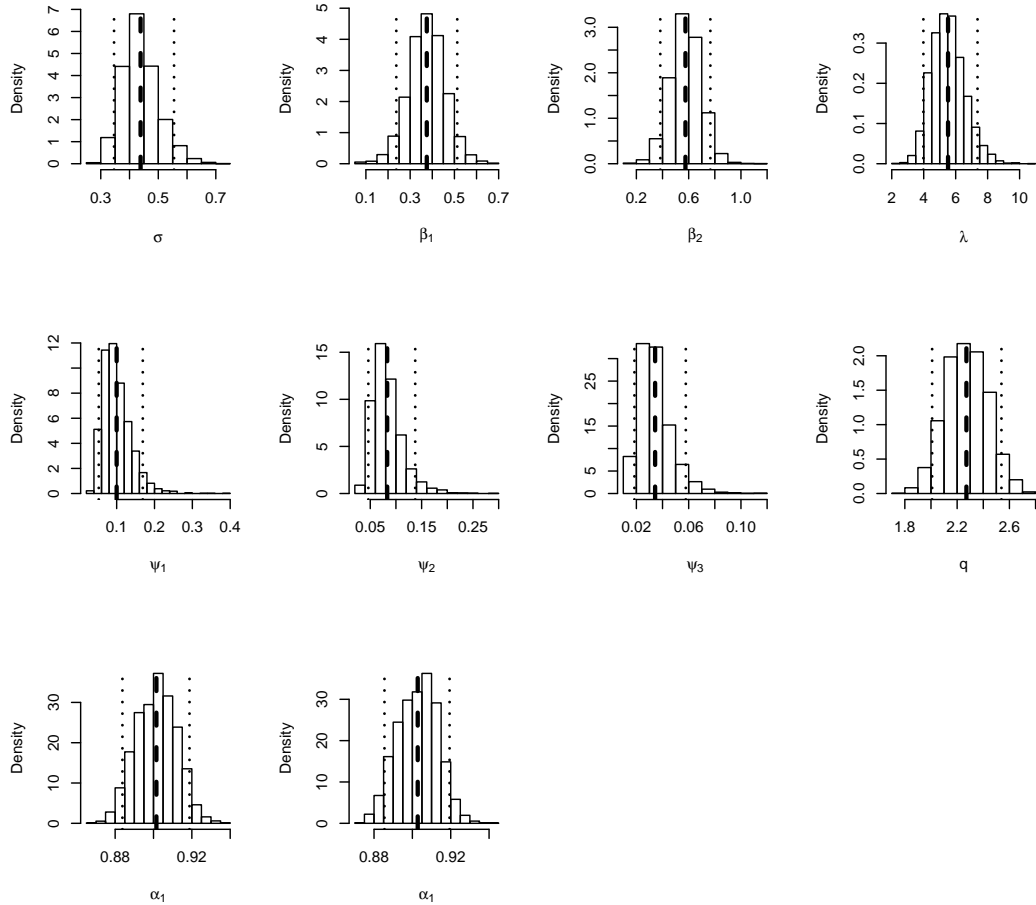


Figura 4.9: Gráficos de dispersão do valor esperado do fator versus o valor verdadeiro.

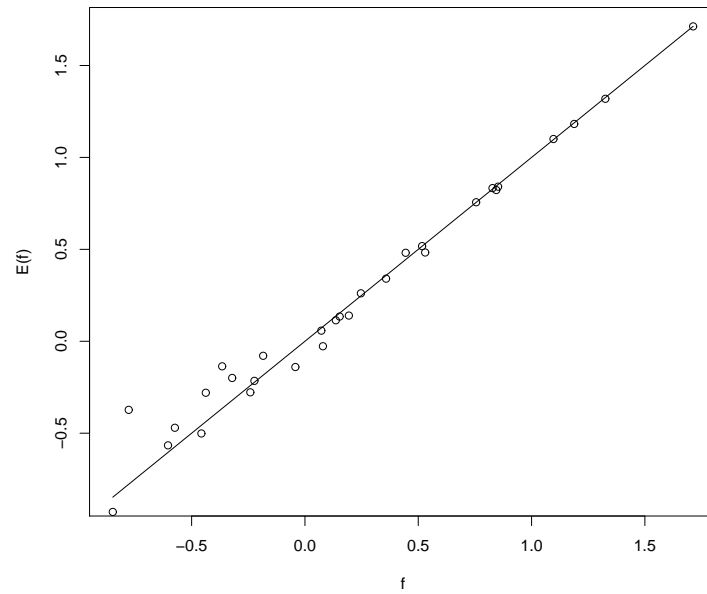
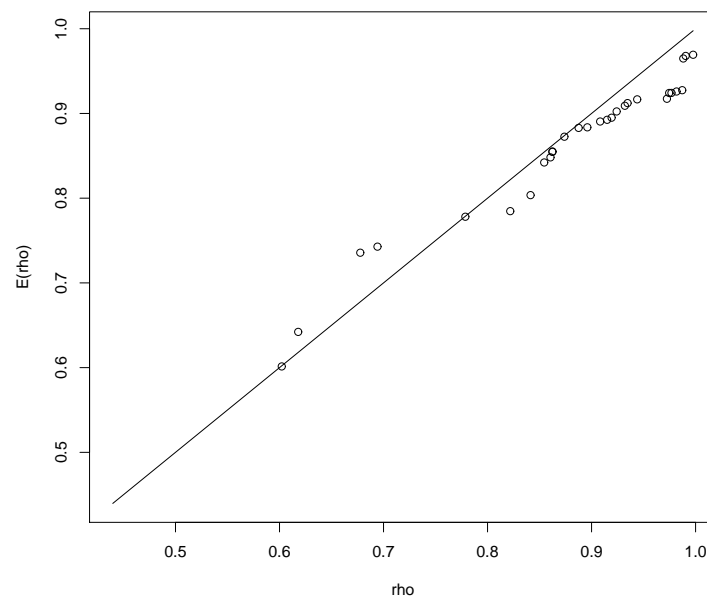


Figura 4.10: Gráficos de dispersão do valor esperado da eficiência versus o valor verdadeiro.



4.7.4 Exemplo 4: Uma outra alternativa de fronteira estocástica para múltiplos-produtos.

Nesta exemplo será aplicado o modelo apresentado na Seção 4.6 a dados simulados. Os dados foram gerados a partir do modelo proposto por Fernández *et al.* (2000) através da seguinte relação:

$$y_{i,j} = \frac{\eta_{i,j}^{1/q} \theta_i}{\alpha_j} \quad (4.23)$$

onde $j = 1, \dots, p$, com $p = 3$, $i = 1, \dots, N$, $N = 30$, $\alpha = (1/3, 1/3, 1/3)$, $q = 2$ e $\eta_i \sim D_p(s)$ com parâmetro $s = (1, 1, 1)$. A fronteira estocástica $\log g(y)$ foi gerada a partir da seguinte distribuição normal:

$$(\log \theta_i | \beta, \sigma^2, u_i) \sim N(x'_i \beta - u_i, \sigma^2), \quad (4.24)$$

onde $\beta = (0.5, 0.5)$, $\sigma = 0.5$ e u foi gerada a partir da distribuição exponencial com parâmetro $\lambda = 5$.

A inferência será feita com base na amostra da distribuição a posteriori gerada a partir do algoritmo MCMC.

Modelo1	Modelo de fronteira estocástica para múltiplos-produtos proposto por Fernández <i>et al.</i> (2000)
Modelo 2	Modelo de fronteira estocástica fatorial com transformação de Box-Cox

Para cada modelagem foi gerada uma cadeia de tamanho 15.000. A amostra a posteriori foi obtida utilizando um burn-in de 5.000. As seguintes especificações a priori foram utilizadas: $\alpha \sim D(a)$ com $a = (1, 1, 1)$, $s_j \sim G(0.01, 0.01)$, $q \sim \text{Exp}(0.01)I_{(1,\infty)}(q)$, $\lambda \sim G(0.01, .01)$, $\beta \sim NM(\underline{0}, \text{diag}(0.001, 0.001))$ e $\sigma_v^2 \sim GI(0.01, 0.01)$. Para o modelo fatorial com transformação de Box-Cox foram usadas as mesmas distribuições a priori definidas na seção anterior.

As Figuras 4.11 e 4.12 apresentam os histogramas da amostra da distribuição a posteriori de todos os parâmetros em cada modelo. Fazendo um paralelo entre os dois modelos, observa-se que alguns parâmetros não são comparáveis. Os resultados mostraram que os dois modelos forneceram estimativas diferentes para o parâmetro q . Porém, observamos que as médias a posteriori dos parâmetros relacionados à fronteira estocástica β_1 , β_2 , σ e λ são parecidas nos dois modelos. Ou seja, apesar de estimativas diferentes para o parâmetro q é obtido um produto agregado similar.

Na Figura 4.13 são apresentados os gráficos de dispersão da média a posteriori de ρ versus os valores verdadeiros para os dois modelos. Nota-se que na Figura 4.13 (b) as estimativas das eficiências, utilizando o modelo proposto, foram mais próximas aos valores verdadeiros do que as estimativas a posteriori obtidas utilizando a abordagem proposta por Fernández *et al.* (2000).

Uma forma de comparar as estimativas obtidas dos dois modelos é através das distribuições dos ranks associados a medida de eficiência. Essas distribuições são diretamente obtidas usando as cadeias de MCMC. Se denotarmos por $e_j^{(m)} = \exp(-u_j^{(m)})$ como a medida de eficiência da j^{th} firma na m^{th} replicação da cadeia de MCMC e colocarmos os $e_j^{(m)}$ em ordem para $j = 1, \dots, N$, o rank associado à j^{th} firma na m^{th} replicação é a posição ocupada por $e_j^{(m)}$ na ordenação. As Figuras 4.14 e 4.15 apresentam os histogramas do ranking da eficiência para duas firmas segundo o modelo de fronteira fatorial e, também, usando o modelo proposto por Fernández *et al.* (2000). Observamos que os dois modelos conseguiram diferenciar as firmas mais eficientes das menos eficientes.

Figura 4.11: Histograma dos parâmetros - Modelo de fronteira estocástica para múltiplos-produtos proposto por Fernández et al (2000). As retas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%, respectivamente.

