

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Marcus Vinicius Pereira de Souza

**Uma Abordagem Bayesiana para o Cálculo
dos Custos Operacionais Eficientes das
Distribuidoras de Energia Elétrica**

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio como parte dos requisitos parciais para obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Prof. Reinaldo Castro Souza

Rio de Janeiro

Maio de 2008

Livros Grátis

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

Milhares de livros grátis para download.



Marcus Vinicius Pereira de Souza

**Uma Abordagem Bayesiana para o Cálculo
dos Custos Operacionais Eficientes das
Distribuidoras de Energia Elétrica**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Reinaldo Castro Souza
Orientador

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Madiagne Diallo
Co-orientador

Departamento de Engenharia Industrial

Prof. José Francisco Moreira Pessanha
CEPEL

Prof. Tufi Machado Soares
Universidade Federal de Juiz de Fora

Profa. Mariane Branco Alves
UERJ

Prof. Marcos Azevedo da Silveira
Departamento de Engenharia Elétrica – PUC-Rio

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico

Rio de Janeiro, 13 de maio de 2008

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Marcus Vinicius Pereira de Souza

Nascido em Andrelândia-MG, 1972. Graduou-se em Engenharia Elétrica (1999) pela Universidade Federal de Juiz de Fora (UFJF). Mestre em Engenharia Aeronáutica e Mecânica, Área de Produção (2002) pelo Instituto Tecnológico de Aeronáutica (ITA). Suas pesquisas de interesse incluem as áreas de redes neurais artificiais, inteligência artificial, estatística aplicada e pesquisa operacional.

Ficha Catalográfica

Souza, Marcus Vinicius Pereira de

Uma abordagem bayesiana para o cálculo dos custos operacionais eficientes das distribuidoras de energia elétrica / Marcus Vinicius Pereira de Souza ; orientador: Reinaldo Castro Souza. – 2008.

153 f. ; 30 cm

Tese (Doutorado em Engenharia Elétrica) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Análise envoltória de dados. 3. Fronteiras estocásticas. 4. Rede neural de Kohonen. 5. Regulação econômica. I. Souza, Reinaldo Castro. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

*Esta Tese é dedicada à memória do meu
amado Pai, Sebastião Teixeira de Souza, que
certamente está num plano maior
intercedendo por mim; e também para minha
querida Mãe Elizabeth Meer Pereira de
Souza, essencial em todos os momentos da
minha vida.*

Agradecimentos

A Deus, à Nossa Senhora Aparecida, aos Mestres da Grande Fraternidade Branca e a todos os Gloriosos Santos por mais esta conquista.

Consigno meus sinceros agradecimentos e respeito ao professor Dr. Reinaldo Castro Souza pela oportunidade; orientando-me com paciência, dedicação e incentivo em prol desta tese.

Também quero externar o meu apreço sincero ao professor Dr. José Francisco Moreira Pessanha que aconselhou-me e encorajou-me durante toda essa trajetória.

Ao grande amigo e professor Dr. Madiagne Diallo, pelo apoio moral e constante estímulo na concretização desta importante etapa da minha vida.

Quero ainda expressar minha gratidão à Helaine Aparecida da Silva pelos conselhos, incentivos e sua valiosa ajuda na efetivação desta tese.

Muitas outras pessoas contribuíram para o sucesso deste trabalho, em especial: Airam Regina de Aquino Martins, Alexandra Mello Schmidt, Alexandre Bessa dos Santos, Alexandre Zanini, Álvaro Veiga, Ana Paiva, André Luis Marques Marcato, Arildo de Jesus Batista, Armando Zeferino Milioni, Cairo Lúcio Nascimento Jr., Carlos Galeno Ladeira Pereira, Célia Maria de Souza Pereira, Cristiano Augusto C. Fernandes, Eduardo Eliphaz Pereira, Ernesto Fagundes Neto, Eurides Moura, Fátima Mendes de Carvalho, Felipe Fortes Carvalho Silva, Fernando César Coelli, Fernando Neves Breseguello, Gisele Teixeira Rocha, Hélio Francisco da Silva, Isnarde Antônio Ernesto, Jacqueline Furtado Vital, José Luiz Bellini Leite, José Marcelino de Jesus, Juliana Garcia Cespedes, Lúcia Helena Lima Miranda e Silva, Lúcia Margareth Pereira, Luiz Tarciso de Andrade, Marcia Brum Guerra Gomes, Márcia Doná, Márcio Vicente Rizzo, Maria Alcina Portes, Maria Aparecida Gonçalves, Mariane Antunes Maia, Marley M. B. Rebuzzi Vellasco, Mischel Carmen Neyra Belderrain, Mônica Barros, Nair de Carvalho Pereira, Neimar Fernandes, Paula Adriana Soares, Pedro Gomes Barbosa, Ricardo Tanscheit, Sebastião Marcos M. P. Vasconcelos, Tatiane Cristina da Silva, Vander Menengoy da Costa, Wilma Galgane.

À **PUC-Rio**, pelo suporte financeiro do projeto.

Finalmente, mas não menos importante, desejo registrar a minha admiração a todos que me honraram apresentando críticas e sugestões concernentes ao meu trabalho, em particular: Elizabeth Meer Pereira de Souza, Dr. Marcos Azevedo da Silveira, Dra. Mariane Branco Alves e Dr. Tufi Machado Soares.

Resumo

Souza, Marcus Vinicius Pereira de; Souza, Reinaldo Castro (Orientador). **Uma Abordagem Bayesiana para o Cálculo dos Custos Operacionais Eficientes das Distribuidoras de Energia Elétrica**. Rio de Janeiro, 2008. 153p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro

Esta tese apresenta os principais resultados de medidas de eficiência dos custos operacionais de 60 distribuidoras brasileiras de energia elétrica. Baseado no esquema *yardstick competition*, foi utilizado uma Rede Neural de Kohonen (KNN) para identificar grupos de empresas similares. Os resultados obtidos pela KNN não são determinísticos, visto que os pesos sinápticos da rede são inicializados aleatoriamente. Então, é realizada uma simulação de Monte Carlo para encontrar os *clusters* mais frequentes. As medidas foram obtidas por modelos DEA (*input oriented*, com e sem restrições aos pesos) e modelos Bayesianos e frequencistas de fronteira estocástica (utilizando as funções Cobb-Douglas e Translog). Em todos os modelos, DEA e SFA, a única variável *input* refere-se ao custo operacional (OPEX). Os índices de eficiência destes modelos representam a potencial redução destes custos de acordo com cada concessionária avaliada. Os *outputs* são os *cost drivers* da variável OPEX: número de unidades consumidoras (uma *proxy* da quantidade de serviço), montante de energia distribuída (uma *proxy* do produto total) e a extensão da rede de distribuição (uma *proxy* da dispersão dos consumidores na área de concessão). Finalmente, vale registrar que estas técnicas podem mitigar a assimetria de informação e aprimorar a habilidade do agente regulador em comparar os desempenhos das distribuidoras em ambientes de regulação incentivada.

Palavras-chave

Análise Envoltória de Dados; Fronteiras Estocásticas, Rede Neural de Kohonen, Regulação Econômica.

Abstract

Souza, Marcus Vinicius Pereira de; Souza, Reinaldo Castro (Advisor). **A Bayesian Approach to Estimate the Efficient Operational Costs of Electrical Energy Utilities**. Rio de Janeiro, 2008. 153p. Doctorate Thesis – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro

This thesis presents the main results of the cost efficiency scores of 60 Brazilian electricity distribution utilities. Based on yardstick competition scheme, it was applied a Kohonen Neural Networks (KNN) to identify and to group the similar utilities. The KNN results are not deterministic, since the estimated weights are randomly initialized. Thus, a Monte Carlo simulation was used in order to find the most frequent clusters. Therefore was examined the use of the DEA methodology (input oriented, with and without weight constraints) and Bayesian and non-Bayesian Stochastic Frontier Analysis (centered on a Cobb-Douglas and Translog cost functions) to evaluate the cost efficiency scores of electricity distribution utilities. In both models the only input variable is operational cost (OPEX). The efficiency measures from these models reflect the potential of the reduction of operational costs of each utility. The outputs are the cost-drivers of the OPEX: the number of customers (a proxy for the amount of service), the total electric power supplied (a proxy for the amount of product delivered) and the distribution network size (a proxy of the customers scattering in the operating territory of each distribution utility). Finally, it is important to mention that these techniques can reduce the information asymmetry to improve the regulator's skill to compare the performance of the utilities in incentive regulation environments.

Keywords

Data Envelopment Analysis; Stochastic Frontier Analysis, Kohonen Neural Networks, Economic Regulation.

Sumário

1. Introdução	18
1.1. Considerações gerais	18
1.2. Revisão bibliográfica	23
1.3. Objetivos	24
1.4. Estrutura da tese	26
2. Redes Neurais Auto-Organizáveis	29
2.1. Introdução	29
2.2. Componentes de uma tarefa de <i>clustering</i>	29
2.3. Aprendizado por competição	31
2.4. Rede Neural de Kohonen (RNK)	32
2.5. Exemplo ilustrativo da Rede Neural de Kohonen	34
2.6. Rede Neural de Kohonen via simulação de Monte Carlo (MCRNK)	38
3. Análise Envoltória de Dados	40
3.1. Introdução	40
3.2. Conjunto de Possibilidades de Produção	42
3.3. Conceito de eficiência	44
3.4. <i>Benchmarks</i>	46
3.5. Retornos de escala	47
3.6. Modelos DEA	48
3.6.1. Modelo DEA CCR <i>input-oriented</i>	49
3.6.2. Modelo DEA BCC <i>input-oriented</i>	53
3.6.3. Outros modelos DEA	56
3.6.4. Seleção de variáveis <i>inputs</i> e <i>outputs</i>	57
3.6.5. Técnicas baseadas em Restrições aos Pesos	58
3.7. <i>Adjusted contingent restrictions on weights</i>	60
4. Conceitos Básicos de Estatística Bayesiana e Simulação Estocástica	62
4.1. Elementos de inferência Bayesiana	62
4.2. Distribuição a priori	63
4.3. Distribuição amostral (Função de Verossimilhança)	64
4.4. Distribuição a posteriori	65
4.5. Exemplo	65
4.6. Obtenção de resumos de interesse através de simulação	68
4.6.1. Integração via Monte Carlo	69
4.7. Monte Carlo via Cadeias de Markov (MCMC)	72
4.7.1. Algoritmo de Metropolis-Hastings	74
4.7.2. Amostrador de Gibbs	77
4.7.3. Monitoração da convergência	82
5. Modelos Econométricos	84
5.1. Fronteira Estocástica de Produção	84
5.2. Fronteira Estocástica de Custos	89
5.3. Fronteira Estocástica sob o paradigma Bayesiano	91
5.3.1. Modelos Bayesianos para Fronteira Estocástica	92

5.3.2. <i>Varying Efficiency Distribution (VED) models</i>	93
5.4. Inferência Bayesiana usando Amostrador de Gibbs	95
6. <i>Clustering</i> : Definindo Grupos Estratégicos	98
6.1. Introdução	98
6.2. Análise de dados	99
6.3. Resultados obtidos via <i>cluster analysis</i>	99
7. Mensuração da Eficiência Relativa utilizando Modelos DEA	105
7.1. Introdução	105
7.2. Resultados DEA e discussões	105
8. Estimação da Eficiência Técnica usando Fronteiras Estocásticas	110
8.1. Introdução	110
8.2. Resultados dos modelos econométricos com enfoque clássico	110
8.3. Resultados dos modelos econométricos com enfoque Bayesiano	115
9. Análise Multivariada de Dados	123
9.1. Introdução	123
9.2. Matriz de correlação de Pearson e Spearman	123
9.3. Análise Fatorial	127
10. Conclusões e Perspectivas	133
10.1 Conclusões	133
10.2 Perspectivas	135
Referências bibliográficas	138
Anexo 1 - Conjunto de Dados	149
Anexo 2 - Modelo Bayesiano	150

Lista de figuras

Figura 1.1: Processo de revisão tarifária [4]	19
Figura 2.1: Etapas de um processo de <i>clustering</i>	30
Figura 2.2: Rede de Kohonen com as unidades de saída organizadas como grade bi-dimensional	33
Figura 2.3: Disposição bi-dimensional das entradas	35
Figura 2.4: Disposição das unidades no espaço das entradas	36
Figura 2.5: Resultado do processo de identificação dos <i>clusters</i>	38
Figura 3.1: Comparação de DEA e regressão [18]	41
Figura 3.2: Eficiências Técnica, Alocativa e Econômica	45
Figura 3.3: Envoltória determinada pelo modelo CCR <i>input-oriented</i> [18]	52
Figura 3.4: Envoltória determinada pelos modelos CCR e BCC (<i>input-oriented</i>) [18]	55
Figura 4.1: Densidades a priori do parâmetro θ	66
Figura 4.2: Função de Verossimilhança para θ	67
Figura 4.3: Densidades a posteriori do parâmetro θ	68
Figura 4.4: <i>Scatterplot</i> dos valores simulados	71
Figura 4.5: Valores simulados através do algoritmo Metropolis-Hastings	76
Figura 4.6: Algoritmo Metropolis-Hastings na geração da distribuição de Cauchy	77
Figura 4.7: Dados simulados para λ e β utilizando o algoritmo de Gibbs	80
Figura 4.8: Resultado das médias amostrais para o parâmetro λ	80
Figura 4.9: Distribuição marginal estimada de $p(x)$	81
Figura 4.10: Gráficos de autocorrelação amostral dos parâmetros λ e β	83
Figura 5.1: Densidades de probabilidade da combinação de $v_j + u_j$	91
Figura 6.1: Resultado de um processo de <i>clustering</i> em uma simulação	100
Figura 6.2: Histograma dos <i>clusters</i>	101
Figura 6.3: Histograma dos 13 <i>clusters</i> mais frequentes	103
Figura 7.1: <i>Scatterplot Matrix</i> dos índices de eficiência (Modelos DEA)	108
Figura 8.1: <i>Boxplot</i> dos índices de eficiência (Modelos SFA Clássica)	114
Figura 8.2: Representação gráfica dos índices de eficiência (Modelos Bayesianos)	119

Figura 8.3: Estimativa pontual dos coeficientes de regressão para diferentes valores de r^*	122
Figura 9.1: Coeficientes de correlação de Pearson e Spearman	125
Figura 9.2: Coeficientes de correlação de Pearson e Spearman	126
Figura 9.3: Coeficientes de correlação de Pearson e Spearman	127
Figura 9.4: Gráfico qui-quadrado (Q-Q <i>plot</i>)	128
Figura 9.5: Gráfico de dispersão dos <i>loadings</i> (sem rotação ortogonal).	130
Figura 9.6: Gráfico de dispersão dos <i>loadings</i> (com rotação ortogonal).	130
Figura 9.7: Gráfico de dispersão de objetos segundo os <i>factor scores</i>	131

Lista de tabelas

Tabela 2.1 - Centro dos <i>clusters</i> com seus respectivos pontos de treinamento	34
Tabela 3.1 - Representação algébrica de retornos de escala	47
Tabela 3.2 - Relação entre retornos de escala e elasticidade total da produção	48
Tabela 6.2 - <i>Clusters</i> obtidos pela Rede de Kohonen em uma simulação	101
Tabela 6.3 - Sequência ordenada dos <i>clusters</i> mais frequentes	102
Tabela 6.4 - Agrupamentos finais	104
Tabela 7.1 - Variáveis consideradas	105
Tabela 7.2 - Índices de eficiência θ_j^{DEA}	107
Tabela 7.3 - Estatística descritiva dos modelos DEA	108
Tabela 8.1 - Índices de eficiência $\theta_j^{SFA-C/HN}$	112
Tabela 8.2 - Índices de eficiência $\theta_j^{SFA-C/NT}$	113
Tabela 8.3 - Índices de eficiência r_j^{SFA-B}	117
Tabela 8.4 - Índices de eficiência r_j^{SFA-B}	118
Tabela 8.5 - Estimativa dos coeficientes de regressão e parâmetros do modelo de fronteira estocástica (enfoque Bayesiano)	120
Tabela 8.6 - Índices de eficiência $r_j^{SFA-B/NT-TL}$; ($r^* = 0,95$)	121
Tabela 9.1 - Coeficientes de Correlação de Pearson	124
Tabela 9.2 - Coeficientes de Correlação de Spearman	124
Tabela 9.3 - Comparação entre os <i>factor loadings</i> não rotacionados e rotacionados	129
Tabela 9.4 - <i>Ranking</i> das concessionárias	132

Símbolos, Acrônimos e Abreviaturas

d.p. - desvio-padrão;

e - vetor de 1's;

e - Elasticidade Total;

exp - função exponencial;

$\text{Exp}(\cdot)$ - distribuição exponencial com parâmetro (\cdot) ;

$E(X)$ - valor esperado da variável aleatória X ;

$f(x|\theta)$ - densidade de X , condicional ao parâmetro θ ;

$h(\cdot)$ - função Translog (TL) ou Cobb-Douglas (CD);

i - indexador de *inputs*, $(i=1, \dots, m)$;

i.i.d - independentes e identicamente distribuídos;

j - indexador de DMU's, $(j=1, \dots, n)$;

j_0 - indexador da DMU que está sendo analisada;

jt - indexador de DMU's eficientes;

l - função de verossimilhança;

ln - logaritmo natural;

L - logaritmo natural da função de verossimilhança;

m - número de *inputs*;

max(\cdot) - máximo de (\cdot) ;

min(\cdot) - mínimo de (\cdot) ;

n - número de DMU's;

$\text{Poi}(\cdot)$ - distribuição de Poisson com parâmetro (\cdot) ;

r - indexador de *outputs*, $(r=1, \dots, s)$;

\mathbf{R}_*^+ - Conjunto dos números Reais inteiros positivos;

\mathbf{R}_+^s - Conjunto dos números Reais inteiros positivos de dimensão s ;

\mathbf{R}_+^m - Conjunto dos números Reais inteiros positivos de dimensão m ;

s - número de *outputs* (DEA);

T - Conjunto de Possibilidades de Produção;

u_r - peso atribuído ao r -ésimo *output*, modelos (DEA) *input-oriented*;

v_i - peso atribuído ao i -ésimo *input*, modelos (DEA) *input-oriented*;
 \mathbf{W}_i^k - vetor de pesos sinápticos do neurônio (i) no instante (k);
 X_j = vetor de variáveis aleatórias (*inputs*) da DMU j com dimensão $1 \times m$;
 \mathbf{X} - matriz de *inputs* observados de dimensão $n \times m$;
 \mathbf{x}_j = vetor de *inputs* observados da DMU j com dimensão $1 \times m$;
 \mathbf{x} = vetor de realizações;
 Y_j = vetor de variáveis aleatórias (*outputs*) da DMU j com dimensão $1 \times s$;
 \mathbf{Y} - matriz de *outputs* observados de dimensão $n \times s$;
 \mathbf{y}_j = vetor de *outputs* observados da DMU j com dimensão $1 \times s$;
 \mathbf{y} = vetor de realizações;
 $\Gamma(a,b)$ ou Gama(a,b) - distribuição Gama com parâmetros a e b ;
 $Unif(a,b)$ - distribuição Uniforme com parâmetros a e b ;
 $\pi(\cdot)$ - distribuição a posteriori;
 $[\cdot]^T$ - transposto de $[\cdot]$;
 β - coeficientes da fronteira estocástica;
 θ - índice de eficiência no modelo DEA CCR e vetor de parâmetros (na abordagem Bayesiana) dependendo do contexto;
 θ^* - valor ótimo do índice de eficiência no modelo CCR *input-oriented*;
 $\phi(\cdot)$ - função de densidade da Normal-padrão;
 $\Phi(\cdot)$ - função de distribuição acumulada da Normal-padrão;
 ε - infinitésimo positivo ou erro composto dependendo do contexto;
 \prod - produto;

ANEEL - Agência Nacional de Energia Elétrica;
BCC - Banker, Charnes e Cooper;
CB - Cobb-Douglas;
CCR - Charnes, Cooper e Rhodes;
CED - *Common Efficiency Distribution*;
DEA - *Data Envelopment Analysis*;
DMU - *Decision Making Unit*;
EA - Eficiência Alocativa;

EE - Eficiência Econômica;

EEF - Equilíbrio Econômico Financeiro;

ET - Eficiência Técnica;

HN - Distribuição *Half-Normal*;

IRT - Índice de Reajuste Tarifário;

LINDO - *Linear Interactive and Discrete Optimizer*;

MCMC – Monte Carlo via Cadeias de Markov;

MCRNK – Redes Neurais de Kohonen via simulação de Monte Carlo;

NT - Distribuição Normal-Truncada;

PL - Programação Linear;

PPL - Problema de Programação Linear;

RCE - Retorno Constante de Escala;

RNA - Redes Neurais Artificiais;

RVE - Retorno Variável de Escala;

SEB - Sistema Elétrico Brasileiro;

SFA – *Stochastic Frontier Analysis*;

TL - Translog;

VED - *Varying Efficiency Distribution*;

WinBUGS - *Bayesian inference Using Gibbs Sampling for Windows*.

"Bem-aventurado o homem que acha Sabedoria, e o homem que adquire
Conhecimento."
(Provérbios 3-13)

“Ninguém ignora tudo, ninguém sabe tudo. Por isso, aprendemos sempre”.

(Paulo Freire)

1

Introdução

1.1

Considerações gerais

O Setor Elétrico Brasileiro (SEB) teve por muitos anos características de indústria em monopólio. Nos meados da década de 90, este setor passou por uma profunda reestruturação cujos objetivos incluem o aumento da eficiência das atividades de distribuição e transmissão, e a introdução de um mercado competitivo nas atividades de comercialização da energia elétrica.

A Agência Nacional de Energia Elétrica (ANEEL), encarregada da regulação do setor elétrico brasileiro, cumpre tanto regular o serviço público de distribuição de energia elétrica, quanto o processo de revisão tarifária.

A partir de 1995, a União estabeleceu contratos de concessão com 61 das 64 empresas de distribuição de energia elétrica. Entre outras exigências, esses contratos definem os mecanismos para a alteração das tarifas de energia elétrica, quais sejam: i) reajuste tarifário anual; ii) revisão tarifária extraordinária; iii) revisão tarifária periódica, que ocorre, em média, a cada 4 anos [6].

No ano de realização da revisão tarifária periódica, em acordo com a ANEEL, as tarifas são reposicionadas em um nível compatível com a cobertura dos custos operacionais e com a remuneração adequada dos investimentos realizados pelas distribuidoras e reconhecidos pela agência reguladora.

Tal revisão tarifária periódica objetiva garantir a manutenção do Equilíbrio Econômico-Financeiro (EEF) da concessão. Levando-se em conta a corrosão do processo inflacionário, no período entre duas revisões tarifárias periódicas, tem-se o reajuste tarifário anual, (ii), pelo Índice de Reajuste Tarifário (IRT). Este índice é definido matematicamente pela Norma Técnica NT97 [3] como:

$$\text{IRT} = \frac{\text{VPA}_1}{\text{RA}_0} + \frac{\text{VPB}_0(\text{IGPM} - \text{X})}{\text{RA}_0} \quad (1.1.1)$$

onde VPA_1 é a parcela referente aos custos não-gerenciáveis pela distribuidora (compras de energia e encargos setoriais) na data do reajuste, RA_0 é a receita anual calculada considerando-se as tarifas vigentes na data de referência anterior e o mercado de referência, e VPB_0 é a parcela relativa aos custos gerenciáveis pela distribuidora (mão-de-obra, serviços de terceiros, material, depreciação e remuneração adequada dos ativos investidos e do capital de giro) na data de referência anterior ($VPB_0 = RA_0 - VPA_0$).

De acordo com (1.1.1), os custos não gerenciáveis (VPA) são repassados integralmente às tarifas finais, enquanto a parcela relativa aos custos gerenciáveis (VPB) é atualizada pelo IGP-M, descontado do Fator X. Este fator incide apenas sobre os custos gerenciáveis e constitui o mecanismo que permite repassar aos consumidores, por meio da redução das tarifas, os ganhos de produtividade das distribuidoras de energia elétrica.

Para ilustrar estes conceitos, a Figura 1.1 apresenta os processos de revisão tarifária e o reajuste tarifário. Para simplificar, admite-se que as variações do IGP-M sejam nulas.

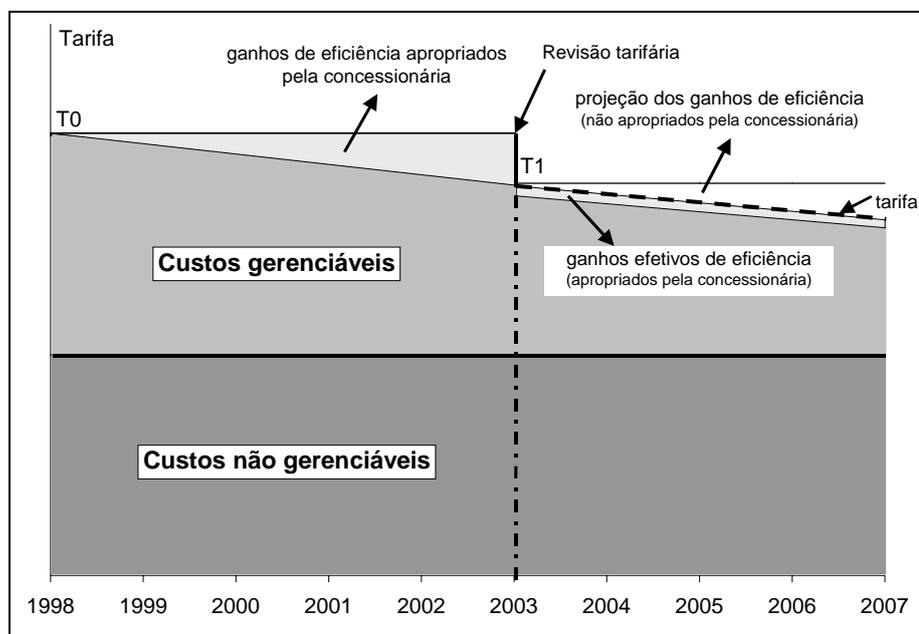


Figura 1.1: Processo de revisão tarifária [4].

A tarifa inicial T_0 , fixada no contrato de concessão, permanece constante (em termos reais) até a primeira revisão tarifária (2003). Esta tarifa garante o EEF da concessão no primeiro período tarifário (1998-2003), cobrindo os custos não gerenciáveis pela distribuidora (compras de energia e encargos setoriais) e os custos gerenciáveis pela distribuidora (mão-de-obra, serviços de terceiros, material, depreciação e remuneração adequada dos ativos investidos e do capital de giro). Neste período inicial os ganhos de produtividade decorrentes do crescimento do mercado e da redução dos custos operacionais alcançadas pela gestão eficiente da concessionária, conforme indicado pelo triângulo cinza na Figura 1.1, são integralmente apropriados pela concessionária, aumentando sua remuneração.