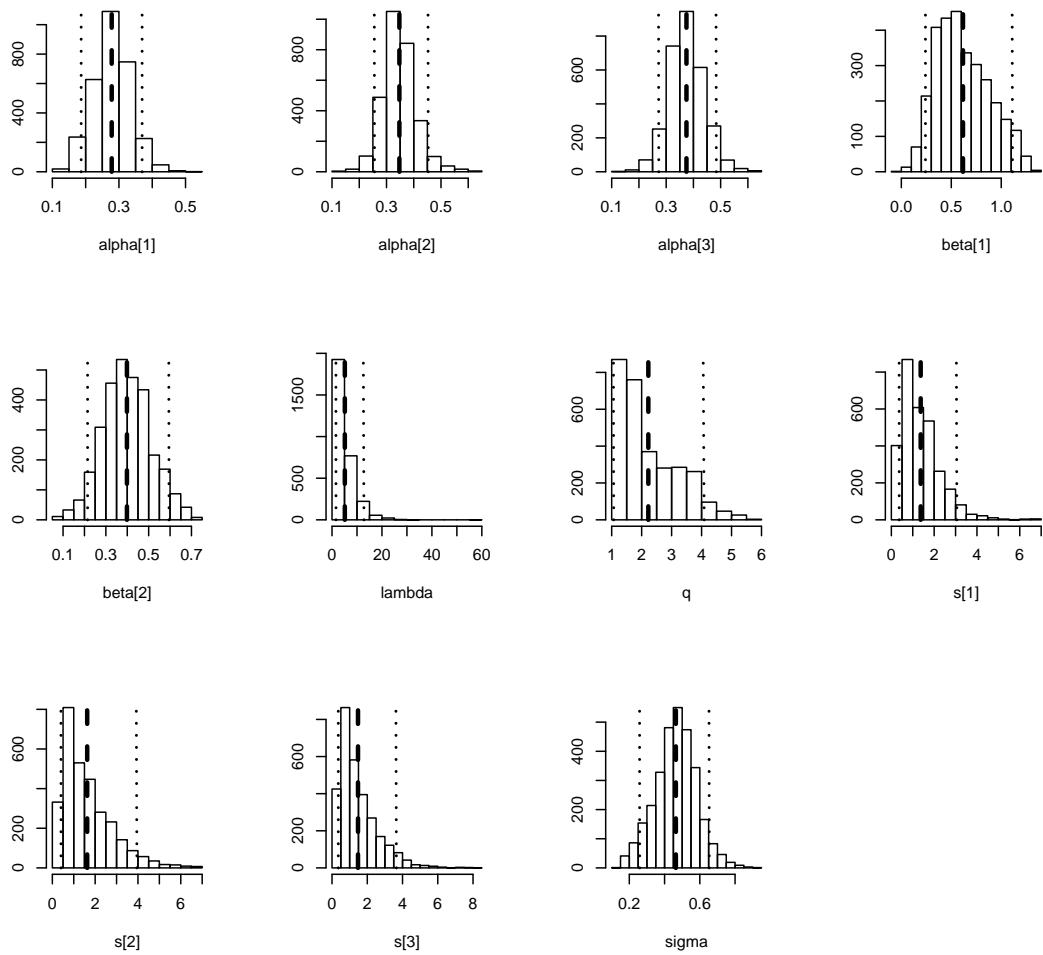


Figura 4.12: Histograma dos parâmetros - Modelo de fronteira estocástica fatorial com transformação de Box-Cox. As retas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%, respectivamente.

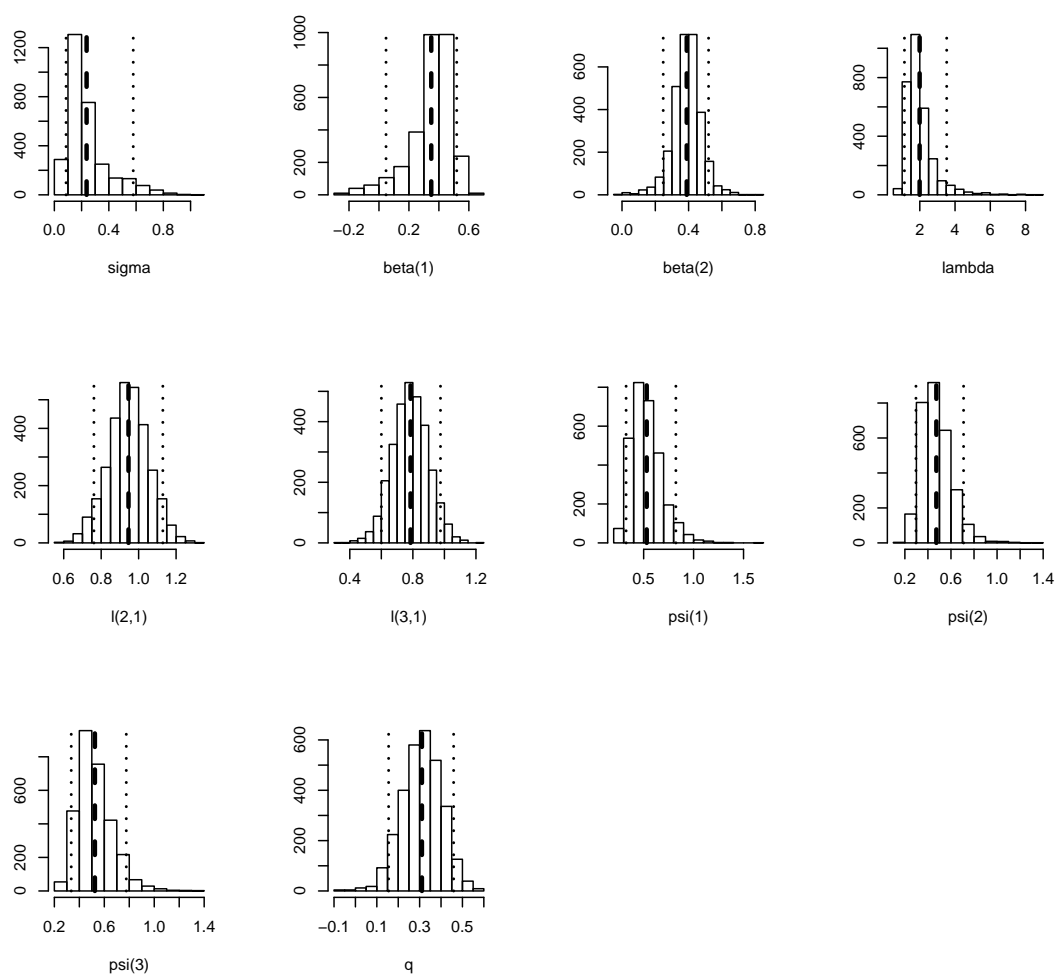
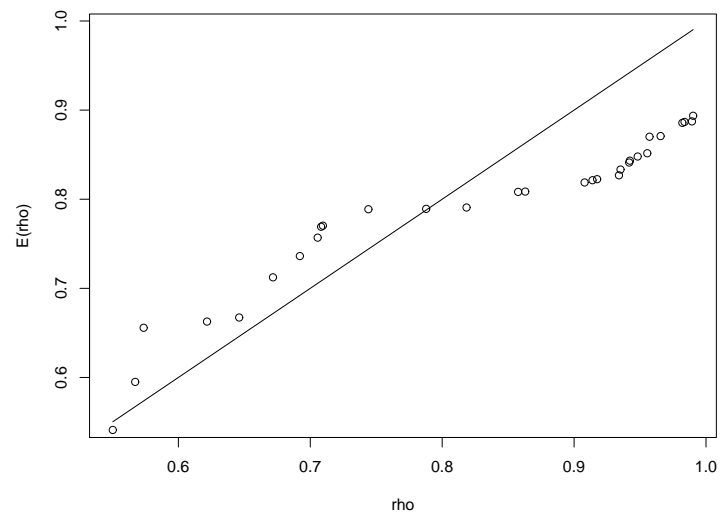
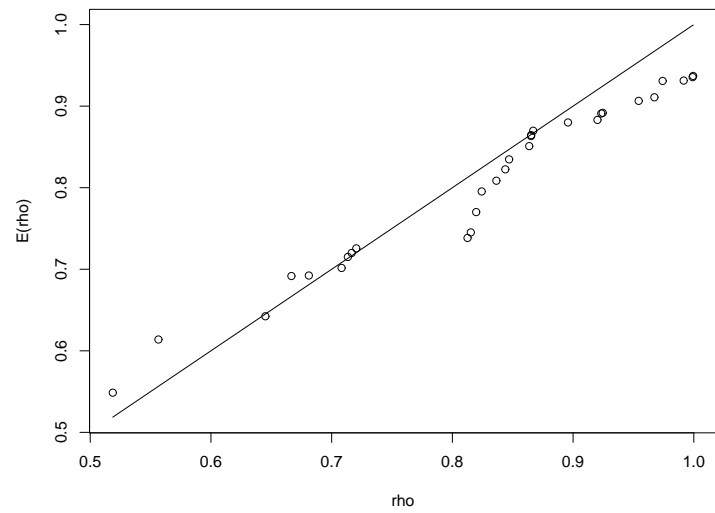


Figura 4.13: Gráfico de dispersão do valor esperado de ρ versus o seu valor verdadeiro.



(a) Modelo de fronteira proposto pela Fernández



(b) Modelo de fronteira estocástica fatorial

Figura 4.14: Histogramas dos *ranking* da eficiência para as firmas 10 e 17 segundo o modelo de fronteira estocástica fatorial

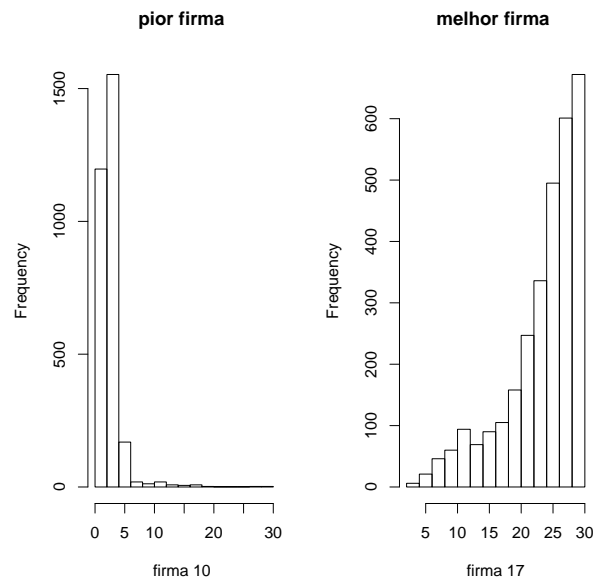
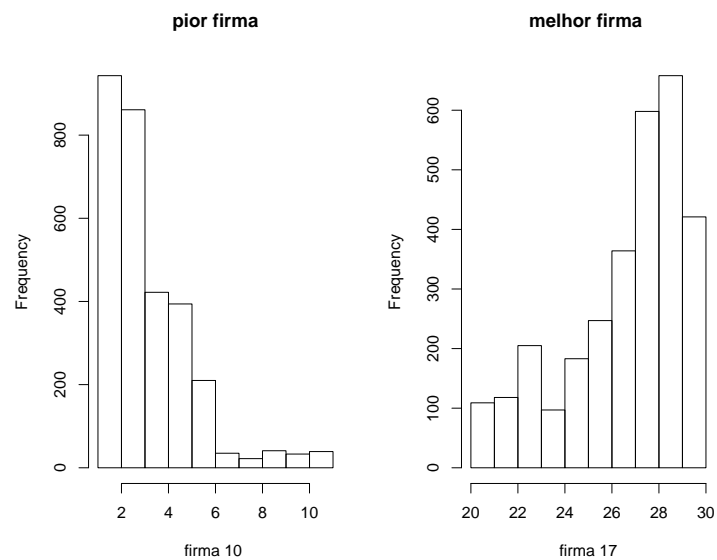


Figura 4.15: Histogramas dos *ranking* da eficiência para as firmas 10 e 17 segundo o modelo proposto por Fernández *et al.* (2000)