No final do primeiro período tarifário acontece a primeira revisão tarifária periódica (2003), a qual transcorre em duas etapas. Primeiro, o regulador reposiciona a tarifa em um nível (T_1) compatível com a receita necessária para atender as condições eficientes de manutenção e operação da atividade de distribuição, bem como uma taxa de retorno considerada adequada ao risco do negócio e que remunere o capital investido. Em seguida, a partir dos ganhos de eficiência projetados ou metas de produtividade para o segundo período tarifário (2004-2007) o regulador define o Fator X, conforme indicado em (1.1.1).

Na distribuição de energia elétrica os ganhos de produtividade decorrem principalmente do crescimento do mercado atendido pela concessionária, tanto pelo maior consumo das unidades consumidoras existentes como pela ligação de novas unidades. Em função da presença de economias de escala, a expansão do mercado é atendida com custos incrementais decrescentes, resultando em um ganho de produtividade para a distribuidora que não decorre de uma gestão mais eficiente da empresa e que, portanto, deve ser repassado para os consumidores com a finalidade de promover a modicidade tarifária.

Examinando a Figura 1.1, no segundo período tarifário (2004-2007) a tarifa segue a trajetória decrescente representada pela linha tracejada, cobrindo o repasse integral dos custos não gerenciáveis e uma parte dos custos gerenciáveis que se reduz anualmente em função do seu reajuste pelo IGP-M descontado do Fator X.

O resultado deste mecanismo é que a projeção dos ganhos de produtividade decorrentes do crescimento do mercado no período 2004/2007,

conforme indicado pelo triângulo em branco na Figura 1.1, não são apropriados pela concessionária, mas compartilhados com os consumidores.

No segundo período tarifário, a aplicação efetiva do Fator X reduz a parcela dos ganhos de eficiência que podem ser apropriados pela concessionária, limitando-os aos ganhos que ultrapassem as metas de produtividade definidas na data da revisão tarifária, conforme indicado na Figura 1.1, pela estreita faixa em cinza. Assim a concessionária é incentivada a explorar seu potencial de eficiência, e quanto maior for a eficiência da concessionária maior será seu benefício, dado pela diferença entre a meta de produtividade e o custo efetivamente realizado.

Conforme pode-se ver, o cálculo do Fator X desempenha um papel elementar para efeito do aperfeiçoamento regulatório, pois ele serve para a repartição dos ganhos de produtividade entre acionistas e consumidores. Neste cenário, é interessante citar a metodologia proposta por Zanini [97], para o cálculo do Fator X, que consiste na seguinte decomposição: $\mathbf{X} = \mathbf{X}_E + \mathbf{X}_O$, onde a componente \mathbf{X}_E determina a eficiência econômica em relação às mudanças tecnológicas do setor, refletindo, assim, os ganhos esperados de produtividade da concessionária em virtude do crescimento do seu segmento de mercado. A componente \mathbf{X}_O está diretamente relacionada com o nível de eficiência de uma determinada empresa. Vale lembrar que uma empresa é eficiente se consegue converter, de forma ótima, seus insumos em produtos. Pelo exposto, não é difícil conjecturar que uma determinada empresa só consegue atingir níveis mais acentuados de produtividade caso minimize seus custos operacionais (custos de operação, manutenção, administração e gestão comercial). Já a Resolução ANEEL Nº 55/2004 [5], define o Fator X como sendo a combinação de três componentes (\mathbf{X}_E , \mathbf{X}_A e \mathbf{X}_C), conforme a expressão abaixo:

$$\mathbf{X} = (\mathbf{X}_E + \mathbf{X}_C)(\text{IGPM} - \mathbf{X}_A) + \mathbf{X}_A \quad (1.1.2)$$

A componente \mathbf{X}_A reflete o efeito da aplicação do IPCA na componente mão-de-obra do VPB da concessionária. A componente \mathbf{X}_C relaciona-se com a qualidade percebida pelos consumidores atendidos pela concessionária. A última componente, \mathbf{X}_E , é a mais importante, sendo definida com base no método do fluxo de caixa descontado do tipo *forward looking* de maneira a igualar o valor

presente líquido do fluxo de caixa da concessionária no período tarifário, acrescido do valor residual, com o valor dos ativos da concessionária no início do período tarifário:

$$A_0 = \sum_{t=1}^N \left[\frac{(\mathbf{RO}_t \cdot (1 - \mathbf{X}_E)^{t-1} - \mathbf{T}_t - \mathbf{OM}_t - \mathbf{d}_t)(1 - \mathbf{g}) + \mathbf{d}_t - \mathbf{I}_t}{(1 + \mathbf{r}_{\mathbf{WACC}})^t} \right] + \frac{A_N}{(1 + \mathbf{r}_{\mathbf{WACC}})^N} \quad (1.1.3)$$

onde:

- N é o período em anos entre duas revisões tarifárias;
- A_0 é o valor dos ativos da concessionária na data da revisão tarifária;
- A_N é o valor dos ativos da concessionária no final do período tarifário;
- \mathbf{g} são as alíquotas do imposto de renda e da Contribuição Social sobre o Lucro Líquido (CSLL);
- $\mathbf{r}_{\mathbf{WACC}}$ ¹ é o custo médio ponderado do capital;
- \mathbf{RO}_t é a receita operacional da concessionária;
- \mathbf{T}_t são os valores dos tributos (PIS/PASEP, COFINS e P&D);
- \mathbf{OM}_t são os custos operacionais da concessionária (OPEX);
- \mathbf{I}_t é o montante de investimentos realizados;
- \mathbf{d}_t é a depreciação, todos referentes ao ano t .

As parcelas que compõem o fluxo de caixa em (1.1.3) são projetadas segundo os critérios descritos em [5]. Por exemplo, a projeção da receita operacional é determinada pelo produto entre o mercado previsto e a tarifa média do reposicionamento tarifário, enquanto os custos operacionais (custos de operação, manutenção, administração e gestão comercial) são projetados com base nos custos da empresa de referência referenciados à data do reposicionamento tarifário.

¹Calculado conforme os procedimentos descritos na Nota Técnica No 122/2005/SRE/ANEEL.

A empresa de referência é uma distribuidora virtual onde se simula a prestação do serviço de distribuição de energia elétrica nas mesmas condições que opera uma distribuidora real. Para um dado nível de qualidade exigido pelo regulador, a empresa de referência é orientada para a utilização eficiente dos recursos, estabelecendo assim um padrão de desempenho (*benchmark*) a ser perseguido pela empresa real. Esta estratégia segue as linhas gerais da regulação por comparação (*yardstick competition*), pois emula a competição entre uma distribuidora real e a respectiva empresa de referência.

A metodologia da empresa de referência, adotada no primeiro ciclo de revisões tarifárias [7], segue uma abordagem tipo *bottom-up* que se inicia com a identificação de todos os processos inerentes à atividade de distribuição de energia elétrica nas áreas comercial e técnica, passa pelo estabelecimento da eficiência dos custos de cada um dos processos identificados e se encerra com uma estimativa da eficiência dos custos operacionais totais. Portanto, trata-se de uma metodologia complexa e que abre a possibilidade de envolver o regulador em uma espécie de micro-gestão da empresa, justamente o contrário do que recomenda a boa prática da ação reguladora.

1.2

Revisão bibliográfica

Conforme pôde ser observado, a atual metodologia da empresa de referência é um tanto quanto complexa. Assim, para tornar mais objetivo o processo de estabelecimento da eficiência dos custos operacionais, a ANEEL não descarta a possibilidade de utilizar técnicas de *benchmarking*, entre as quais, fronteiras de eficiência, conforme o modelo atual adotado na quantificação da eficiência dos custos operacionais das empresas de transmissão [8]. A fronteira é o lugar geométrico dos melhores desempenhos e a simples comparação desta com as concessionárias, permite quantificar quanto cada uma deve melhorar sua posição em relação às demais distribuidoras.

Jasmab & Pollit [51] reportam que as principais metodologias para determinação destes índices de eficiência são: (i) Análise Envoltória de Dados (*Data Envelopment Analysis* - DEA), (ii) Análise por Modelos de Fronteira

Estocástica (*Stochastic Frontier Analysis* - SFA), (iii) Mínimos Quadrados Ordinários Corrigidos (*Corrected Ordinary Least Square* - COLS).

Vale frisar que as abordagens DEA [23] e econométrica (SFA e COLS) [58] partem de pressupostos diferentes e apresentam vantagens e desvantagens de acordo com a aplicação em que são empregados, portanto, não há um método superior, a priori. Resultados obtidos por distintos modelos DEA na avaliação da eficiência das distribuidoras do SEB foram apresentados por Resende [75], Vidal & Távora Júnior [95], Sollero & Lins [81] e Pessanha *et al.* [72]. Índices de eficiência obtidos por modelos SFA são apresentados por Zanini [97] e Arcoverde *et al.* [9]. Em Souza *et al.* [83], foi proposta uma comparação dos modelos DEA e SFA para o cálculo da eficiência dos custos operacionais das distribuidoras. Ressalta-se que são poucos os estudos semelhantes utilizando modelos SFA com enfoque Bayesiano [62]. Como exemplo, Kleit & Dek [53], por meio de uma abordagem Bayesiana, estimam uma fronteira de custo para unidades geradoras a gás natural. A partir de uma amostra de 78 usinas, os autores concluem que em média os custos poderiam ser reduzidos em torno de 13% eliminando-se a ineficiência. Zhao [98] apresenta três modelos de fronteira estocástica sob a ótica Bayesiana para avaliar a eficiência de unidades geradoras térmicas a carvão nos Estados Unidos. O primeiro modelo apresentado por Zhao [98] avalia os custos da geração de eletricidade de térmicas a carvão. Já o segundo modelo examina os custos da geração de eletricidade de térmicas a gás e, a seguir, compara estes com aqueles obtidos para o primeiro modelo. Por fim, por meio de um terceiro modelo, desta vez para os dados em painel, Zhao [98] mostra que nos Estados Unidos as unidades que operam em mercados desregulados são mais eficientes que as que operam no mercado regulado.

1.3

Objetivos

Nos tópicos anteriores procurou-se ressaltar a importância de elaborar um estudo voltado para a questão da revisão tarifária das empresas de distribuição de energia elétrica no Brasil, basicamente no que se refere ao cálculo da eficiência dos custos operacionais.

Como se sabe, o processo de privatização dessas empresas envolveu um significativo montante de recursos e, certamente, nenhuma das partes (estado e empresas privadas) quer que haja prejuízo. Entre outros fatores, um que merece ser sublinhado refere-se propriamente à definição e execução da técnica de revisão tarifária que, sendo bem elaborada, pode possibilitar a continuidade dos investimentos no setor. Afinal, é fundamental entender as possíveis conseqüências que esta aplicação pode ter sobre o desempenho das empresas, bem como o bem-estar dos consumidores.

Como o tema é bastante amplo e moderno, é plenamente justificável fornecer uma metodologia com fins científicos que contemple diversas áreas do conhecimento conseguindo, assim, reduzir possíveis incertezas relacionadas ao procedimento aplicado pelo regulador.

Em concordância com o que foi exposto, esta tese elucidada, discute e edita os principais resultados obtidos através das técnicas abaixo discriminadas:

1. A técnica DEA:

Técnica considerada como não-paramétrica por envolver o uso de programação linear para a construção de superfícies que permitam a estimação das eficiências de interesse. A principal vantagem é o fato de não ser necessária a imposição de uma forma funcional para a eficiência; porém, a fronteira estimada será muito sensível a *outliers* e pontos discrepantes. Qualquer desvio com relação à fronteira é considerado como medida de ineficiência.

2. A técnica SFA com enfoque Clássico:

Esta ferramenta, com característica econométrica, captura *outliers* e os desvios com relação à fronteira são devido a fatores não gerenciáveis. Dessa forma, é possível eliminar o efeito desses fatores sobre a medida de eficiência. Analisando por um outro prisma, esta técnica é muito restritiva no sentido de que impõe uma forma funcional para a tecnologia.

3. A técnica SFA com enfoque Bayesiano:

Esta metodologia permite realizar inferência a partir de dados utilizando modelos probabilísticos, tanto para as quantidades observadas quanto para as não observadas. Uma outra particularidade desta metodologia é possibilitar ao especialista a inclusão de seus conhecimentos prévios no modelo estudado. Por estas razões, os modelos Bayesianos são considerados mais flexíveis e, desta forma, na maioria das vezes não são analiticamente tratáveis. Para contornar esse problema, faz-se necessário a utilização de métodos de simulação. Os mais utilizados são os métodos de Monte Carlo baseados em cadeias de Markov (MCMC).

É importante assinalar que devido à heterogeneidade das concessionárias, é interessante agrupá-las em *clusters*. Afinal, é plausível conjecturar que, procedendo desta forma, cada empresa poderá ser melhor avaliada. Para tal, inicialmente as empresas foram agrupadas através de uma Rede Neural de Kohonen.

1.4

Estrutura da tese

O índice analítico reflete a dimensão desta obra que apresenta, de forma unificada e inédita, os principais métodos relacionados com o estudo de regulação tarifária. Dentre os quais, pode-se citar: Redes Neurais Auto-Organizáveis, Análise Envoltória de Dados e Fronteiras Estocásticas.

Esta tese é composta de dez capítulos e pode ser dividida em três partes. A primeira parte (Capítulos 1, 2, 3, 4 e 5) é devotada à exposição do problema e à abordagem dos aspectos teóricos básicos. Já a segunda parte (Capítulos 6, 7, 8 e 9) é constituída pela apresentação e interpretação dos resultados obtidos. A terceira e última parte (Capítulo 10) trata das conclusões e questões para estudo futuro.

O Capítulo 1 discute a teoria de regulação tarifária. Esta discussão tem como finalidade entender melhor a metodologia atualmente aplicada pelo regulador e, conseqüentemente, os objetivos desta pesquisa.

O Capítulo 2 explica as Redes Neurais Auto-organizáveis (Rede Neural de Kohonen -RNN). Aqui, um exemplo do funcionamento da rede é apresentado de forma clara e objetiva.

O Capítulo 3 aborda a Análise Envoltória de Dados (*Data Envelopment Analysis* – DEA). Conjectura-se necessária esta apresentação, pois na área de Pesquisa Operacional relacionado a métodos quantitativos, DEA é uma teoria relativamente recente e que tem recebido uma atenção muito especial por parte de muitos pesquisadores.

O Capítulo 4 compreende os principais ingredientes usados em estatística Bayesiana. Os conceitos chave de distribuição a priori, verossimilhança e distribuição a posteriori são explicados. Os principais métodos, computacionalmente intensivo, para inferência são abordados. Vale frisar que este capítulo também apresenta uma série de exemplos ilustrativos.

O Capítulo 5 discorre sobre as Fronteiras Estocásticas (*Stochastic Frontier Analysis* - SFA). Inicialmente, esta metodologia é apresentada á luz da escola freqüencista, seguida pela abordagem Bayesiana. É digno de registro que a fronteira estocástica sob o paradigma Bayesiano tem muitas vantagens; porém ainda não é muito explorada e, sobretudo, ela requer alguma maturidade matemática.

O Capítulo 6 é dedicado à apresentação dos resultados encontrados pela Rede Neural de Kohonen.

O escopo do Capítulo 7 trata dos resultados encontrados por três modelos DEA.

O Capítulo 8 publica os resultados auferidos por diversos modelos de fronteira estocástica. É importante assinalar que são dezoito modelos, dentre os quais dez são de natureza Bayesiana.

O Capítulo 9 traz a análise multivariada de dados via matriz de correlação de Pearson e Spearman e Análise Fatorial. Vale informar que este Capítulo não se propõe a ser exaustivo, as duas abordagens foram selecionadas de acordo com os conhecimentos e interesses do autor.

O Capítulo 10 tem o propósito de suscitar os resultados mais relevantes e sumarizar os principais pontos deste trabalho. Também são registradas recomendações de pesquisas futuras.

Este trabalho é finalizado com o registro das referências bibliográficas, Anexo 1 (dados utilizados) e Anexo 2 (*script* do programa de um modelo Bayesiano).

2

Redes Neurais Auto-Organizáveis

2.1

Introdução

Problemas de *clustering* estão presentes nos mais variados contextos, como por exemplo: classificação de padrões, mineração de dados e recuperação de informações de documentos (por exemplo pode-se citar busca na *internet*). Nesse contexto, *clustering* é considerado como uma das principais etapas da análise de dados, denominada análise de *cluster*, e consiste em classificar os dados de forma não supervisionada. A análise de *cluster* envolve, portanto, a organização de um conjunto de padrões (usualmente representados na forma de vetores de atributos ou pontos em um espaço multidimensional – espaço de atributos) em *clusters* (agrupamentos), de acordo com alguma medida de similaridade. Intuitivamente, padrões pertencentes a um dado *cluster* devem ser mais similares entre si do que em relação a padrões pertencentes a outros *clusters*. Há diversas técnicas utilizadas para a construção destes conglomerados (*clusters*) as quais são classificadas como hierárquicas e não-hierárquicas. De acordo com Mingoti [66], as técnicas hierárquicas, na maioria das vezes, são utilizadas em análise exploratória de dados com o intuito de identificar possíveis agrupamentos e o valor provável do número de grupos g . Com relação às técnicas não hierárquicas, é necessário que o valor do número de grupos já esteja pré-especificado pelo especialista. Conceitos teóricos, algoritmos e exemplos dessas metodologias podem ser encontrados em [26, 49, 60, 66, 67].

2.2

Componentes de uma tarefa de *clustering*

Segundo Jain et al [50], atividades de *clustering* normalmente envolvem os seguintes passos:

1. Representação dos padrões (podendo incluir extração ou seleção de características);
2. Definição de uma medida de similaridade apropriada ao domínio da aplicação;
3. *Clustering* ou Agrupamento;
4. Apresentação dos resultados.

A Figura 2.1 mostra a seqüência típica dos três (3) primeiros passos registrados acima.

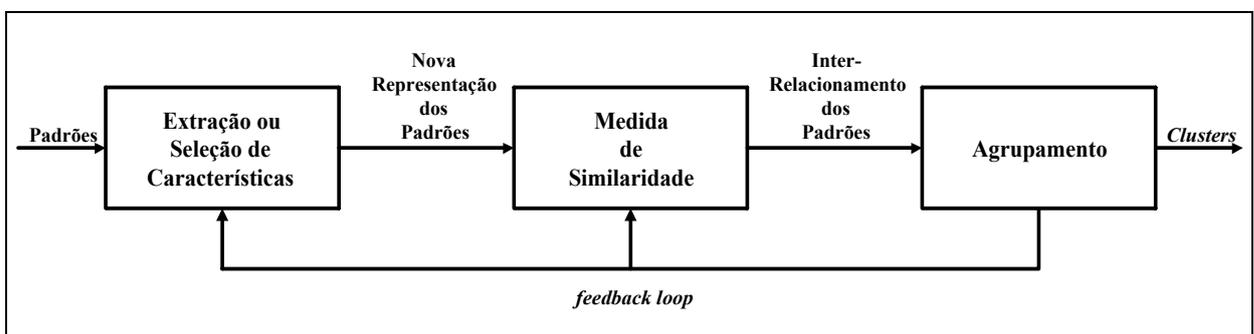


Figura 2.1: Etapas de um processo de *clustering*.

Em geral, pode-se dizer que:

- Representação dos padrões: envolve definição do número, tipo e modo de apresentação dos atributos que descrevem cada padrão;
- Seleção de características: processo de identificação do subconjunto mais efetivo dos atributos disponíveis para descrever cada padrão;
- Extração de características: uso de uma ou mais transformações junto aos atributos de entrada de modo a salientar uma ou mais característica dentre aquelas que estão presentes nos dados;
- Medida de Similaridade: fornecida por uma função de distância definida entre pares de dados ou padrões. É possível incluir na medida de distância, aspectos conceituais (qualitativos) ou então numéricos (quantitativos);
- Agrupamento: os grupos podem ser definidos como conjuntos *crisp* (um padrão pertence ou não-pertence a um determinado grupo) ou *fuzzy* (um padrão pode apresentar graus de pertinência aos grupos).

O processo de agrupamento pode ser hierárquico, com um processo recursivo de junções ou separações de grupos, ou não-hierárquico, com o emprego direto de técnicas de discriminação de *clusters*;

- A realimentação (*feedback loop*) do resultado do processo de *clustering* pode levar, se necessário, à redefinição dos módulos de “extração ou seleção de características” e de medida de similaridade.

2.3

Aprendizado por competição

As pesquisas na área de Redes Neurais Artificiais (RNA) têm passado por uma rápida expansão e desfrutado de grande popularidade tanto na comunidade de pesquisa industrial como acadêmica. Este sucesso é dedicado ao fato das RNA serem uma técnica não-linear bastante sofisticada capaz de modelar funções extremamente complexas. Com isso, resultados bastante expressivos estão sendo obtidos em uma gama muito grande de áreas, como exemplo, finanças, medicina, engenharia, geologia e física. De fato, qualquer área onde há problemas de predição, classificação e controle, RNA são apresentadas como uma ferramenta muito interessante [82].

É importante ressaltar que RNA são estruturas maciçamente paralelas inspiradas no conhecimento sobre os sistemas nervosos biológicos. Vale lembrar que alguns modelos de RNA são tão grosseiros, quando comparados a sistemas nervosos biológicos, que qualquer comparação a estes se torna improcedente [67].

De acordo com Haykin [49], as RNA são capazes de aprender a partir de seu ambiente e de melhorar o seu desempenho através da aprendizagem. A melhoria do desempenho ocorre com o tempo de acordo com alguma medida pré-estabelecida. Uma Rede Neural Artificial aprende acerca do seu ambiente através de um processo iterativo de ajustes aplicados a seus pesos sinápticos e níveis de bias. Idealmente, a Rede Neural Artificial se torna mais instruída sobre o seu ambiente após cada iteração do processo de aprendizagem.

No contexto de RNA, aprendizagem pode ser definida como:

“Aprendizagem é um processo pelo qual os parâmetros livres de uma Rede Neural Artificial são adaptados através de um processo de estimulação pelo

ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre.”

Um conjunto pré-estabelecido de regras bem-definidas para a solução de um problema de aprendizagem é denominado um algoritmo de aprendizagem.

Nascimento Jr. & Yoneyama [67] mencionam que os diversos métodos de aprendizado podem ser classificados de acordo com a participação do supervisor.

Este trabalho utiliza um tipo de aprendizagem cujo grau de supervisão é muito pequeno sendo, portanto, denominado de supervisão muito fraca. Neste caso, o algoritmo de treinamento da Rede Neural Artificial tenta descobrir categorias entre os dados de entrada e o supervisor participa apenas fornecendo os rótulos para estes agrupamentos. Na literatura, esse tipo de aprendizado é denominado aprendizado não supervisionado ou auto-organizado, um nome que, estritamente falando, não estaria totalmente apropriado, pois ainda ocorre uma pequena (mas efetiva) participação do supervisor no processo de aprendizagem [67]. Para a realização desta aprendizagem, é utilizado a regra de aprendizado competitivo.

Nesse processo de aprendizagem competitivo, como o nome indica, as unidades de saída de um mapeamento competem entre si para se tornar ativos (disparar). É esta característica que torna a aprendizagem competitiva muito adequada para descobrir características estatisticamente similares e, assim sendo, ter a capacidade de fazer classificações em um conjunto de padrões de entrada.

Há na literatura diversas medidas para determinar a similaridade entre entradas. Uma medida de similaridade usada freqüentemente é baseada no conceito de distância euclidiana.

2.4

Rede Neural de Kohonen (RNK)

Rede Neural de Kohonen utiliza aprendizado com supervisão muito fraca. Essa particularidade faz com que sua utilização seja bastante diferente em relação aos outros tipos de RNA em virtude de todos os outros tipos serem projetadas para executar tarefas mediante um processo de aprendizado supervisionado.

Esta Rede Neural é formada por grades de unidades uni ou bidimensionais (maiores dimensões são possíveis, mas pouco comuns), que modificam seus pesos

sinápticos em um processo de aprendizagem competitivo, formando sobre a grade de saída um sistema de coordenadas significativas para diferentes características de entrada (mapa topográfico dos padrões de entrada). A Figura 2.2 ilustra o caso de uma grade bidimensional, que é mais usual. Se duas unidades estão topologicamente próximas nesse espaço, elas são ditas unidades vizinhas [67]:

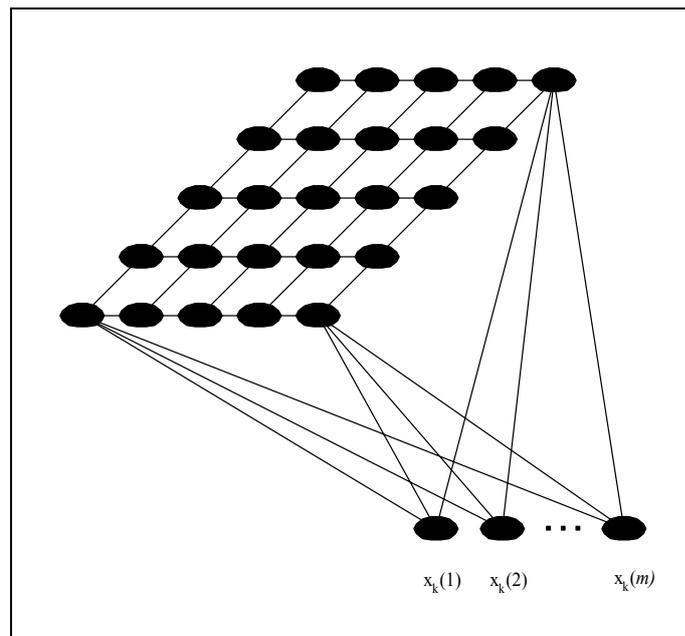


Figura 2.2: Rede de Kohonen com as unidades de saída organizadas como grade bi-dimensional.

O objetivo da Rede Neural de Kohonen é obter um mapeamento entrada-saída, onde a topologia do espaço de entrada seja preservada no espaço de saída. Assim, as entradas que são próximas no espaço de entrada serão mapeadas para saídas próximas no espaço de saída, ou seja, ativarão unidades de saída que estão fisicamente próximas.

Uma consequência interessante da Rede de Kohonen é a alocação das unidades de saída de acordo com a distribuição de probabilidade das entradas no espaço de entrada. Por exemplo, se não for usado durante o treinamento da Rede Neural nenhum ponto de uma determinada região do espaço de entrada, nenhuma unidade será modificada de forma a aumentar a similaridade de seu vetor de pesos com as entradas desta região. Analogamente, mais unidades de saída serão atraídas para as regiões do espaço de entrada com mais pontos de treinamento.

2.5

Exemplo ilustrativo da Rede Neural de Kohonen

Este t3pico tem como objetivo elucidar a implementa33o de uma Rede Neural de Kohonen conforme registrado em Souza [82]. Para tal, seja uma rede com 9 unidades (neur3nios), 20 padr3es (vetores de treinamento). 3 importante assinalar que cada padr3o possui 3 entradas (atributos).

A Tabela 2.1 fornece o centro de cada *cluster* com seus respectivos pontos de treinamento:

Tabela 2.1 - Centro dos *clusters* com seus respectivos pontos de treinamento.