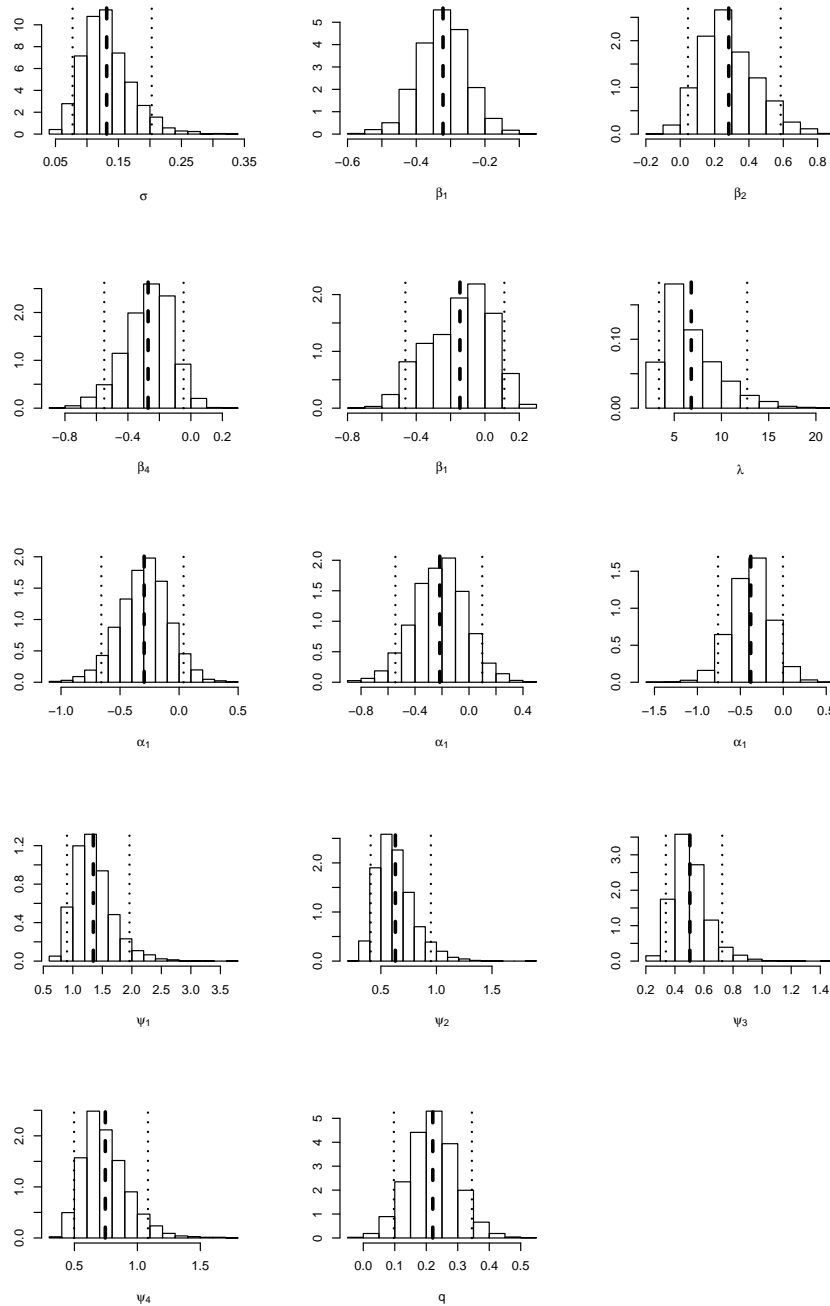
4.8 Aplicação a dados reais

Nesta seção será utilizado o modelo de fronteira estocástica fatorial proposto na Seção 4.6 para modelar dados com múltiplos produtos. O conjunto de dados refere-se à produção em pesquisas agrícolas de $N = 34$ unidades de pesquisas da Embrapa no ano de 1996. O conjunto de dados contém $p = 4$ produtos: (1) produção científica, (2) produção de publicações técnicas, (3) produção de desenvolvimento de tecnologias, produtos e processos e (4) produção de difusão de tecnologias e imagem, 3 insumos (pessoal, outras despesas e capital). Os detalhes sobre os dados são descritos em Souza *et al.* (1996).

Souza *et al.* (1999) e Medici (2000) ajustaram um modelo de fronteira estocástica para o caso de um único produto, que foi construído através de uma soma ponderada de y_i , $i = 1, 2, 3, 4$ onde os pesos são fixos. Nesta tese será ajustado o modelo de fronteira estocástica fatorial com transformação de Box-Cox. A inferência será feita com base na amostra da distribuição a posteriori gerada usando o algoritmo MCMC proposto. Foi gerada uma cadeia de tamanho 20.000. A amostra a posteriori foi obtida utilizando um burn-in de 10.000. Amostras foram guardadas a cada 5 iterações e, portanto, os resultados estão baseados em cadeias de tamanho 2.000. A verificação da convergência foi verificada através do critério de convergência de Gelman & Rubin, que ficaram próximas a 1 o que sugere que as cadeias convergiram (ver estatísticas no Apêndice). As distribuições a priori foram as mesmas especificadas na seção 4.6.

A Figura 4.16 apresenta os histogramas das distribuições marginais a posteriori geradas via MCMC para o modelo de fronteira estocástica fatorial. As linhas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%. Observamos que a média a posteriori do parâmetro da transformação de Box-Cox, q , é 0.13 e o IC de 95% não inclui o valor 0.

Figura 4.16: Histograma das distribuições marginais a posteriori segundo o Modelo de fronteira estocástica fatorial



As Figuras 4.17 e 4.18 mostram, respectivamente, o gráfico de dispersão da média a posteriori para as eficiências individuais ρ_i versus as estimativas apresentadas no artigo de Souza *et al.* (1996) e versus as estimativas obtidas utilizando o modelo proposto por Fernández *et al.* (2000). Observamos que as estimativas das eficiências utilizando o modelo proposto foram maiores aos resultados que utilizam as outras abordagens. No entanto, a ordenação das firmas foram similares as

outras abordagens, se observamos a relação linear existente entre as estimativas das eficiências.

Figura 4.17: Gráfico de dispersão do valor esperado de ρ apresentado em Souza *et al.* (1996) versus o valor esperado de ρ utilizando o modelo de fronteira fatorial.

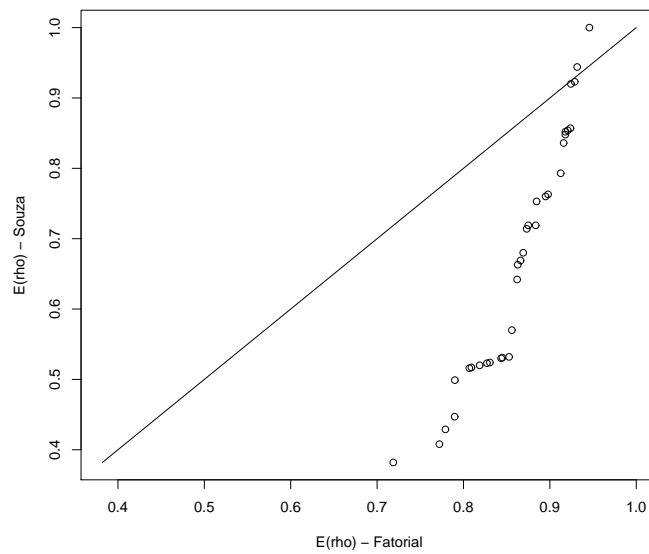
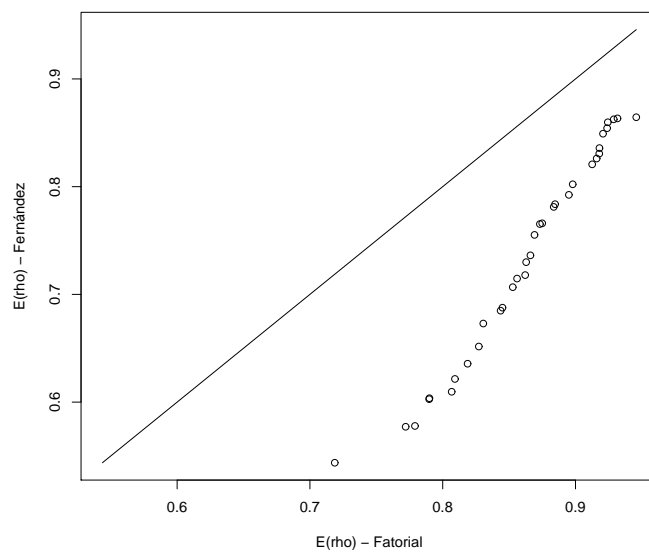
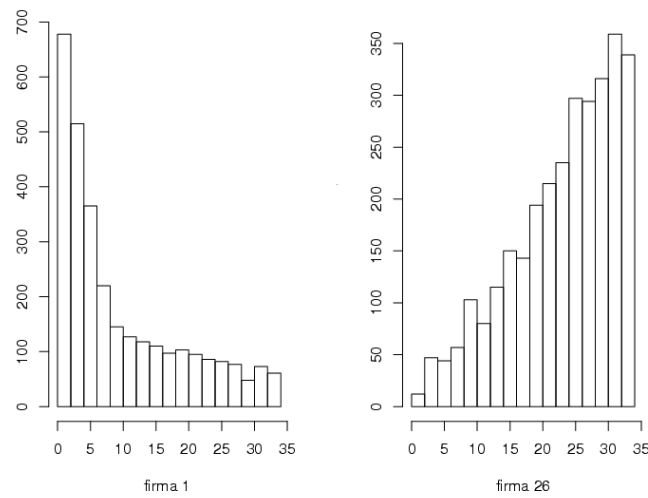


Figura 4.18: Gráfico de dispersão do valor esperado de ρ obtido pelo modelo da Fernández *et al.* (2000) versus o valor esperado de ρ utilizando o modelo de fronteira fatorial.

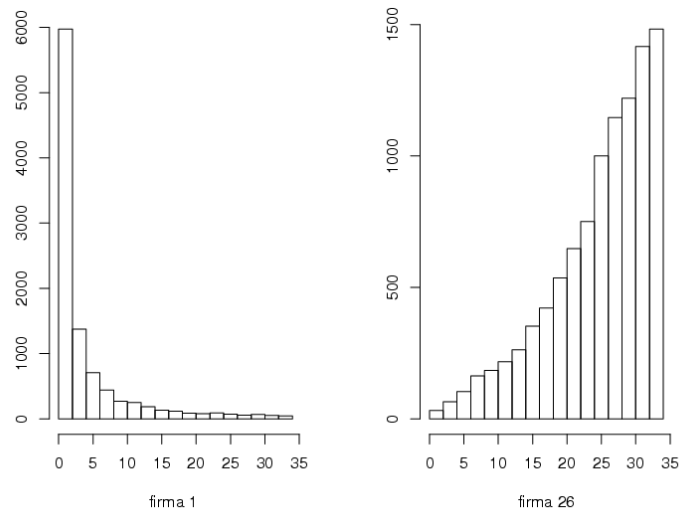


A Figura 4.19 apresenta os histogramas dos *rankings* das eficiências de duas unidades de pesquisas da Embrapa segundo o modelo de fronteira estocástica fatorial e, também, o modelo proposto por Fernández *et al.* (2000). Observa-se que os dois modelos conseguiram diferenciar as firmas mais eficientes das menos eficientes.

Figura 4.19: Histograma dos *rankings* da eficiência para as unidades de pesquisas da Embrapa UD1 e UD26



(a) modelo de fronteira estocástica fatorial



(b) Modelo de fronteira proposto pela Fernández *et al.* (2000)

Capítulo 5

Modelo de Fronteira Estocástica Fatorial Dinâmico

Neste capítulo, um modelo de fronteira estocástica com dados de painel para múltiplos produtos é proposto. Este modelo, é uma extensão do modelo de fronteira fatorial proposto no capítulo anterior, e permite que as ineficiências técnicas variem no tempo de forma dinâmica. Esta especificação permite a inclusão de diversas estruturas temporais, tais como tendência e sazonalidade na evolução das ineficiências. A inferência para este modelo é realizado utilizando uma abordagem Bayesiana através do algoritmo MCMC. Finalmente, um estudo simulado e um exemplo usando os dados da Embrapa são apresentados para mostrar a performance do modelo proposto.