Centro dos <i>Clusters</i>	Pontos de Treinamento
$C_1 = [0 \ 0 \ 0]^T;$	$x_1 = [0,1033 \ -0,3161 \ -1,1968]^T;$ $x_2 = [-0,5450 \ -0,4317 \ 0,0667]^T;$ $x_3 = [0,2247 \ 0,5830 \ -1,1134]^T;$ $x_4 = [0,8058 \ -0,2275 \ -0,2512]^T;$ $x_5 = [1,6383 \ -0,8553 \ 0,1821]^T;$
$C_2 = [0 \ 10 \ 0]^T;$	$x_6 = [-0,0510 \ 8,2008 \ 0,5728]^T;$ $x_7 = [-0,8182 \ 11,0073 \ 0,9793]^T;$ $x_8 = [-0,5798 \ 10,1410 \ 1,8104]^T;$ $x_9 = [-1,6670 \ 9,0699 \ 0,4962]^T;$ $x_{10} = [0,1776 \ 10,9324 \ -0,5245]^T;$
$C_3 = [10 \ 0 \ 0]^T;$	$x_{11} = [10,3432 \ -0,6592 \ 1,1919]^T;$ $x_{12} = [9,1674 \ -1,4384 \ 1,0644]^T;$ $x_{13} = [9,6810 \ -0,8515 \ 0,2207]^T;$ $x_{14} = [9,1564 \ -1,7950 \ 0,7760]^T;$ $x_{15} = [9,5935 \ 0,2390 \ 1,2727]^T;$
$C_4 = [10 \ 10 \ 10]^T;$	$x_{16} = [8,9785 \ 9,2988 \ 8,8607]^T;$ $x_{17} = [10,8186 \ 8,8467 \ 9,8218]^T;$ $x_{18} = [10,7659 \ 9,9824 \ 10,3016]^T;$ $x_{19} = [8,7124 \ 9,9586 \ 10,3646]^T;$ $x_{20} = [9,0056 \ 8,6034 \ 9,3408]^T;$

Continuando, tem-se a disposição das unidades em uma grade bi-dimensional:

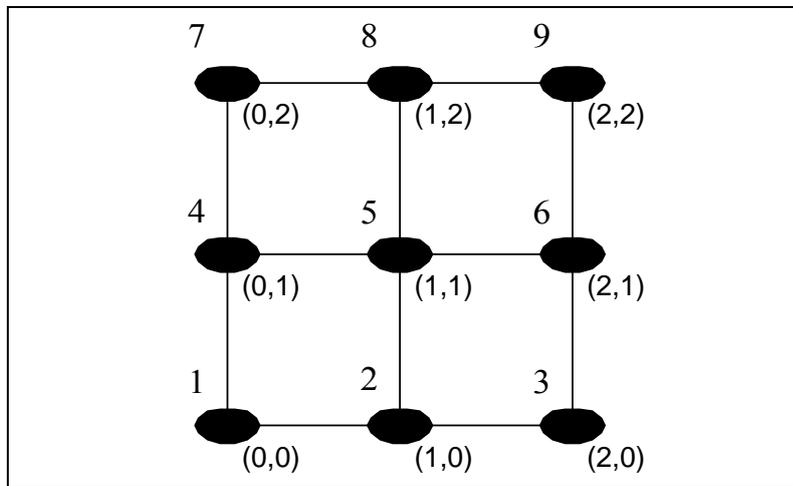


Figura 2.3: Disposição bi-dimensional das unidades.

Antes de mostrar a disposição da grade das unidades no espaço das entradas, é necessário atribuir, aleatoriamente, valores iniciais para os pesos da rede. Dessa forma, tem-se

$$\begin{aligned} \mathbf{w}_1^0 &= [1 \ 1 \ 3]^T, & \mathbf{w}_2^0 &= [2 \ 3 \ 1]^T, & \mathbf{w}_3^0 &= [3 \ 0 \ 5]^T, & \mathbf{w}_4^0 &= [3 \ 2 \ 2]^T, \\ \mathbf{w}_5^0 &= [1 \ 4 \ 3]^T, & \mathbf{w}_6^0 &= [5 \ 1 \ 1]^T, & \mathbf{w}_7^0 &= [5 \ 5 \ 5]^T, & \mathbf{w}_8^0 &= [3 \ 1 \ 0]^T, \\ \mathbf{w}_9^0 &= [4 \ 1 \ 3]^T \end{aligned}$$

A Figura 2.4 ilustra este processo.

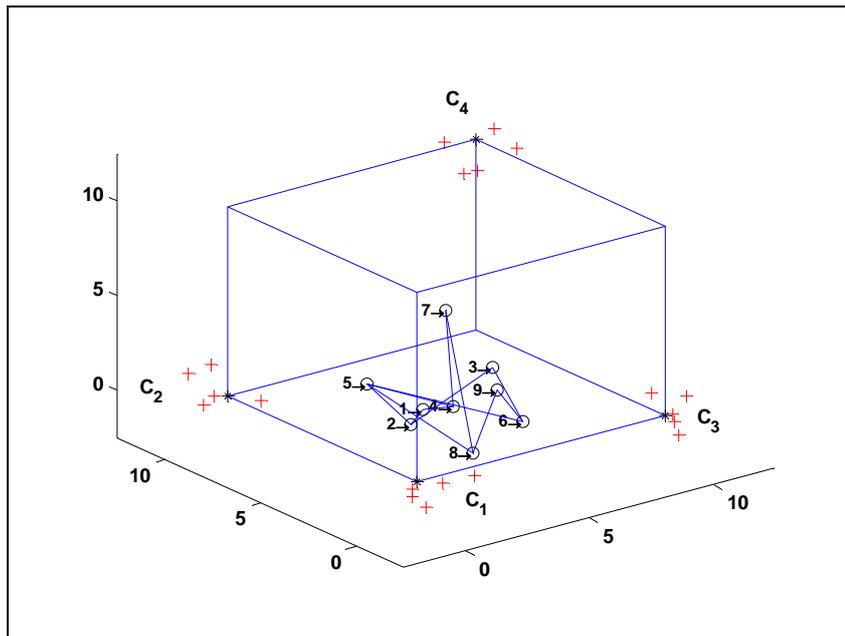


Figura 2.4: Disposição das unidades no espaço das entradas.

Neste ponto, é importante descrever como é feito o algoritmo de treinamento:

1. Inicie os pesos da rede \mathbf{W}_i^0 , ($i = 1, \dots, 9$), Figura 2.4;
 Defina o raio inicial da vizinhança, $R_{v_0} = 3\sqrt{2}$;
 Defina o número de iterações, $Niter = 50$;
 Defina a taxa de aprendizado inicial $\eta_0 = 0,5$;
 Defina o η mínimo durante o aprendizado, $\eta_{min} = 0,05$;
 Defina $\eta = \eta_0$;
2. *Loop* em k (número de iterações) de 1 a 50;
3. Reordene, de forma aleatória, os 20 vetores de treinamento para esta iteração;
4. *Loop* em j (número do vetor de treinamento na seqüência aleatória) de 1 a 20;
5. Calcule a distância euclidiana entre \mathbf{x}_j^k e todos os pesos da rede:

$$d(\mathbf{x}_j^k, \mathbf{W}_i^k) = \|\mathbf{x}_j^k - \mathbf{W}_i^k\| \text{ para } j \text{ fixo e } i = 1, \dots, 9;$$

6. Determine o índice ν da unidade vencedora, a unidade com a maior similaridade em relação ao vetor \mathbf{x}_j^k , ou seja:

$$\mathbf{W}_\nu^k = \min_i \left\{ \left\| \mathbf{x}_j^k - \mathbf{W}_i^k \right\| \right\}, i = 1, \dots, 9;$$

7. Atualize os pesos da unidade vencedora \mathbf{W}_ν^k e da sua vizinhança, usando o critério da distância, ou seja:

loop em i ($i = 1, \dots, 9$);

$$\text{calcule } d_{i\nu} = \left\| \mathbf{p}_i - \mathbf{p}_\nu \right\|$$

$$\text{se } d_{i\nu} < R_\nu \text{ então faça: } \mathbf{W}_i^{k+1} = \mathbf{W}_i^k + \eta (\mathbf{x}_j^k - \mathbf{W}_i^k)$$

fim do *loop* em i ;

8. Fim do *loop* em j . Retorne ao passo 4 até que todos os 20 pontos de treinamento sejam apresentados 1 única vez em sequência aleatória;
9. Reduza o tamanho da vizinhança, $R_\nu = R_{\nu 0} (Niter - k + 1) / Niter$;
10. Reduza a constante de aprendizado $\eta, \eta = \eta_{\min} + \eta_0 \cdot (Niter - k + 1) / Niter$;
11. Fim do *loop* em k (*loop* de iterações)

No passo 7, \mathbf{p}_i e \mathbf{p}_ν indicam, respectivamente, os vetores com a localização das unidades i e ν no arranjo geométrico das unidades de saída. Como exemplo, sejam: $\mathbf{p}_1 = [0 \ 0]^T$, $\mathbf{p}_2 = [1 \ 0]^T$, $\mathbf{p}_5 = [1 \ 1]^T$, $\mathbf{p}_9 = [2 \ 2]^T$; assumindo a grade com distância horizontal e vertical unitária.

A identificação dos *clusters* ocorre no espaço das entradas. Em contrapartida, todo o processo referente ao critério de vizinhança como seus respectivos ajustes, se dão no espaço topológico das unidades verificado na Figura 2.3.

Após implementar o algoritmo em Matlab (www.mathworks.com), tem-se o seguinte resultado:

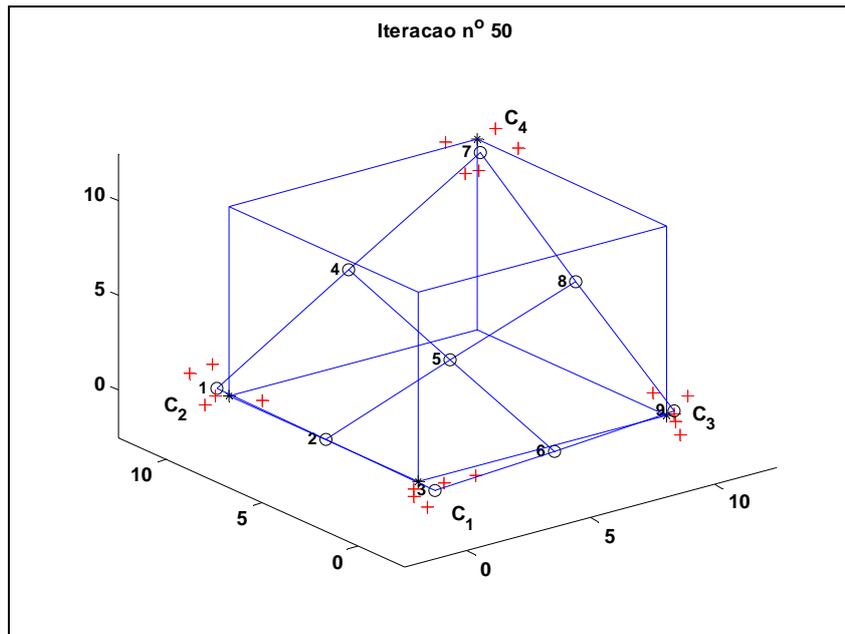


Figura 2.5: Resultado do processo de identificação dos *clusters*.

Através da Figura 2.5, pode-se perceber claramente que a rede identificou os 4 *clusters* registrados na Tabela 2.1:

- Unidade 1 \leftrightarrow *cluster* 2;
- Unidade 3 \leftrightarrow *cluster* 1;
- Unidade 7 \leftrightarrow *cluster* 4;
- Unidade 9 \leftrightarrow *cluster* 3.

2.6

Rede Neural de Kohonen via simulação de Monte Carlo (MCRNK)

É pertinente enfatizar que os pesos sinápticos, da Rede Neural de Kohonen, muito influenciam nos resultados dos *clusters* encontrados. Afinal, como eles (pesos sinápticos) são inicializados com valores aleatórios diferentes, a cada simulação podem-se encontrar *clusters* bem distintos. Esta conjectura foi comprovada por Souza [82] que propôs fazer uma simulação de Monte Carlo utilizando-se dos resultados obtidos a partir da Rede Neural de Kohonen. É importante salientar que, procedendo desta forma, resultados estatisticamente

mais significativos são obtidos. Por esta razão, esta tese faz uso desta técnica a qual foi denominada de Rede Neural de Kohonen via simulação de Monte Carlo (MCRNK).

3

Análise Envoltória de Dados

3.1

Introdução

O processo de avaliar a eficiência relativa de diversas organizações denominadas DMU's (do inglês *Decision Making Units*) foi originalmente proposto por Farrell [31]. Charnes et al. [19] implementaram e popularizaram esta técnica denominando-a *Data Envelopment Analysis* (DEA) que é uma generalização do modelo proposto por Farrell baseado em programação fracionária [68].

DEA foi utilizada pela primeira vez para avaliação das escolas públicas norte-americanas e hoje é largamente aplicada tanto em problemas diversos de cunho empresarial quanto em áreas que até então apresentavam algumas dificuldades para serem avaliadas. Como exemplo pode-se citar saúde (hospitais, doutores), educação (escolas, universidades), bancos, fábricas, eficiência energética, *fast food*, restaurantes, lojas de varejo, etc. A justificativa para o crescente uso da metodologia DEA é o conjunto de possibilidades de análise dos dados que a ferramenta disponibiliza. Charnes et al [18] discorrem algumas razões que motivaram o intenso uso da aplicação. Como exemplo, tem-se:

- poucas técnicas disponíveis para avaliar a eficiência das organizações;
- confiança por parte dos pesquisadores nos resultados obtidos de sua utilização;
- possibilidade de usar simultaneamente múltiplos *outputs* e múltiplos *inputs* mesmo sendo eles dimensionados em unidades diferentes;
- capacidade de ajustar para variáveis exógenas e incorporar variáveis dicotômicas (*dummy*);
- facilidade de utilização, a existência de *softwares* que reduzem o esforço computacional;

- habilidade em identificar possíveis DMU's ineficientes em virtude de excessos de *inputs* como carência de *outputs* para cada entidade;
- capacidade de identificar *benchmarks*, etc.