5.1 Introdução

Um importante problema nos modelos de medida de eficiência técnica diz respeito a como modelar o comportamento temporal da ineficiência. Principalmente quando temos o caso de múltiplos produtos. Numa abordagem clássica para

um único produto, os primeiros modelos (Pitt e Lee, 1981; Schmidt e Sickles, 1984; Kumbhakar, 1987; entre outros) trataram a eficiência técnica invariante no tempo. Pesquisas subseqüentes permitiram a eficiência técnica variar com o passar do tempo (Kumbhakar, 1990; Cornell, Schmidt and Sickles, 1990; Battese and Coelli, 1992; Lee and Schmidt, 1993). Detalhes desses modelos, no contexto de estimar a função de produção, podem ser encontrados no capítulo 3 de Kumbhakar e Lovell (2000). Ahn *et al.* (2002) e Desli, Ray e Kumbhakar (2003) propuseram um modelo para dados em painel com a ineficiência técnica dinâmica. Estes estudos propõem um modelo especificando uma mudança na eficiência por meio de um intercepto específico para cada firma. Este intercepto evolui com o passar do tempo através de um processo auto-regressivo de primeira ordem (AR(1)). Dois fatos podem ser responsáveis por ignorar os modelos dinâmicos. O primeiro refere-se a complexidade da função de verossimilhança e a dificuldade de realizar inferência sobre a ineficiência de uma firma não observada. Assim, os investigadores preferem frequentemente assumir que as ineficiências são invariantes no tempo.

Num contexto Bayesiano, Koop *et al.* (2000 a,b) assume que os parâmetros relacionados a fronteira estocástica são os mesmo para todas as unidades com o passar do tempo. Tal suposição pode ser questionável, especialmente em aplicações onde o número de observações no tempo (T) é grande. Uma extensão desta abordagem é assumir que os parâmetros sejam invariantes no tempo e permitir uma heterogeneidade *cross-sectional* através dos coeficientes aleatórios (Tsionas 2002). As outras extensões assumem que os parâmetros são constantes no nível das firmas, mas podem variar no tempo (Koop *et al.* 2000a,b). Migon e Medici (2002) apresentam um modelo mais geral, em que os parâmetros relacionados a fronteira estão variando no tempo e no nível das firmas e, ainda definem algumas estruturas impostas na ineficiência técnica. Outra contribuição em dados de painel é

apresentado em Fernández *et al.* (1997), onde este mostra uma discussão formal da existência da distribuição a posteriori em modelos de fronteira estocástica com priori imprópria. Recentemente, Tsionas (2006) propõe um modelo de fronteira estocástica com uma ineficiência técnica dinâmica, definindo esta através de um processo auto-regressivo de primeira ordem (AR(1)). Em todos estes modelos de fronteira estocástica para dados de painel, a aplicação é para caso das firmas utilizem vários insumos para produzir um único produto.

Neste capítulo, o modelo de fronteira estocástica fatorial proposto no capítulo anterior é generalizado num contexto de dados de painel. Suponha agora que temos um conjunto de p produtos medidos em vários instantes de tempo $t = 1, \dots, T$. Assim, vamos estimar, ao longo do tempo, a eficiência técnica de firmas que possuem múltiplos produtos, através da fronteira estocástica obtida pelo modelo fatorial com transformação de Box-Cox. O comportamento temporal é descrito através das ineficiências u_{it} . Permitir que as ineficiências variem no tempo de forma dinâmica no contexto de múltiplos produtos, representa uma das principais contribuições deste capítulo no contexto da modelagem temporal.

Este capítulo está organizado da seguinte forma. Na Seção 5.2 é apresentado o modelo proposto. Na Seção 5.3, um algoritmo MCMC é proposto para fazer inferência a posteriori de todos os parâmetros do modelo. Por último, na Seção 5.4, um estudo simulado é apresentado para mostrar a qualidade da aplicação do modelo.

5.2 Modelo Proposto

Seja p o número de produtos para uma determinada firma i e $y_{it} = (y_{it1}, \dots, y_{itp})$ o vetor de observações de dimensão p produzido pela firma i ($i = 1, \dots, N$) no tempo t ($t = 1, \dots, T$). A primeira parte do modelo de fronteira estocástica fatorial dinâmico (MFEFD) proposto é definido pelo seguinte modelo fatorial

com transformação de Box-Cox com apenas um fator

$$y_{it}^{(q)} = Lf_{it}^{(q)} + \epsilon_{it} \quad \epsilon_{it} \sim NM(0, \Psi), \quad (5.1)$$

onde $L = (1, l_2, \dots, l_p)'$ é o vetor de cargas do fator comum de dimensão $p \times 1$ e a matriz Ψ representa a variância de observação. Por simplicidade é assumido que $\Psi = \text{diag}(\psi_1, \psi_2, \dots, \psi_p)$. $y_{it}^{(q)}$ e $f_{it}^{(q)}$ representam a transformação de Box-Cox de y_{it} e f_{it} , respectivamente.

O modelo (5.1) define a transformação de um vetor de output multivariado y_{it} para uma quantidade univariada f_{it} . Dada esta transformação, o problema básico de encontrar uma fronteira estocástica é essencialmente o mesmo que no caso de um único produto. Da mesma forma que no modelo de fronteira estocástica fatorial, o valor de $f_{it}^{(q)}$ pode ser considerado um “produto agregado”, então é sensato definir $\delta_{it} = \log f_{it}^{(q)}$. Assim, o modelo de fronteira estocástica para δ_{it} será

$$\delta_{it} = x'_{it}\beta + v_{it} - u_{it} \quad v_{it} \sim N(0, \sigma_v^2), \quad (5.2)$$

onde x_{it} é um vetor de regressores de dimensão $k \times 1$, β é um vetor de parâmetros relacionados a fronteira estocástica de dimensão $k \times 1$, v_{it} é a componente do erro de medida e u_{it} é o termo de erro não negativo representando a ineficiência técnica. É usual assumir que v_{it} e u_{it} são mutuamente independentes, como também, independentes de x_{it} .

Para o logaritmo da ineficiência técnica especificamos

$$\log u_{it} = E'\underline{\theta}_{it} + z'_{it}\gamma_{it} \quad (5.3)$$

$$\underline{\theta}_{it} = G\underline{\theta}_{i,t-1} + \omega_{it1} \quad \omega_{it1} \sim N(0, W_1)$$

$$\gamma_{it} = \gamma_{i,t-1} + \omega_{it2} \quad \omega_{it2} \sim N(0, W_2),$$

onde z_{it} é um vetor de covariáveis de dimensão $m \times 1$, $E = (1, 0, \dots, 0)'$ e $\theta_{it} = (\theta_{i,t}, \theta_{i,t-1}, \dots, \theta_{i,t-r+1})'$ são vetores de dimensão $r \times 1$ e γ_{it} é unidimensional. A

matriz G de dimensão $r \times r$ caracteriza a evolução dinâmica das ineficiências, enquanto que W_1 , de dimensão $r \times r$, representa a variância da evolução. Por simplicidade é assumido que $W_1 = \text{diag}(\sigma_u^2, 0, \dots, 0)$. A evolução dinâmica das log-ineficiências é caracterizada por

$$G = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_r \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}.$$

A escolha da dimensão de E e G depende do modelo desejado e da natureza das ineficiências técnicas que se pretende descrever. Por exemplo se $r = 3$ temos

$$\begin{aligned} \theta_{i,t} &= \phi_1 \theta_{i,t-1} + \phi_2 \theta_{i,t-2} + \phi_3 \theta_{i,t-3} + \omega_{it1} \\ \theta_{i,t-1} &= \theta_{i,t-1} \\ \theta_{i,t-2} &= \theta_{i,t-2} \end{aligned}$$

Neste caso, o logaritmo da ineficiência é definido através de um processo autorregressivo de terceira ordem (AR(3)). No caso de um único produto para $r = 1$ e $\omega_{it2} = 0$, obtemos o modelo de fronteira dinâmico proposto por Tsionas (2006). E, também, se $\phi_1 = 0$ obtemos o modelo de fronteira estocástica apresentado em Battese e Coelli (1995) e Kumbhakar *et al.* (1991).

O interesse principal nos modelos de fronteira estocástica é estimar a eficiência que é definida por $\exp(-u_{it})$. Na prática, os modelos diferem de acordo com a estrutura imposta a u_{it} . Em muitas situações, é razoável pensar que as ineficiências só dependem das firmas, porém, outras situações podem ser pensadas como, por exemplo, pode existir algum tipo de influência temporal ou até mesmo de variáveis exógenas como características destas firmas.

Comparação com outros modelos

Para acomodar as várias estruturas assumidas, em Migon e Medici (2002) são apresentadas algumas estruturas impostas ao termo de ineficiência u_{it} através de uma função $g(u_{it})$.

$$g^{-1}(u_{it}) = \begin{cases} u_{it}^* + z_{it}'\gamma_i, \\ u_i^* f(\eta_i^* t) + z_{it}'\gamma_i, \end{cases} \quad (5.4)$$

onde z_{it} é um vetor de covariáveis $m \times 1$, γ é um vetor de parâmetros $m \times 1$, η_i^* é um parâmetro escalar desconhecido e g é tal que as ineficiências estejam definidas no \mathbb{R}^+ e tenham as seguintes distribuições:

1. $NT(\mu, \sigma_u^2), \sigma_u^2 > 0$
2. $Ga(p, \theta), p, \theta > 0$
3. $LN(\mu, \sigma_u^2)$

Obviamente, existe um grande número de modelos a serem considerados para cada aplicação em particular. De acordo com as especificações em g é possível determinar as seguintes estruturas para as ineficiências

- $g^{-1}(u_{it}) = u_{it}^*$. Esse caso implica que cada firma em cada período de tempo tem sua própria ineficiência (Pitt e Lee, 1981).
- Battese e Coelli(1992) propõem um modelo variando no tempo, para as ineficiências técnicas, definido por

$$g^{-1}(u_{it}) = u_i^* \exp(-\eta(t - T))$$

- $g^{-1}(u_{it}) = u_i^*$ significa que cada ineficiência da firma i foi assumida invariante no tempo (Pitt e Lee, 1981 e Schmidt e Sickles, 1984).

- $g^{-1}(u_{it}) = u_{it}^* + z_{it}'\gamma_i$, permite que as ineficiências variem sobre as firmas e o tempo, e admite a inclusão de variáveis exógenas.

No modelo aqui proposto assume-se que u_{it} assume uma distribuição lognormal e neste caso $g^{-1}(u_{it}) = \log(u_{it})$.

Função de verossimilhança

Sejam $Y = (y_{11}, \dots, y_{NT})$ a matriz de observações de dimensão $NT \times p$ e $f = (f_{11}, \dots, f_n')$ o vetor do fator único de dimensão $NT \times 1$, assim a função de verossimilhança para $\Omega = (f, L, \Psi, q)$ é dada por

$$p(Y|\Omega) = \frac{|\Psi|^{-\frac{NT}{2}}}{(2\pi)^{\frac{NTp}{2}}} \exp \left\{ -0.5 \sum_{t=1}^T \sum_{i=1}^N (y_{it}^{(q)} - f_{it}^{(q)} L')' \Psi^{-1} (y_i^{(q)} - f_{it}^{(q)} L') \right\} \prod_{i,j,t} y_{itj}^{q-1}.$$

5.3 Inferência Bayesiana e algoritmo de MCMC

Nesta seção é apresentado o procedimento de inferência proposto seguindo uma abordagem completamente Bayesiana. Inicialmente distribuições a priori para todos os parâmetros são apresentadas. Em seguida, o algoritmo MCMC é utilizado para obter amostras a posteriori de todos os parâmetros do modelo.