De acordo com Sueyoshi [85], esta técnica bastante robusta pode ser considerada não-paramétrica pois permite eliminar a necessidade de especificar explicitamente uma relação funcional básica entre *inputs* e *outputs*, a qual é esperada para descrever analiticamente em estudos paramétricos convencionais.

Neste contexto, DEA representa um princípio alternativo de extrair informações sobre um conjunto de observações tal como na Figura 3.1:

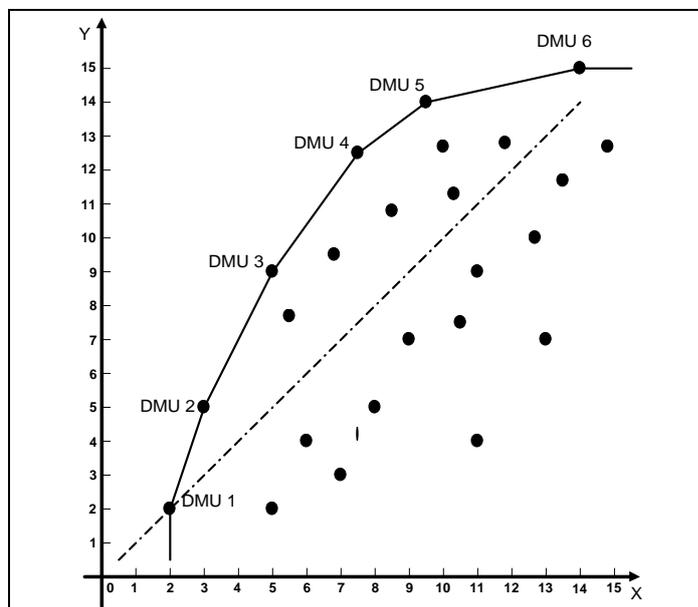


Figura 3.1: Comparação de DEA e regressão [18].

Na Figura 3.1, acima, tem-se a ilustração de uma situação simples em que há somente um *input* e um *output*. Pode-se observar que a linha pontilhada corresponde à linha de regressão passando através da origem e calibrada pelo método dos mínimos-quadrados ordinários. Já a linha superior, dita de fronteira de produção ou função de produção, determina o lugar geométrico das DMU's ditas eficientes (*best-practice*).

Ao contrário das técnicas paramétricas, cujo objetivo é otimizar um simples plano de regressão através dos dados, DEA otimiza sobre cada observação com o propósito de calcular uma fronteira de produção determinada

pelas DMU's eficientes. Ambas as técnicas, paramétrica e não-paramétrica, usam as informações contidas nos dados. Em uma análise paramétrica, a simples equação de regressão é aplicada em cada DMU. O procedimento DEA otimiza o desempenho medido de cada DMU em relação a todas as outras DMU's em um sistema de produção o qual transforma múltiplos insumos (*inputs*) em múltiplos produtos (*outputs*), através do uso de Programação Linear (PL) onde é resolvido um conjunto de Problemas de Programação Linear (PPL's) inter-relacionados, tantas quantas forem as DMU's, objetivando, assim, determinar a eficiência relativa de cada uma delas. Desta forma, a entrada ou retirada de uma ou mais unidades no conjunto de observações pode implicar em mudança no valor da eficiência relativa.

3.2

Conjunto de Possibilidades de Produção

Conforme foi mencionado, DEA é uma ferramenta utilizada para avaliar DMU's que realizam tarefas similares e para as quais dispomos de medidas de *inputs* (insumos) e *outputs* (produtos).

Um Conjunto de Possibilidades de Produção, \mathbf{T} , é um conjunto de combinações viáveis dos vetores de *inputs* \mathbf{x}_j e de *outputs* \mathbf{y}_j .

Considere, então, que cada uma das n DMU's tem transformado seus m *inputs* em s *outputs*.

Matematicamente, tem-se:

- $\mathbf{x}_j = [x_{j1} \ \dots \ x_{jm}] \in \mathbf{R}_+^m$ - conjunto de *inputs* observados consumidos pela j -ésima DMU, $j = 1, \dots, n$;
- $\mathbf{y}_j = [y_{j1} \ \dots \ y_{js}] \in \mathbf{R}_+^s$ - conjunto de *outputs* observados produzidos pela j -ésima DMU, $j = 1, \dots, n$.

Sejam também:

- $\mathbf{T} = \{(x, y) \in \mathbf{R}_+^{m+s} \mid x \text{ pode produzir } y\}$, Conjunto de Possibilidades de Produção;

- $P(x) = \{y | (x, y) \in T\}$ é o conjunto de todos *outputs* observados y que podem ser produzidos usando o *input* x ;
- $L(y) = \{x | (x, y) \in T\}$ é o conjunto de todos *inputs* observados x que podem produzir um dado *output* y .

Em muitas aplicações, T é desconhecido e necessita ser estimado. Com este propósito, a técnica DEA estima T a partir dos dados observados.

Neste ponto, é importante listar algumas propriedades para este conjunto T . Também é assumido que todo $x', x'' \in \mathbf{R}_+^m$ e $y', y'' \in \mathbf{R}_+^s$.

- **P(1) – Disponibilidade.**

$$x' \in L(y') \text{ e } x'' \geq x' \Rightarrow x'' \in L(y'),$$

$$y' \in P(x') \text{ e } y'' \leq y' \Rightarrow y'' \in P(x').$$

- **P(2) – Raio Ilimitado.**

$$(x', y') \in T, \text{ então } (kx', ky') \in T \quad \forall k > 0.$$

- **P(3) – Convexidade.**

Se $(x_j, y_j) \in T, j = 1, \dots, n$; e $\lambda \geq 0$ é um vetor coluna de dimensão

$(n \times 1)$ tal que $\sum \lambda = 1$ então $(X^T \lambda, Y^T \lambda) \in T$. É importante registrar que X^T é uma matriz de *inputs* de dimensão $(m \times n)$, Y^T é uma matriz de *outputs* de dimensão $(s \times n)$ e \underline{e} é um vetor linha de “1” com dimensão $(1 \times n)$.

- **P(4) – Extrapolação Mínima.**

T é o conjunto interseção de todos os \hat{T} estimados que satisfazem as propriedades **P(1)**, **P(2)** e **P(3)** sujeitos à condição que $(x_j, y_j) \in \hat{T}$.

Isto é, T é o menor conjunto consistente com os dados observados em **P(1)**, **P(2)** e **P(3)**. Forni [33] explica as considerações matemáticas que provam

que \mathbf{T} satisfaz as quatro propriedades acima. Assim sendo, este conjunto pode ser definido como:

$$\mathbf{T} = \left\{ (x, y) \mid x \geq \mathbf{X}^T \lambda, y \leq \mathbf{Y}^T \lambda, \lambda \geq 0 \right\} \quad (3.2.1)$$

Muitas vezes, o termo raio ilimitado é referenciado como retorno constante de escala. Basicamente, o modelo desenvolvido por Charnes et al [19], denominado CCR, procura estimar este conjunto e a partir dele calcular os índices de eficiência relativa das DMU's.

Se a propriedade **P(2)** é desconsiderada, o **Conjunto de Possibilidades de Produção**, \mathbf{T} , passa a ser definido como:

$$\mathbf{T} = \left\{ (x, y) \mid x \geq \mathbf{X}^T \lambda, y \leq \mathbf{Y}^T \lambda, e \lambda = 1, \lambda \geq 0 \right\} \quad (3.2.2)$$

Este conjunto está relacionado com o modelo proposto por Banker et al [11]. Mais adiante, estas definições serão abordadas novamente com suas respectivas utilidades.

3.3

Conceito de eficiência

Coelli et al [21] *apud* Farrell [31] o qual propôs duas medidas de eficiência: Eficiência Técnica, a qual reflete a habilidade de uma empresa obter o máximo de *output* a partir de um dado *mix* de *inputs*, e Eficiência Alocativa, a qual reflete a habilidade de uma firma usar seus *inputs* em proporções ótimas, dado seus respectivos preços e a produção tecnológica. Estas duas medidas são então combinadas para fornecer a medida de Eficiência Econômica Total.

Estes conceitos podem ser compreendidos, de uma maneira mais clara, através de uma análise gráfica. Para tal, seja a Figura 3.2 que considera um processo de produção em que um único *output* é produzido por dois *inputs*:

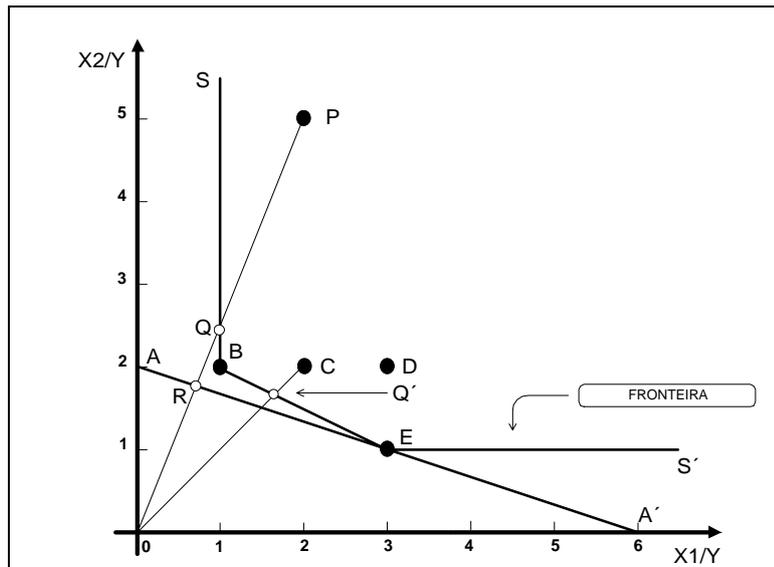


Figura 3.2: Eficiências Técnica, Alocativa e Econômica.

A curva representada por SS' que conecta B e E, representa uma fronteira de produção DEA a qual nos permite calcular a Eficiência Técnica de todas as DMU's envolvidas no processo. As empresas localizadas sobre a fronteira são identificadas empiricamente como DEA-Eficientes. Já a DMU C é rotulada como DEA-Ineficiente. Pois bem, o *score* de ineficiência da DMU C pode ser representado pela distância CQ' . É possível expressar Q' como sendo uma combinação convexa ² de B e E projetada sobre a fronteira de eficiência DEA. Averiguando fisicamente a distância CQ' , pode-se dizer que esta equivale à quantidade de *inputs* que pode-se reduzir proporcionalmente sem a redução de uma unidade de *output*.

Badin [10], generalizando o problema para o caso de múltiplos *inputs* e múltiplos *outputs*, menciona que Ineficiência Técnica pode ser definida como o fracasso de alcançar a fronteira de produção, isto é, o fracasso em alcançar o máximo possível de *outputs* dado um certo *mix* de *inputs*.

² Os modelos DEA assumem que o Conjunto de Possibilidades de Produção é um conjunto convexo e, de fato, se $\mathbf{x}_B = [x_{B1} \ x_{B2}]$, $\mathbf{y}_B = [y_{B1}]$ e $\mathbf{x}_E = [x_{E1} \ x_{E2}]$, $\mathbf{y}_E = [y_{E1}]$ pertencem a \mathbf{T} , então cada ponto sobre o segmento de reta conectando estas duas atividades também pertencem a \mathbf{T} .

Usualmente esta grandeza pode ser expressa de forma percentual. Neste caso tem-se que a Eficiência Técnica (**ET**) DEA da firma **P** é dada por:

$$ET = \frac{\overline{OQ}}{\overline{OP}}$$

(3.3.1)

Conhecendo-se a informação da razão dos preços dos *inputs* pelos *outputs*, pode-se determinar o segmento de reta AA' e dessa forma calcular a Eficiência Alocativa

(**EA**). Novamente, a Eficiência Alocativa de uma DMU operando no ponto **P** é:

$$EA = \frac{\overline{OR}}{\overline{OQ}} \quad (3.3.2)$$

Finalizando, tem-se que Eficiência Econômica (**EE**) de **P** é equivalente ao produto das duas eficiências anteriores, ou seja:

$$EE = \frac{\overline{OR}}{\overline{OP}} \quad (3.3.3)$$

Fundamentalmente, Eficiência Técnica está relacionada ao aspecto físico-operacional da produção, enquanto a Eficiência Econômica se preocupa com o aspecto monetário da produção.

Um comentário importante em relação à Eficiência Técnica baseia-se no fato de que ela é determinada ao longo de um segmento de reta que interliga a origem ao ponto de produção observado. Desta forma ela é considerada como uma medida de eficiência radial cuja vantagem é que as unidades são invariantes, não interferindo na medida de eficiência em si.

3.4

Benchmarks

Badin [10] cita que *Benchmarking* é um processo contínuo e sistemático de avaliação de empresas e serviços através da sua comparação com unidades consideradas eficientes, levando ao estabelecimento de ações gerenciais efetivas com o objetivo de aprimorar os resultados (redução dos custos, aumento de produção, etc). É um dos modernos instrumentos de gerência que possibilita a melhoria do desempenho técnico-econômico das empresas de forma comparativa.