5.3.1 Distribuição a priori

As distribuições a priori para os parâmetros relacionados ao modelo fatorial (L, Ψ, q) são as mesmas que foram definidas na Equação (4.16) no Capítulo 3. Considerando uma independência a priori (L, Ψ, q) com os parâmetros relacionados à fronteira $(\beta, \sigma^2, \sigma_u^2, \phi, \gamma)$, onde $\phi = (\phi_1, \dots, \phi_r)$, obtemos a seguinte distribuição a priori conjunta

$$p(L, \Psi, q, \beta, \sigma_v^2, \sigma_u^2, \phi, \gamma) = \frac{p(L)p(\Psi)p(q)}{J_q^{k/N}} p(\beta) p(\sigma_v^2) p(\phi | \sigma_u^2) p(\sigma_u^2) p(\gamma) \quad (5.5)$$

onde $l_j \sim N(\mu_0, C_0)$ para $j = 2, \dots, p$, $\psi_j \sim GI(v_0/2, s_0/2)$ para $j = 1, \dots, p$, $\beta \sim N(b_0, H_0^{-1})$, $\sigma_v^2 \sim GI(n_0/2, a_0/2)$, $q \sim U(a, b)$ e $\gamma \sim N(\bar{\gamma}, \bar{V}_\gamma)$. De forma similar à usada por Broemeling e Cook (1993) para um processo $AR(r)$, as seguintes distribuições a priori foram usadas: $\sigma_u^2 \sim GI(n_u/2, a_u/2)$ e $\phi|\sigma_u^2$ tem distribuição normal multivariada $NM(\underline{a}_\phi, \Sigma_\phi \sigma_u^2)$. Os hiperparâmetros são escolhidos de forma que as distribuições a priori sejam vagas.

5.3.2 Inferência a posteriori

A distribuição a posteriori conjunta de $(f, L, \Psi, q, u, \beta, \sigma_v^2, \sigma_u^2, \phi, \gamma)$ é proporcional a

$$\begin{aligned}
p(f, L, \Psi, q, u, \beta, \sigma_v^2, \sigma_u^2, \phi, \gamma | Y) &\propto \prod_{i=1}^N \prod_{t=1}^T p(y_{it}^{(q)} | f_{it}, L, \Psi, q) p(L) p(\Psi) p(q) \\
&\times \prod_{i=1}^N \prod_{t=1}^T p(\log f_{it}^{(q)} | u_{it}, \beta, \sigma_v^2) p(\beta) p(\sigma_v^2) \\
&\times \prod_{i=1}^N \prod_{t=1}^T p(\log u_{it} | \sigma_u^2, \phi, \gamma) p(\sigma_u^2) p(\phi) p(\gamma) \\
&\times \prod_{i=1}^N \prod_{t=1}^T \frac{1}{u_{it}} \frac{f_{it}^{q-1}}{f_{it}^{(q)}} \prod_{j=1}^p y_{itj}^{q-1} \quad (5.6)
\end{aligned}$$

a qual é analiticamente intratável e, portanto, o algoritmo MCMC será utilizado para amostrar os parâmetros de interesse. Neste caso, as distribuições condicionais completas a posteriori para parâmetros $L, \Psi, \beta, \sigma_v^2, \sigma_u^2$ e γ são conhecidas e disponíveis para amostragem. Apenas as distribuições condicionais de q, f_{it}, u_{it} e ϕ não são conhecidas. Portanto, será utilizado o algoritmo *Slice Sampling* para a amostragem destas distribuições condicionais. O algoritmo de MCMC será desenvolvido a partir das seguintes distribuições condicionais:

- Distribuição condicional completa de L

- Para as linhas $j = 2, \dots, p$, seja $L_j = (l_{j,1}, \dots, l_{j,j-1})'$, então

$$(L_j|Y, f, \Psi, q) \sim N(\mu_j, C_j),$$

onde $\mu_j = C_j^{-1}(C_0^{-1}\mu_0 1_j + \psi_j^{-2} P_j' z_{1j})$ e $C_j = C_0^{-1} 1_j + \psi_j^{-2} P_j' P_j$. Denota-se por P_j é a matriz $NT \times (j-1)$ contendo as primeiras

$(j-1)$ colunas de $f^{(q)} = (f_{11}^{(q)}, \dots, f_{NT}^{(q)})$ e z_{1j} é a coluna j de $Y^{(q)} = (y_{11}^{(q)}, \dots, y_{NT}^{(q)})$.

- Para $j = k+1, \dots, p$, seja $L_j = (l_{j,1}, \dots, l_{j,k})'$, então

$$(L_j|Y, f, \Psi, q) \sim N(\mu_j, C_j),$$

onde $\mu_j = C_j^{-1}(C_0^{-1}\mu_0 1_k + \psi_j^{-2} f^{(q)'} z_{1j})$ e $C_j = C_0^{-1} 1_k + \psi_j^{-2} f^{(q)'} f^{(q)}$.

- Distribuição condicional completa de ψ_j^2 . Para $j = 1, \dots, p$,

$$(\psi_j^2|Y, L, f, q) \sim GI\left(\frac{(v_0 + N)}{2}, \frac{(s_0 + (z_{1j} - f^{(q)} L_j)'(z_{1j} - f^{(q)} L_j))}{2}\right).$$

- Distribuição condicional completa de q

$$\begin{aligned} p(q|Y, L, f, \Psi, \beta, \sigma_v^2, u) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp\left(\frac{1}{2}(y_{it}^{(q)} - L f_{it}^{(q)})' \Psi^{-1} (y_{it}^{(q)} - L f_{it}^{(q)})\right) \\ &\times \exp\left[-\frac{1}{2\sigma_v^2}(\log f_{it}^{(q)} - x_{it}'\beta + u_{it})^2\right] / f_{it}^{(q)} \\ &\times \prod_{i=1}^N \prod_{t=1}^T f_{it}^{q-1} \prod_j^p y_{itj}^{q-1} \end{aligned}$$

- Distribuição condicional completa de f_i . Para $i = 1, \dots, N$

$$\begin{aligned} p(f_{it}|Y, L, \Psi, q, \beta, \sigma_v^2, \lambda, u) &\propto \exp\left(\frac{1}{2}(y_{it}^{(q)} - L f_{it}^{(q)})' \Psi^{-1} (y_{it}^{(q)} - L f_{it}^{(q)})\right) f_{it}^{q-1} \\ &\exp\left[-\frac{1}{2\sigma_v^2}(\log f_{it} - x_{it}'\beta + u_{it})^2\right] \frac{1}{f_{it}^{(q)}}. \end{aligned}$$

- Distribuição condicional completa de β

$$(\beta|\delta, \sigma_v^2, \sigma_u^2, u, \gamma) \sim N(b_*, H_*^{-1}),$$

onde $b_* = H_*^{-1}\{H_0 b_0 + \sigma_v^{-2} x'(\delta + u)\}$ e $H_* = H_0 + \sigma_v^{-2}(x'x)$.

- Distribuição condicional completa de σ_v^2

$$(\sigma_v^2|\delta, \beta, \sigma_u^2, u, \gamma) \sim GI(c_*, d_*),$$

onde $c_* = \frac{(N + N_0)}{2}$ e $d_* = \frac{[(\delta + u - X\beta)'(\delta + u - X\beta) + a_0]}{2}$.

- Distribuição condicional completa de σ_u^2

$$(\sigma_u^2|\delta, \beta, \sigma_v^2, u, \gamma) \sim G\left(N + v_0, \sum_{i=1}^N \sum_{t=1}^T u_{it} + \omega_0\right).$$

- Distribuição condicional completa de u

$$\begin{aligned} p(u|\delta_{it}, \beta, \sigma_v^2, \sigma_u^2, \gamma, \phi) &\propto \exp\left[-\frac{1}{2\sigma_v^2} \sum_{i=1}^N \sum_{t=1}^T (u_{it} + \delta_{it} - \mathbf{x}_{it}\beta)^2\right] \\ &\times \exp\left[-\frac{1}{2\sigma_u^2} \sum_{i=1}^N \sum_{t=1}^T (\delta_{it} - \mathbf{z}_{it}\gamma - \phi_1 \log u_{i,t-1})^2\right] \\ &\times \prod_{i=1}^N \prod_{t=1}^T \frac{1}{u_{it}}. \end{aligned}$$

- Distribuição condicional completa de γ

$$(\gamma|\delta, \sigma_v^2, \sigma_u^2, u) \sim N(a_*, C_*),$$

onde $a_* = C_*^{-1}\{\bar{\gamma}\bar{V}_\gamma^{-1} + \sigma_u^{-2} z' \log u_1 + \sigma_u^{-2} z'(\log u_2 - \phi \log u_3)\}$ e $C_* = \bar{V}_\gamma^{-1} + \sigma_u^{-2} z'z$. Denota-se por $\log u_1$ o vetor

A seguir serão apresentados três exemplos com dados simulados e uma aplicação com dados de produção em pesquisas agrícolas de unidades pesquisa da Embrapa.

5.4 Estudo simulado

Neste estudo simulado é gerado um conjunto de dados a partir do modelo estocástico fatorial dinâmico com $p = 3$ produtos e $N = 30$ firmas para $T = 10$ instantes de tempo. Os parâmetros considerados foram: $\Psi = \text{diag}(0.01, 0.01, 0.01)$, $L = (1, 2, 2)'$, $q = 2.2$, $\beta = (1, 1)$, $\sigma_v = 0.6$, $\phi = 0.9$, $\gamma = (0.02, 0.02)$ e $\sigma_u = 0.05$.

Foram realizadas 5.000 iterações no qual foram descartadas as 2.000 primeiras para o período de burn-in. A verificação da convergência foi verificada através do critério de convergência de Gelman & Rubin, que ficaram próximas a 1 o que sugere que as cadeias convergiram (ver estatísticas no Apêndice). Os hiperparâmetros foram escolhidos de tal forma que as distribuições a priori sejam vagas.

Na Figura 5.1 temos o histograma da distribuição a posteriori obtida para todos os parâmetros do modelo e retas verticais representando a média e o IC 95% a posteriori. Claramente observa-se que todas as densidades estão concentradas em torno dos valores usados para gerar os dados.

As Figuras 5.2 e 5.3 apresentam os histogramas dos rankings das eficiências para a melhor e a pior firma nos 8 períodos de tempo. Neste caso não necessariamente a melhor e a pior firma será a mesma ao longo do tempo. Observamos que o modelo conseguiu diferenciar bem as firmas mais eficientes das menos eficientes.

Figura 5.1: Histograma dos parâmetros - modelo de fronteira estocástica dinâmico fatorial. As retas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%, respectivamente.

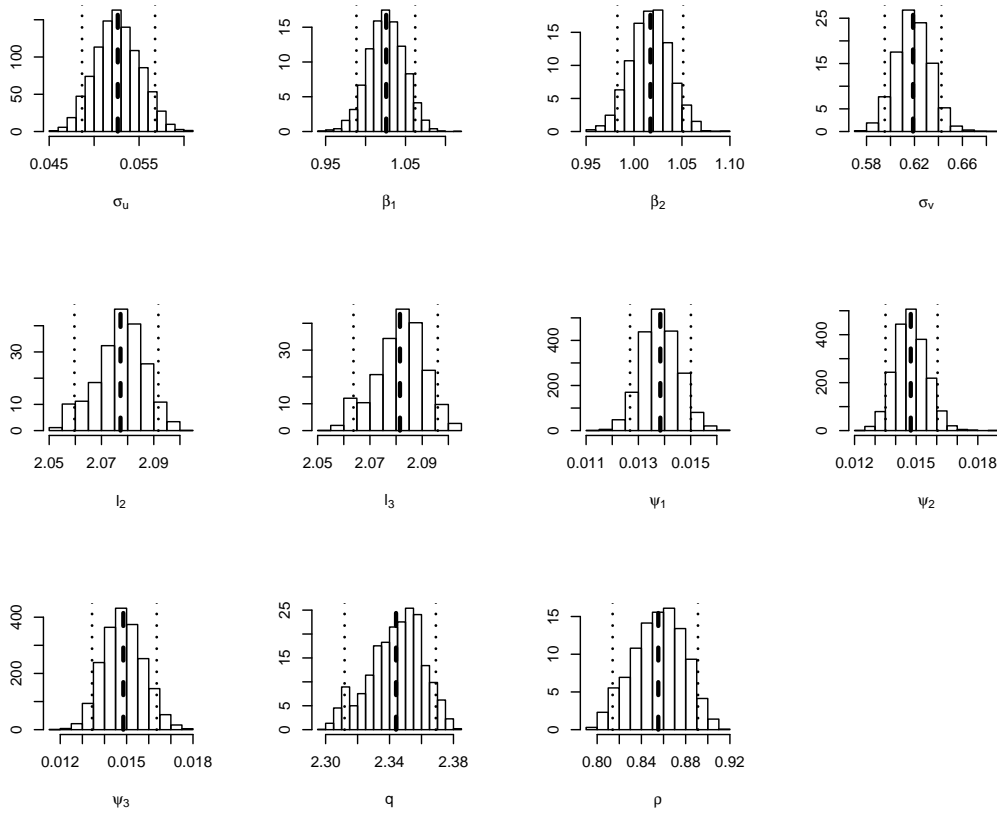


Figura 5.2: Histograma do rank das ineficiências para $t = 1, 2, 3, 4$

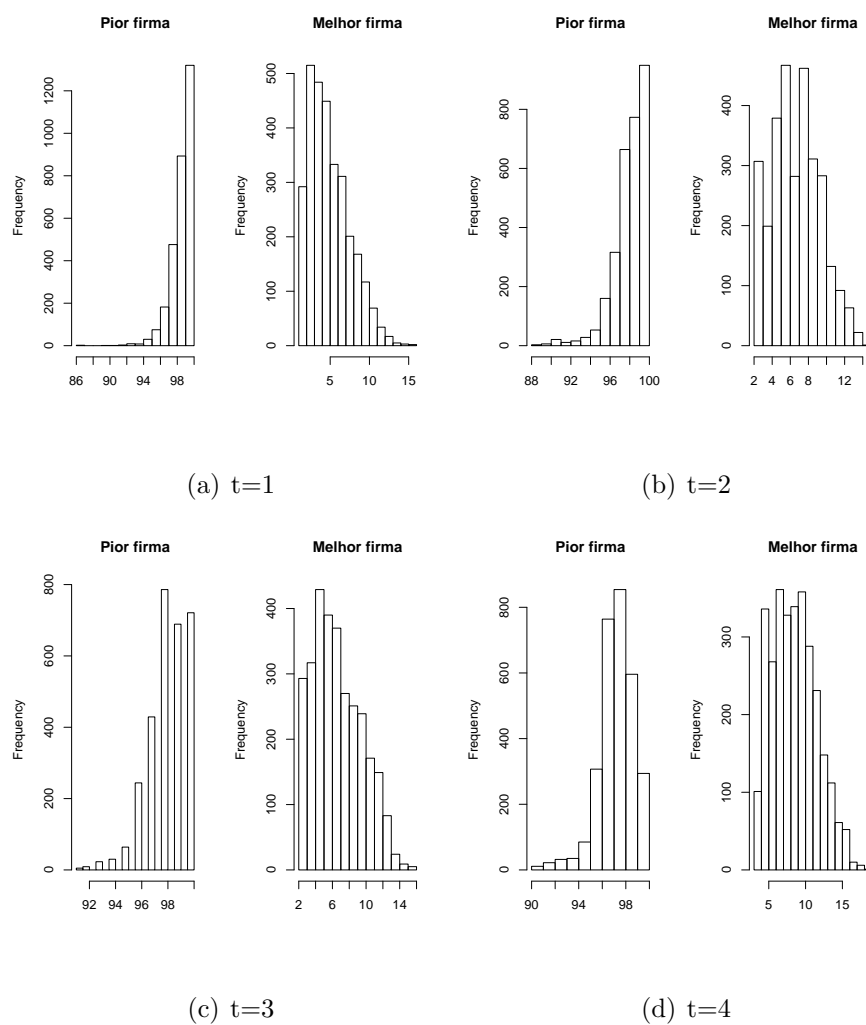
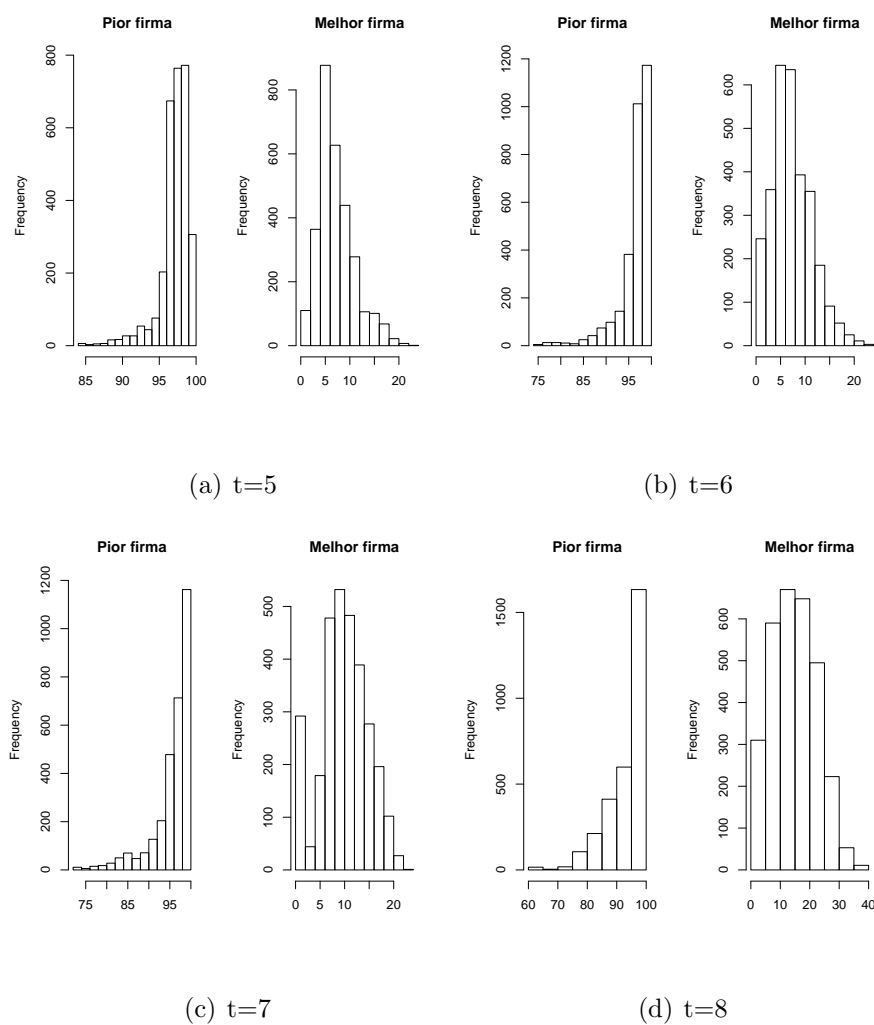


Figura 5.3: Histograma do rank das ineficiências para $t = 5, 6, 7, 8$



5.5 Aplicação a dados da Embrapa

O modelo proposto aqui é aplicado para um painel de unidades da Embrapa. O conjunto de dados refere-se a produção em pesquisas agrícolas de $N = 34$ unidades (DMUs) de pesquisas da Embrapa no ano de 1996 – 1999 ($T = 4$). O sistema de produção possui 28 produtos, que são agrupados em quatro categorias e 3 insumos. Os 3 produtos são: (1) produção científica, (2) produção de publicações técnicas e (3) produção de desenvolvimento de tecnologias, produtos e processos. E os 3 insumos definidos representam os índices de despesas com empregados, os custos operacionais e as despesas com capital, respectivamente.

Foi gerada uma cadeia de tamanho 15.000. A amostra a posteriori foi obtida utilizando um burn-in de 5.000. A verificação da convergência foi verificada através do critério de convergência de Gelman & Rubin, que ficaram próximas a 1 o que sugere que as cadeias convergiram (ver estatísticas no Apêndice).

A Figura 5.4 apresenta os histogramas da amostra da distribuição a posteriori dos parâmetros do modelo de fronteira estocástica fatorial dinâmico. As linhas tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%. Observamos que a média a posteriori do parâmetro da transformação de Box-Cox q é 0.2 e o IC de 95% não inclui o valor 0.

A Figura 5.5 apresenta os histogramas dos *rankings* das eficiências de duas unidades de pesquisas da Embrapa no período de 1996 – 1999. Observa-se que o modelo proposto capta as unidades de pesquisa que em um determinado ano é eficiente (ineficiente) e passam a ser ineficiente (eficiente). Por exemplo, a unidade de pesquisa 13, no ano de 1996 foi considerada ineficiente e a cada ano, sua eficiência foi aumentando e no último período foi uma das mais eficientes.

Figura 5.4: Histograma dos parâmetros segundo o Modelo de fronteira estocástica fatorial dinâmico para dados da Embrapa. As retas verticais tracejada e pontilhada representam a média a posteriori e o intervalo de credibilidade de 95%, respectivamente.