Neste processo, parte-se do pressuposto de que as empresas que atuam num certo ramo representam atividades semelhantes, com padrões que podem ser aplicados a todas, como por exemplo, processamento de pedidos, processamento de dados ou lançamento de novos produtos. Algumas delas conseguem combinar melhor seus insumos (*inputs*), utilizando processos e técnicas de gestão mais aprimorados, e gerando produtos (*outputs*) com maior eficiência e eficácia.

Essas empresas servem, então, de referência para as demais, que devem rever seus processos de forma a caminhar no sentido de um melhor desempenho gerencial.

3.5

Retornos de escala

O conceito de retornos de escala é algo muito amplo que reflete o quanto um aumento proporcional em todos os *inputs* implica em um aumento no *output*. Baseado em Coelli et al [21], os conceitos em termos algébricos são apresentados. Considerando x_1 e x_2 duas variáveis de *input* observadas, tem-se:

Tabela 3.1 - Representação algébrica de retornos de escala.

Retornos de escala	Definição
Constante	$f(\alpha x_1, \alpha x_2) = \alpha f(x_1, x_2)$
Crescente	$f(\alpha x_1, \alpha x_2) > \alpha f(x_1, x_2)$
Decrescente	$f(\alpha x_1, \alpha x_2) < \alpha f(x_1, x_2)$

Nota: ($\alpha > 1$).

Há muitas razões pelas quais um setor de atividade econômica pode apresentar diferentes características de retornos de escala. O exemplo mais comum corresponde ao de uma pequena empresa que apresenta retornos de escala crescentes por ter contratado um funcionário adicional, especialista em uma determinada tarefa. Uma possível razão para retornos de escala decrescentes é o caso onde uma firma tornou-se tão grande, que a gerência não é capaz de exercer um total controle sobre todos os aspectos do processo de produção.

Análises empíricas de produção rotineiramente investigam retornos de escala por estimar a elasticidade total da produção (e). Frequentemente, são

encontradas estimativas da elasticidade parcial da produção, (E_i), que mede a mudança proporcional na variável *output*, resultante de um acréscimo proporcional no i -ésimo *input*, com todos os outros *inputs* mantidos constantes. Pode-se defini-la como:

$$E_i = \frac{\partial y}{\partial x_i} \cdot \frac{x_i}{y} \quad (3.5.1)$$

A elasticidade total da produção (e) mede a mudança proporcional na variável *output* resultante de um acréscimo proporcional unitário em todos os *inputs*, ou seja:

$$e = E_1 + E_2 + \dots + E_m \quad (3.5.2)$$

O valor de e está relacionado com retornos de escala da seguinte maneira:

Tabela 3.2 - Relação entre retornos de escala e elasticidade total da produção.

Retornos de escala	Elasticidade Total da Produção (e)
Constante	=1
Crescente	>1
Decrescente	<1

Vale sublinhar que estes conceitos são muito utilizados no desenvolvimento de modelos DEA, assunto do próximo tópico.

3.6

Modelos DEA

Observa-se que um grande desafio consiste em estimar, através dos dados disponíveis muitas vezes constituídos de múltiplos *inputs* e *outputs*, uma função de produção não paramétrica.

Essencialmente, dentre os vários modelos DEA, cada um procura estabelecer quais subconjuntos de n_0 DMU's determinará esta envoltória. Como será visto, a geometria desta envoltória está altamente relacionada à escolha do modelo DEA. Desta forma, é necessário que o especialista estude, de forma criteriosa, vários modelos DEA, antes de um veredicto final. Afinal, cada empresa

pode utilizar, nos seus diversos setores, diferentes versões de tecnologia da produção, a qual pode ser considerada ideal neste período do tempo e, mesmo assim, devido ao modelo DEA empregado, ser considerada tecnicamente ineficiente. É importante, então, que os modelos sejam desenvolvidos com o suporte DEA objetivando não somente capturar a relação entre as quantidades de *input* e *output*, mas também outros fatores externos que influenciam na análise de eficiência.

Os dois modelos considerados como os pilares de toda teoria são:

- **CCR** - modelo pioneiro desenvolvido por Charnes, Cooper e Rhodes [19];
- **BCC** – modelo sugerido por Banker, Charnes e Cooper [11], relacionado à ausência da propriedade **P(2)**, com o objetivo de sanar algumas carências do primeiro modelo.

Vale frisar que estes modelos fornecem ao gestor informações importantes sobre assuntos econômicos, gerenciando possibilitando, dessa forma, tomar medidas pertinentes em relação ao desempenho de sua empresa.

3.6.1

Modelo DEA CCR *input-oriented*

Seja o caso mais trivial onde cada unidade tem apenas um *input* e um *output*. A eficiência técnica de cada unidade, segundo o conceito clássico de engenharia, pode ser expresso como a razão do *output* pelo *input*:

$$ET = \frac{\textit{output}}{\textit{input}} \quad (3.6.1.1)$$

No entanto, na grande maioria das vezes, as unidades organizacionais possuem múltiplos *inputs* e *outputs*, tornando o conceito um pouco mais complexo. Desta forma, a medida de eficiência pode ser definida como a razão da soma ponderada de *outputs* (*outputs* virtuais) pela soma ponderada de *inputs* (*inputs* virtuais):

$$ET = \frac{\text{soma ponderada de } \textit{outputs}}{\text{soma ponderada de } \textit{inputs}} \quad (3.6.1.2)$$

O objetivo é maximizar a eficiência de cada empresa de acordo com suas possibilidades e levando-se em consideração todas as demais empresas. Tem-se, então, o “CCR *ratio form*” o qual pode ser matematicamente expresso por **(M3.1)**:

$$\mathbf{Max}_{u,v} \quad \theta = \frac{u_1 y_{01} + u_2 y_{02} + \dots + u_r y_{0r}}{v_1 x_{01} + v_2 x_{02} + \dots + v_i x_{0i}} = \frac{\text{output virtual}}{\text{input virtual}}$$

S.T.

$$\frac{u_1 y_{j1} + u_2 y_{j2} + \dots + u_r y_{jr}}{v_1 x_{j1} + v_2 x_{j2} + \dots + v_i x_{ji}} \leq 1, \quad j = 1, \dots, n$$

$$u_r \geq \varepsilon, \quad r = 1, \dots, s$$

$$v_i \geq \varepsilon, \quad i = 1, \dots, m$$

(M3.1)

onde:

- y_{jr} = quantidade do r -ésimo *output*, da j -ésima DMU;
- x_{ji} = quantidade do i -ésimo *input*, da j -ésima DMU;
- r = indexador de *outputs*, ($r=1, \dots, s$);
- i = indexador de *inputs*, ($i=1, \dots, m$);
- j = indexador de DMU's, ($j=1, \dots, n$);
- 0 = DMU que está sendo analisada (j_0);
- u_r = peso atribuído ao r -ésimo *output*;
- v_i = peso atribuído ao i -ésimo *input*;
- ε = infinitésimo positivo.

A principal característica deste modelo é que os pesos u e v são as incógnitas ou variáveis de decisão. Eles serão escolhidos de forma a maximizar a eficiência da unidade j_0 . Neste ponto, vale mencionar uma carência metodológica do modelo, ou seja, ele pode produzir um número infinito de soluções ótimas. Por exemplo, se (u^*, v^*) for a solução ótima, então é fácil ver

que $(\alpha u^*, \alpha v^*)$ também é uma outra solução ótima $\forall \alpha \in \mathbf{R}_+^*$. Para resolver este problema, é definida uma relação equivalente que separa o conjunto de soluções viáveis do modelo DEA (M3.1) em classes equivalentes. A transformação desenvolvida em [23] para o problema de programação linear fracionária, procura selecionar uma solução representativa (isto é, uma solução (u, v) que satisfaça (i) – M3.2) de cada DMU equivalente e produza o seguinte problema de programação linear, que pode ser resolvido pelo método simplex:

$$\mathbf{Max}_{u,v} \quad w_0 = \sum_r u_r y_{0r}$$

S.T.

$$\sum_i v_i x_{0i} = 1 \quad \text{(i)}$$

$$\sum_r u_r y_{jr} - \sum_i v_i x_{ji} \leq 0, \quad j = 1, \dots, n$$

$$u_r, v_i \geq \varepsilon,$$

(M3.2)

Pelo estudo de Programação Linear, sabe-se que o modelo (M3.2) possui uma forma denominada dual a qual está registrada no modelo (M3.3):

$$\mathbf{Min}_{\theta, \lambda} \quad \theta$$

S.T.

$$\theta x_0^T - X^T \lambda \geq 0,$$

$$Y^T \lambda \geq y_0^T,$$

$$\lambda \geq 0;$$

θ : sem restrição.

(M3.3)

Percebe-se no modelo **(M3.3)** que o conjunto de restrições é equivalente ao conjunto T , referenciado anteriormente como **Conjunto das Possibilidades de Produção** (vide equação 3.2.1). Prosseguindo, como o objetivo é encontrar um valor ótimo θ^* que permita minimizar os *inputs* mantendo os *outputs* constantes, tem-se, então, o modelo *CCR input-oriented*. Este modelo é conhecido na literatura como **Envelope**.

Tanto maximizando o modelo **(M3.2)** como minimizando o modelo **(M3.3)** e considerando um conjunto de restrições, é possível obter o *score* de eficiência da unidade j_0 . Vale assinalar que para os modelos com orientação a *inputs* (*input oriented*), $\theta \in (0, 1]$.

A Figura 3.3 mostra a geometria da envoltória determinada por este modelo para o caso de apenas um *input* e um *output*:

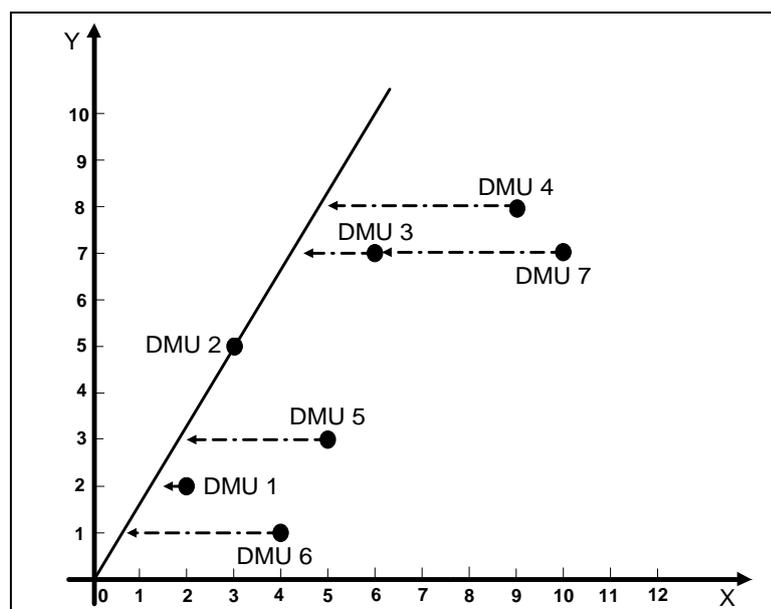


Figura 3.3: Envoltória determinada pelo modelo *CCR input-oriented* [18].

Uma análise na Figura 3.3 permite constatar uma fronteira de produção não paramétrica sobre os pontos dados de tal forma que estes pontos pertençam ou estejam abaixo desta fronteira. O ponto sobre a fronteira indica que a respectiva unidade é tecnicamente eficiente, ou seja, $\theta^* = 1$, valor este conseguido através da resolução do modelo **(M3.3)**.

Verifica-se, também, que cada DMU considerada ineficiente tenta aproximar-se da envoltória dirigindo-se no sentido de minimizar o *input* mantendo inalterado o *output*.

Neste momento, são necessárias algumas definições do modelo CCR:

Definição 3.1 – Eficiência CCR.

A DMU j_0 é dita CCR eficiente se existe no mínimo uma solução ótima (u^*, v^*) , com $v^* > 0$ e $u^* > 0$, para a qual $\theta^* = 1$.

Caso contrário a DMU j_0 é dita CCR ineficiente.

Então, CCR ineficiente significa que ou (i) $\theta^* < 1$ ou (ii) $\theta^* = 1$ e no mínimo um elemento de (u^*, v^*) é zero para cada solução ótima do PPL.

O sub-conjunto composto de DMU's CCR eficientes é denominado **conjunto de referências** ou **grupos pares** para as outras DMU's CCR ineficientes. A partir de conhecimentos envolvendo programação linear, estas unidades são identificadas resolvendo o modelo **(M3.3)** onde cada $\lambda > 0$ corresponde à unidade tecnicamente eficiente.

Uma importante característica de DEA, é que a formulação consegue fazer uma comparação do desempenho da j_0 -ésima DMU com todas as outras DMU's. Assim sendo, a técnica DEA mede a eficiência relativa, não absoluta, das entidades.

Por fim, basicamente o modelo CCR resulta em uma envoltória linear por partes com Retorno Constante de Escala (RCE) significando que um aumento proporcional em todos os *inputs* resulta em um mesmo aumento proporcional nos *outputs*.

3.6.2

Modelo DEA BCC *input-oriented*

Em várias oportunidades, pode-se encontrar empresas que operam em escala ótima adotando o modelo CCR, o qual assume estimar uma fronteira de produção com Retorno Constante de Escala. O índice de eficiência denominado *score* (**Eficiência Técnica Global**), que é obtido pela resolução deste modelo, permite fazer comparações sobre as mesmas. Entretanto, em casos onde o RCE

não prevalece devido a uma competição imperfeita, restrições nas finanças, etc, conjectura-se que as unidades devam ser comparadas, dado suas escalas de operação ou, no mínimo, deve ser mencionado que esta é a razão da sua ineficiência. Em tais circunstâncias a Eficiência Técnica Global pode ser decomposta Eficiência Técnica Pura Local e Eficiência de Escala.