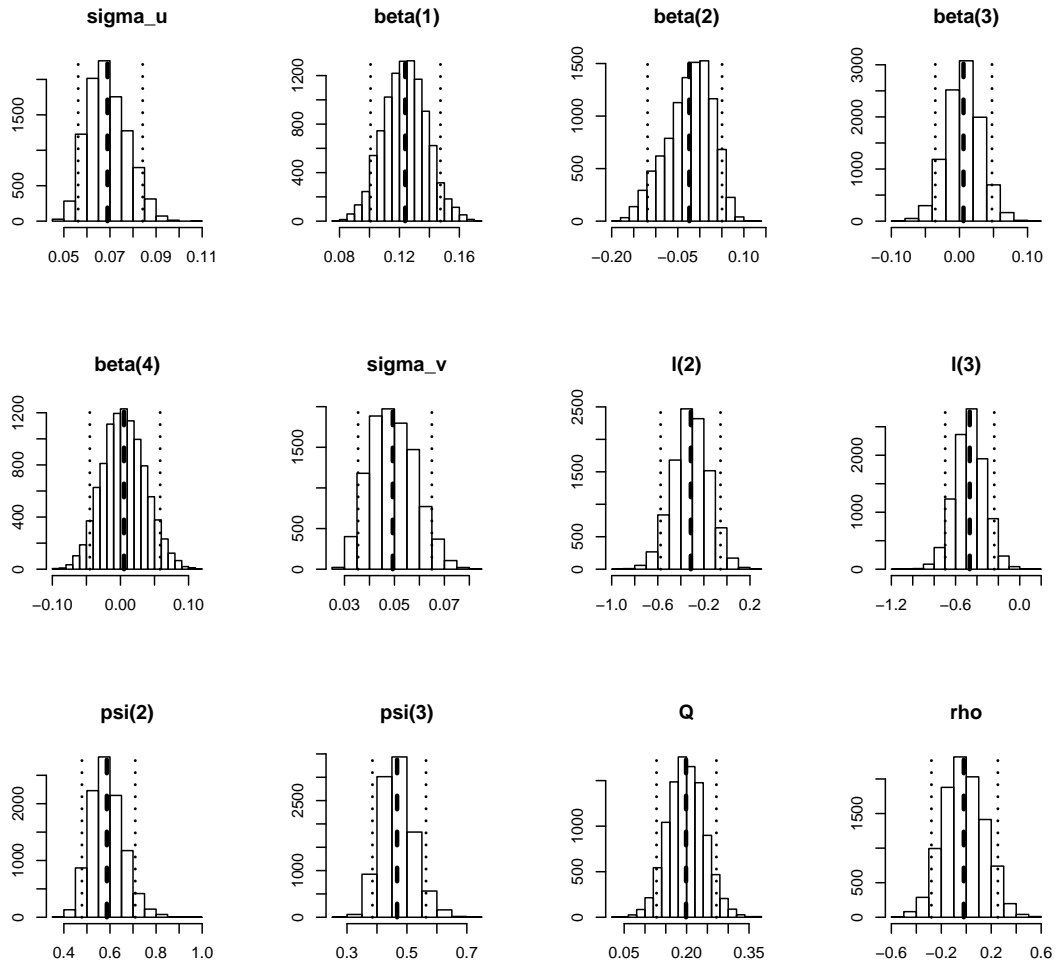
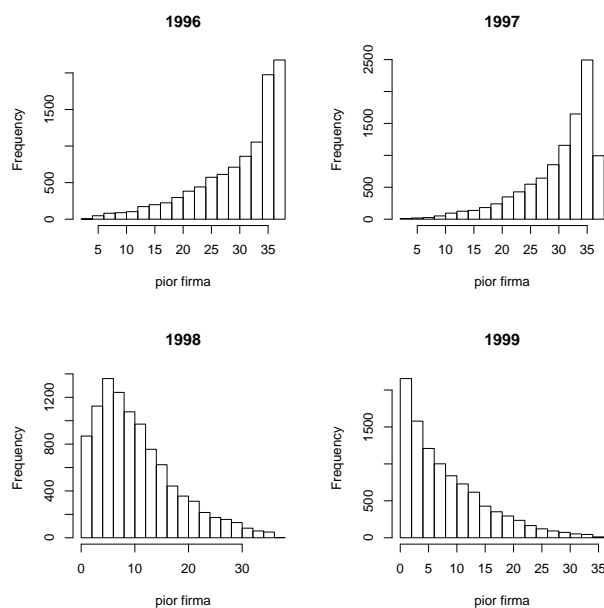
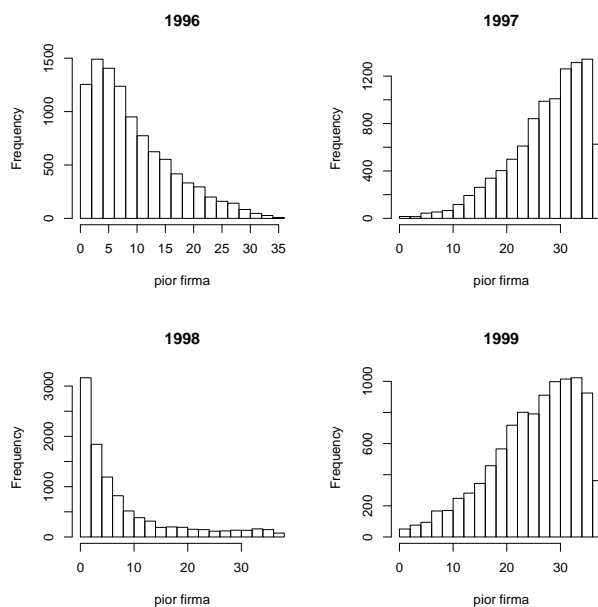


Figura 5.5: Histograma dos *rankings* da eficiência para as unidades de pesquisas da Embrapa *UD13* e *UD21* no período de 1996 – 1999.



(a) *UD13*



(b) *UD21*

Capítulo 6

Conclusões e extensões

Nesta tese foram apresentadas soluções alternativas para diversas questões relacionadas com os modelos de fronteira de produção estocástica. Como mencionado anteriormente, várias decisões importantes estão envolvidas na modelagem estocástica, destacando-se: a escolha da componente de erro associado a medida de ineficiência técnica e a seleção da função de produção, componente determinístico do modelo.

No capítulo 2 discutimos a seleção de modelos de fronteira estocástica através do critério de seleção DIC (*deviance information criterion*). Num estudo simulado mostrou-se que o critério de seleção identifica, com alta probabilidade, o verdadeiro processo gerador dos dados, quando variamos alguns parâmetros que o definem. Na aplicação envolvendo dados reais, identificou-se o modelo de fronteira de produção estocástica com distribuição log-normal para a componente de ineficiência e com função de produção tipo Cobb-Douglas, como sendo a mais apropriada. Notou-se também, que o modelo de fronteira normal – log-normal conseguiu discriminar mais efetivamente as firmas mais eficientes das menos eficientes. Neste capítulo, também, foi apresentado resultados de um estudo que compararam diferentes metodologias de simulação estocástica. Os resultados da

simulação indicam, com precisão, que o algoritmo Slice Sampling é o mais eficiente e exige um menor tempo computacional do que o Metropolis otimizado.

Em seguida, abordou-se alguns problemas que surgem na etapa da inferência em modelos de fronteira estocástica com um único produto. A função de verossimilhança desses modelos apresentam problemas teóricos e numéricos que dificultam a estimação dos parâmetros, sobretudo em amostras pequenas. A abordagem Bayesiana, apresentada no capítulo 3, pode ser pensada como uma maneira de penalizar a função de verossimilhança e, dessa forma, soluciona, praticamente, os problemas encontrados na estimação de máxima verossimilhança. Assim, podemos concluir que a priori de Jeffreys mostrou-se, pragmaticamente, bastante útil na solução deste problema. Investigações ainda precisam ser realizadas no sentido de garantir que as distribuições a posteriori obtidas a partir das distribuições a priori de Jeffreys sejam próprias.

No capítulo 4, abordou-se o problema de se estimar a eficiência de uma firma no contexto de múltiplos produtos. Desenvolveu-se um modelo de fronteira de produção estocástica para múltiplos produtos. Usa-se rudimentos de análise fatorial para se obter um único fator, sobre o qual modela-se a fronteira estocástica. A utilização de transformação na família Box-Cox, no contexto dos modelos de análise fatorial, permitiu que se preservassem propriedades econômicas fundamentais, como por exemplo: elasticidade constante de transformação (CET).

Ressalta-se que o parâmetro associado a transformação de Box-Cox é conjuntamente estimado. Assim, a modelagem proposta é capaz de calcular a eficiência técnica de uma firma, para o caso de múltiplos produtos, com vantagens comparativas em relação às alternativas disponíveis na literatura, quer do ponto de vista do tempo computacional, quer da facilidade de implementação. Os estudos simulados sugerem, claramente, que as técnicas de inferência propostas funcionam adequadamente e, portanto, podem ser utilizadas em situações práticas. A

aplicação aos dados da Embrapa mostrou que a ordenação das firmas, baseada na medida de eficiência decorrentes do modelo proposto, são similares às aquelas obtidas em que se utilizaram a outra abordagem estocástica descrita na literatura.

Por último, a tese considera o caso denominado na literatura de dados em painel. As eficiências técnicas variam no tempo e, continua-se, admitindo a situação de múltiplos produtos. No capítulo 5, o modelo de fronteira estocástica para múltiplos produtos foi estendido ao permitir variações temporais na componente de ineficiência. A variação temporal foi modelada através das ineficiências técnicas que seguem processos auto-regressivos. Está extensão parte do suposto de que cada componente u_{it} pode ser explicada por uma estrutura dinâmica e por variáveis exógenas. Note-se que este modelo é muito parametrizado, por esta razão algoritmos computacionalmente eficientes têm que ser utilizados para amostrar os parâmetros de interesse. Em particular, é proposto um esquema MCMC para fazer inferência do ponto de vista Bayesiano. O exemplo com dados simulados mostrou que os algoritmos de inferência funcionam adequadamente, podendo ser aplicados a dados reais com estrutura de painel para múltiplos produtos. Finalmente, no capítulo 5 uma aplicação a dados de produção em pesquisas agrícolas de unidades de pesquisas da Embrapa foi apresentado. Os resultados mostraram que os modelos propostos se adaptaram bem à dinâmica dos dados.

Possíveis extensões deste estudo seriam: Discutir as propriedades teóricas do capítulo 3 e modelar outras famílias comuns, como por exemplo: normal – log-normal ou normal – t-Student truncada. No capítulo 4, estender o caso de múltiplos produtos para a modelagem espacial. Realizar outras aplicações com dados clássicos da literatura de forma a confirmar, com mais precisão, o valor do modelo proposto no capítulo 5. Por ultimo, introduzir modelos de múltiplos produtos na família exponencial, enquanto modelo linear generalizado, o que de outra forma implica em fazer uma análise fatorial na família exponencial.

Referências Bibliográficas

- [1] Adams, R., Berger, A. e Sickles, R., (1999). Semiparametric approaches to stochastic panel frontiers with applications in the applicatins in the banking industry. *Journal of Business and Economic Statistics* **17**: 349-358.
- [2] Aigner, D., C. Lovell, e P. Schmidt (1977). Formulation and Estimation of Stochastic Frontier Production Function Models, *Journal of Econometrics* **6**: 21-37.
- [3] Aguilar, O. e M. West (2001) Bayesian dynamic factor models and variance matrix discounting for portfolio allocation. *Journal of Business and Economic Statistics* **18**: 338-357.
- [4] Ahn, S.C. Good, D.H. Sickles R. (2002). Estimation of long-run inefficiency levels: a dynamic frontier approach. Unpublished.
- [5] Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In: B. N. PETROV and F. CSAKI, eds. Second International Symposium on Information Theory. Budapest Akademiai Kiado, 267-281.
- [6] Azzalini, A. (1985). A Class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12**: 171-178.

- [7] Azzalini, A. e Capitanio, A. (2003). Distributions generated by perturbations of symmetry with emphasis on a multivariate skew t distribution. *Journal of the Royal Statistical Society, Series B* **65**: 367-389
- [8] Battese, G. e T. Coelli. (1992). Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis* **3(1)**: 153-169.
- [9] Battese GE, Coelli TJ. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* **20**: 325-332.
- [10] Branco, M. D. e Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis* **79**: 99-113
- [11] Broemeling, L.D.; Cook, P. (1993) Bayesian estimation of the mean of an autoregressive process. *Journal of Applied Statistics*, **20**: 25-38.
- [12] Broeck, J. van den, Koop, G., Osiewalski, J., e Steel, M. F. J. (1994). Stochastic frontier models - A Bayesian perspective, *Journal of Econometrics*, **61**: 273-303.
- [13] Box, G.E.P. and D.R. Cox (1964). An analysis of transformations. *Journal of Royal Statistical Society B*, **26**: 211-243.
- [14] Christensen, L. R. e W. H. Greene. (1976), Economies of scale in U.S. electric power generation. *Journal of Political Economy*, **84**: 655-676.
- [15] Charnes, A., W. Cooper, e E., Rhodes (1978) Measuring the efficiency of decision-making units. *European Journal of Operational Research*, **2**: 429-444.

- [16] Cornwell, C., P. Schmidt and R. Sickles. (1990). Production Frontiers with Cross-Sectional and Time- Series Variation in Efficiency Levels. *Journal of Econometrics* **46(1)**: 185-200.
- [17] De Oliveira, V., Kedem, B. e Short, D. A. (1997). Bayesian prediction of transformed Gaussian random fields. *J. Amer. Statist. Assoc.* **92**: 1422-1433.
- [18] De Oliveira, V. (2003). A note on the correlation structure of transformed gaussian random fields. *Aust. N. Z. J. Stat.* **45**: 353-366
- [19] Desli, E., Ray, S. C., Kumbhakar, S. (2003). Dynamic Stochastic Frontier Analysis. *Applied Economics Letters*, **10**: 623-626.
- [20] Domínguez-Molina , J. A., González-Farías, G. e Ramos-Quiroga, R. (2004), Skew-normality in stochastic frontier analysis in Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality, Genton, M. G., Ed., Chapman & Hall / CRC, Boca Raton, FL, 223-241.
- [21] Fernández, C. e Steel, M.F.J. (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association* **93**: 359-371.
- [22] Fernández, C., Koop, G. e Steel, M. F. J. (2000). A Bayesian Analysis of multiple-output production frontiers. *Journal of Econometrics*. **98**: 47-79
- [23] Fonceca, T. C. O., Ferreira, M. A. R. ; Migon, H. S. (2008). Objective Bayesian analysis for the Student-t regression model. *Biometrika*, **95**: 325-333.
- [24] Gelfand, A. E., Dey, D. K., Chang, H. (1992). Model determination using predictive distributions with implementation via sampling-based methods. In Bayesian statistics 4, (ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith), pp. 147-68. Oxford University Press.

- [25] Gelfand A.E., Ghosh S.K. (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika* **85**: 1-11.
- [26] Genton, M. G. (2004b). Skew-symmetric and generalized skew-elliptical distributions. In *Skew-elliptical distributions and their applications* (ed. M. G. Genton), chapter 5, 81-100. Chapman & Hall/CRC, London.
- [27] Geweke, J.F. e Zhou, G. (1996). Measuring the pricing error of the arbitrage pricing theory. *The Review of Financial Studies* **9**: 557-587.
- [28] Greene, W. H. (1990). A Gamma Distributed Stochastic Frontier Model. *Journal of Econometrics* **84**: 141-163.
- [29] Han C., Carlin B. (2001). Markov Chain Monte Carlo methods for computing Bayes factors: a comparative review. *J. Amer. Statist. Assoc* **96**: 1122-1132.
- [30] Henze, N. (1986). A Probabilistic Representation of the 'Skew-normal' Distribution. *Scand. J. Statist* **13**: 271-275.
- [31] Jeffreys, H. (1961). *Theory of Probability*, Oxford University Press, New York.
- [32] Koop, G., Osiewalski, J., Steel, M.F.J., (1997). Bayesian efficiency analysis through individual effects: hospital cost frontiers. *Journal of Econometrics* **76**: 77-105.
- [33] Koop, G., Osiewalski, J., Steel, M.F.J., (1999). The components of output growth: a stochastic frontier analysis. *Oxford Bulletin of Economics and Statistics* **61**: 455-487.
- [34] Koop G, Osiewalski J, Steel MFJ. (2000a). A stochastic frontier analysis of output level and growth in Poland and Western economies. *Economics of Planning* **33**: 185-202.

- [35] Koop G, Osiewalski J, Steel MFJ. (2000b). Measuring the sources of output growth in a panel of countries. *Journal of Business and Economic Statistics* **18**: 284-299.
- [36] Kumbhakar, S. (1987). The specification of technical and allocative inefficiency of multi-product firms in stochastic production and profit frontiers. *Journal of Quantitative Economics* **3**: 213-223.
- [37] Kumbhakar, S. (1990). Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency. *Journal of Econometrics* **46(1/2)**: 201-212.
- [38] Kumbhakar S, Ghosh S, McGuckin T. (1991). A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms. *Journal of Business and Economic Statistics* 279-286.
- [39] Kumbhakar, S. and K. Lovell. (2000). Stochastic Frontier Analysis. Cambridge: Cambridge University Press.
- [40] Lau, L., (1972). Profit functions of technologies with multiple inputs and outputs. *Review of Economics and Statistics* **54**: 281-289.
- [41] Lee, Y. and P. Schmidt. (1993). A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency. In H. Fried and K. Lovell (eds.), *The Measurement of Productive Efficiency: Techniques and Applications*. New York: Oxford University Press.
- [42] Liseo, B., Loperfido N. (2006). Default Bayesian analysis of the skew-normal distribution. *Journal of Statistical Planning and Inference* **136**: 373-389.
- [43] Lopes, H. F., West, M. (2004). Bayesian model assessment in factor analysis. *Statistica Sinica* **14**: 41-67.

- [44] Lothgren, M., (1997). Generalized stochastic frontier production models. *Economics Letters* **57**: 255-259.
- [45] Medici E., Migon H. S. (2002). Hierarchical Bayesian models applied to stochastic production frontier. *Relatorio técnico*, Departamento de Métodos Estatísticos **151**, UFRJ.
- [46] Medrano, L. A. T. (2003). Análise bayesiana de modelos de fronteira de produção estocástica. UFRJ, (Dissertação de Mestrado).
- [47] Medrano, L. A. T., Migon, H. S.(2007). Sampling Schemes for Asymmetric Models a comparative study. Relatorio técnico, Departamento de Métodos Estatísticos No. 204, UFRJ.
- [48] Medrano, L. A. T., Migon, H. S.(2008). Bayesian Stochastic Frontier: model selection via information based criteria. (submetido).
- [49] Meeusen, W., J. van den Broeck (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* **8**: 435-444.
- [50] Migon, H. S. (2006). Bayesian Hierarchical Models for Stochastic Production Frontier. *Estatística* **56**: 1-19.
- [51] Neal, R. (2003). Slice sampling. *Annals of Statistics* **31**: 705-767.
- [52] Nerlove, M., 1963, Return to scale in electricity supply, in: C. F. Christ, ed., Measurement in economics (Stanford University Press, Stanford, CA).
- [53] Pitt, M. and L. Lee. (1981). The Measurement and Sources of Technical Inefficiency in Indonesian Weaving Industry. *Journal of Development Economics* **9**: 43-64.

- [54] Powell, A., Gruen, F., (1968). The constant elasticity of transformation production frontier and linear supply system. *International Economic Review* **9**: 315-328.
- [55] Press, S.J. (1982). Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference (2nd edition). New York: Krieger.
- [56] Reis, E. A. ; Salazar, E. ; Gamerman, D. (2006). Comparison of sampling schemes for dynamic linear models. *International Statistical Review* **74(2)**: 203-214.
- [57] Schmidt, P., R. Sickles. (1984). Production Frontiers with Panel Data. *Journal of Business and Economic Statistics* **2(4)**: 367-374.
- [58] Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**: 461-464.
- [59] Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. (2002). Bayesian measure of model complexity and fit. *Journal of the Royal Statistical Society Series B* **64**: Part 3, 1-34.
- [60] Souza, G. S., Ávila, A.F.D. e Alves, E. (1999). Technical Efficiency of Production in Agricultural Research. *Scientometrics*, Budapest, **46(1)**: 141 – 160.
- [61] Tancredi, A. (2003). Accounting for Heavy Tails in Stochastic Frontier Models. *Technical Report, Dip. Scienze Statistiche, Università di Padova, Italy*.
- [62] Tsionas, E. G. (2000). Full Likelihood Inference in Normal-Gamma Stochastic Models. *Journal of Productivity Analysis* **13**: 183-205.
- [63] Tsionas, E. G. (2002). Stochastic frontier models with random coefficients. *Journal of Applied Econometrics* **17**: 121-147.

- [64] Tsionas, E. G. (2006). Inference in dynamic stochastic frontier models. *Journal of Applied Econometrics* **21**(5): 669-676.
- [65] Stevenson, R. (1980). Likelihood Functions for Generalized Stochastic Frontier Estimation. *Journal of Econometrics* **13**: 57-66.
- [66] Zellner, A., Revankar, N.S. (1970). Generalized Production Functions. *The Review of Economic Studies* **36**: 241-250.
- [67] Wahed, A., Ali, M.M., (2001). The skew-logistic distribution. *J. Statist. Res.* **35**: 71-80.

Apêndice

Critério de convergência

Para verificar a convergência das cadeias foi utilizada estatística \hat{R} (ver Gelman & Rubin, 1992, para mais detalhes) onde valores próximos de 1 indicam convergência das cadeias. Neste apêndice, estão todos os resultado das estatística \hat{R} para todos as aplicações com dados artificiais e reais utilizado nesta tese.

Tabela 1: Estatística \hat{R} de Gelman & Rubin para Seção 3.6

	Fronteira normal — normal truncada	Fronteira t-student — t-student truncada
β_1	1,03	1,04
β_2	1,02	1,04
β_3	1,04	1,02
β_4	1,04	1,03
β_5	1,02	1,02
α	1,04	1,03
λ	1,10	1,08
η		1,06

Tabela 2: Estatística \widehat{R} de Gelman & Rubin para Seção 4.7.1 - Modelo Fatorial usando a transformação de Box-Cox

	ψ_1	ψ_2	ψ_3	l_2	l_3	q
\widehat{R}	1,01	1,02	1,01	1,01	1,02	1,03

Tabela 3: Estatística \widehat{R} de Gelman & Rubin para Seção 4.7.3 - Modelo Fatorial usando a transformação de Box-Cox

	ψ_1	ψ_3	ψ_3	l_2	l_3	q	β_1	β_2	σ_v	λ
\widehat{R}	1,01	1,00	1,01	1,00	1,00	1,02	1,02	1,02	1,01	1,02

Tabela 4: Estatística \widehat{R} de Gelman & Rubin para Seção 4.8 - Modelo Fatorial usando a transformação de Box-Cox

	ψ_1	ψ_2	ψ_3	ψ_4	l_2	l_3	l_4
\widehat{R}	1,02	1,03	1,03	1,04	1,01	1,00	1,01
	β_1	β_2	β_3	β_4	σ_v	λ	q
\widehat{R}	1,00	1,01	1,03	1,01	1,02	1,05	1,07

Tabela 5: Estatística \widehat{R} de Gelman & Rubin para Seção 5.4 - modelo de fronteira estocástica dinâmico fatorial.

	ψ_1	ψ_2	ψ_3	l_2	l_3	q
\widehat{R}	1,05	1,02	1,02	1,03	1,02	1,04
	β_1	β_2	σ_v	σ_u	ρ	
\widehat{R}	1,01	1,03	1,10	1,08	1,04	

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)