Para permitir este processo, Banker et al [11] sugeriram uma extensão do modelo CCR e formularam um modelo que permitisse Retorno Variável de Escala (RVE). Este modelo, denominado BCC, em referência aos seus idealizadores, é uma outra formulação derivada do modelo **(M3.3)** acrescido da restrição de convexidade.

A seguir, tem-se, então, a formulação do modelo BCC representado por **(M3.4)**:

$$\text{Min } \theta$$

$$\theta, \lambda$$

S. T.

$$\theta x_0^T - X^T \lambda \geq 0,$$

$$Y^T \lambda \geq y_0^T,$$

$$e\lambda = 1; \quad (\text{ii})$$

θ : sem restrição.

(M3.4)

Esta restrição de convexidade representada no modelo **(M3.4)** por **(ii)**, reduz o **Conjunto de Possibilidades de Produção (T)** viável e converte uma tecnologia de ganho constante de escala em uma tecnologia de ganho variável de escala.

Semelhante à definição do modelo CCR, tem-se:

Definição 3.2 – Eficiência BCC.

A DMU j_0 é BCC eficiente se $\theta^* = 1$ e todas as folgas são zero. Caso contrário, a DMU j_0 é identificada como BCC ineficiente.

O sub-conjunto composto de DMU's BCC eficientes é denominado **conjunto de referências** ou **grupo de pares** para as outras DMU's BCC ineficiente.

Pela Figura 3.4 pode-se fazer uma comparação entre as fronteiras de produção estabelecidas pelo modelo CCR e BCC, em um ambiente composto de 7 DMU's cada qual com um *input* e um *output*.

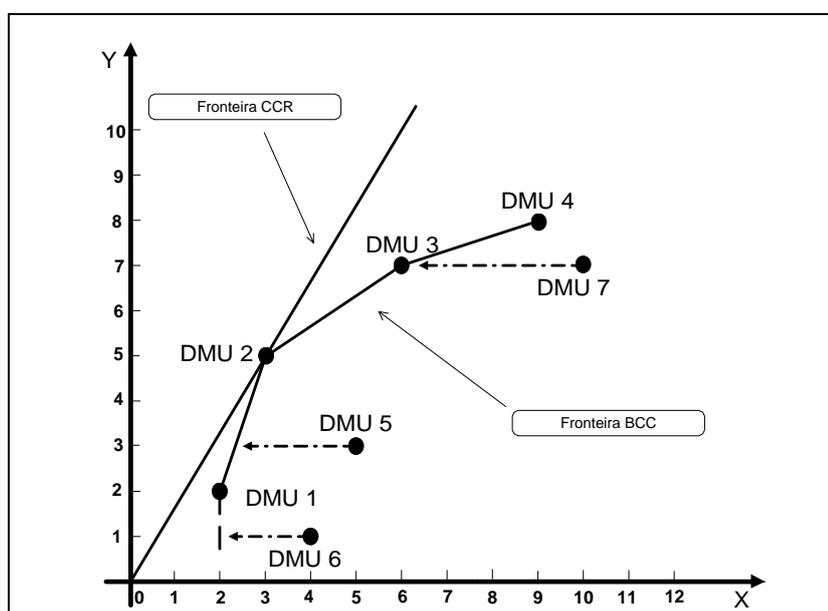


Figura 3.4: Envoltória determinada pelos modelos CCR e BCC (*input-oriented*) [18].

Por inspeção visual da Figura 3.4, percebe-se que no segundo caso, ou seja, no modelo BCC, a técnica DEA forma uma superfície convexa, constituída por planos interligados os quais envelopam os pontos, tornando-os mais próximos da fronteira do que no caso da superfície RCE. O resultado deste processo é a obtenção de uma medida de eficiência maior ou igual a do modelo CCR. Desta forma, é possível avaliar um maior número de DMU's como sendo tecnicamente eficientes.

As DMU's 5, 6 e 7 são tecnicamente ineficientes quando analisadas segundo o modelo BCC o qual adota RVE. Assim sendo, é possível comparar cada uma destas DMU's com as DMU's eficientes projetando-as sobre um ponto

$(\hat{x}_{jt1}, \hat{y}_{jt1})$ da envoltória determinada pelas jt -DMU's eficientes. Cabe ressaltar que este ponto particular de projeção $(\hat{x}_{jt1}, \hat{y}_{jt1})$ tanto é dependente do modelo DEA escolhido quanto da orientação. Por exemplo, em um modelo BCC *input-oriented*, o objetivo é caminhar em direção à fronteira através da redução proporcional de *inputs*. Este conceito pode ser facilmente interpretado na Figura 3.4 onde a orientação para *inputs* consiste nas seguintes projeções:

- A DMU 5 incide sobre um ponto que reflete a combinação linear convexa das DMU's 1 e 2 as quais são seu **conjunto de referência**;
- A DMU 6 incide sobre um ponto que representa a DMU 1 a qual é seu **conjunto de referência**;
- A DMU 7 incide sobre um ponto que representa a DMU 3 a qual é seu **conjunto de referência**;

Se uma DMU é plenamente eficiente à luz dos *scores* de eficiência CCR e BCC, significa que ela está operando com a máxima produtividade. É o caso da DMU 2 na Figura 3.4 acima. No entanto, se ela é somente BCC eficiente, então ela está operando de forma eficiente apenas localmente devido à dimensão de escala da DMU.

3.6.3

Outros modelos DEA

É importante assinalar que há vários outros modelos DEA [18, 23, 99]. No entanto, o objetivo desta obra não é explicar exhaustivamente todas as formulações. Nesse contexto, é elencado outros modelos registrados na literatura:

- Modelo CCR *output oriented*;
- Modelo BCC *output oriented*;
- Modelo Aditivo;
- Modelo Multiplicativo;
- Modelos não-radiais.

É fundamental ressaltar que resultados bastante diferentes podem ser encontrados ao resolver o DEA não somente com relação à seleção de diferentes modelos, mas também com relação à orientação adotada, ou seja, *input-oriented* ou *output-oriented*.

O ponto crucial para o analista é, portanto, escolher entre vários modelos, aquele que melhor atende às suas especificações. Outras questões relevantes referem-se à escolha das variáveis (a seleção das variáveis de *input* e *output* deve ser feita de forma criteriosa), representação dos dados, interpretação dos resultados e conhecimento de limitações.

Devido às importantes conseqüências sobre o estudo e os resultados obtidos, a aplicação de um modelo básico DEA deveria ser feita somente depois de cuidadosas considerações. Como exemplo, é importante que o número de DMU's seja, no mínimo, $2mxs$ onde mxs ³ é o produto do número de *inputs* e *outputs*, pois caso isso não ocorra pode-se identificar uma DMU como sendo eficiente quando na verdade a mesma não é [27].

3.6.4

Seleção de variáveis *inputs* e *outputs*

Embora não seja este o foco principal deste trabalho, é bastante oportuno abordar a questão da escolha das variáveis. Em concordância com o que foi citado, a elaboração de um bom resultado de eficiência está intimamente relacionado com a escolha das mesmas, ou seja, elas devem ser bastante significativas de modo a ter a capacidade de influenciar de forma decisiva no resultado de eficiência de cada unidade.

Boussofiâne et al [15] mencionam que embora os modelos DEA permitam combinar um conjunto de pesos comuns para os *inputs* e *outputs*, naturalmente

³ este assunto é bastante referenciado na literatura. Em [9] encontramos uma outra regra de bolso:

$$n \geq \max\{m \times s, 3(m + s)\}$$

n = número de DMU's;

m = número de *inputs*;

s = número de *outputs*.

eles não conseguem evitar o problema de uma seleção equivocada dos *inputs* e *outputs*. Claramente, qualquer recurso usado por uma organização poderia ser incluído como um *input*.

Uma DMU precisa ter a capacidade de converter *inputs* em *outputs* de modo que estes *outputs* possam ser produzidos em diferentes níveis de qualidade.

Continuando, fatores ambientais os quais podem afetar a produção de *outputs* precisam ser identificados e incluídos no modelo a ser analisado para que a comparação das eficiências seja satisfatória. Estes fatores ambientais podem ser medidos diretamente ou indiretamente através do uso de uma *proxy*. Um fator ambiental o qual adiciona recursos, pode ser incluído como um *input* enquanto um que requer recursos pode ser incluído como um *output*.

3.6.5

Técnicas baseadas em Restrições aos Pesos

Cada vez mais DEA tem sido utilizada para resolver problemas relacionados com o cotidiano. Existem várias pesquisas envolvendo diversos tipos alternativos de análise de restrições dos pesos fazendo deste aspecto uma das áreas mais promissoras de desenvolvimento em DEA, sem mostrar qualquer sinal de saturação [2].

No modelo básico, reporte-se ao modelo primal (M3.2), foi observado que os pesos u 's e v 's eram variáveis do problema e estavam restritas a serem maiores ou iguais a algum infinitésimo positivo ε de forma a não permitir que algum *input* ou *output* fosse totalmente ignorado no processo de calcular a eficiência das respectivas DMU's. Charnes et al [18] advogam que o processo de permitir uma certa flexibilidade na escolha dos pesos é freqüentemente apresentada como uma prática vantajosa em aplicações da metodologia DEA; uma especificação *a priori* dos pesos não é necessária, e cada DMU é avaliada à luz da melhor perspectiva.

No entanto, em algumas situações, esta completa flexibilidade pode levar a conseqüências indesejáveis, ou seja, o especialista pode avaliar uma DMU como eficiente em situações difíceis de justificar como tal. Procurando sanar essa carência metodológica, foi sugerido incorporar a opinião do especialista com relação à importância dos *inputs* e *outputs* bem como sobre as relações entre as

variáveis. Assim sendo, o modelo DEA torna-se mais plausível e uma conclusão mais coerente acerca do desempenho das DMU's é obtida.

A gênese de toda essa teoria é creditada a Thompson et al [89] que modificaram o modelo DEA original incorporando restrições aos pesos através de regiões de segurança (*Assurance Region* - AR). Segundo Thanassoulis & Allen [87], diversos trabalhos têm sido propostos na literatura com este fim. Vale citar que estes métodos têm diferido apenas com relação à especificação dos limites permissíveis dos valores dos pesos evitando, assim, que o problema se torne inviável.

Seguindo o critério adotado por Allen et al [2], as técnicas para impor restrições aos pesos podem ser classificadas em três categorias, a saber:

1. Restrições Diretas sobre os Pesos:

Esta técnica consiste em adicionar ao modelo (M3.2) restrições aos pesos. Há duas maneiras para tal: (i) impor limites absolutos inferiores e superiores sobre u_r e v_i [28, 79]; (ii) Regiões de Segurança 1 e 2 (*Assurance Region*) [88, 89, 90].

2. Ajuste sobre os níveis de *inputs*- *outputs* observados (modelos ajustados com a utilização de dados artificiais):

Para este caso, é gerada uma base de dados artificiais através de uma transformação linear nos dados originais (*inputs/outputs*). O método Cone-Ratio [18] e a técnica de Relações Ordinais [44] são duas técnicas que emulam restrições aos pesos através de transformações lineares nos dados.

3. Restrição sobre a flexibilidade dos *inputs* e *outputs* virtuais.

Os tópicos anteriores utilizam restrições aos pesos sem levar em consideração a magnitude dos *inputs* e *outputs*. Neste contexto, duas técnicas são propostas na literatura para este fim: *Contingent Restrictions on Weights* - Pedraja et al [70] e *Relative Importance of factors to a DMU* – Wong & Beasley [96]. Esta última é baseada em proporções e consiste em estabelecer limites (inferior e superior) para a

razão de determinado *output* (*input*) virtual pela soma de todos os *outputs* (*inputs*) virtuais.

No entanto, há controvérsias sobre este tema. Dyson et al [27] ressaltam que os resultados obtidos dos modelos com restrições aos pesos não podem ter a mesma interpretação que quando na forma original, sem restrições aos pesos. Isto ocorre porque quando são impostas restrições aos pesos, a interpretação do **Conjunto de Possibilidades de Produção**, (3.2.1) e (3.2.2), passa a não ser mais válida. A característica de modelo radial também é perdida.

Este trabalho segue a metodologia proposta por Pedraja et al [70]. Como a mesma foi adaptada, resolveu-se denominá-la: *Adjusted Contingent Restrictions on Weights*.

3.7

Adjusted contingent restrictions on weights

Contingent restrictions on weights é uma técnica DEA de restrições aos pesos proposta por Pedraja et al [70]. Segundo os autores, a razão do termo “*contingent*” é evidenciar o fato de que o valor dos pesos depende da quantidade de *inputs* e *outputs* que cada DMU consome/produz. Em um contexto de *inputs*, tem-se que a contribuição do *i*-ésimo *input* à *j*0 DMU é $v_i x_{0i}$. Procedendo desta maneira, pode-se assegurar que apenas os *inputs* e *outputs* que realmente contribuem significativamente para o desempenho de uma DMU, com relação aos custos totais ou benefícios da referida DMU, são incluídos na análise.

Matematicamente, Pedraja et al [70] propõem a seguinte forma para as restrições de um conjunto de *inputs*:

$$c_i v_1 x_{j1} \leq v_i x_{ji} \leq d_i v_1 x_{j1}, \quad \forall i > 1, \quad j = 1, \dots, j_0, \dots, n \quad (3.7.1)$$

onde c_i e d_i são escolhidos pelo analista. Vale sublinhar que restrições similares podem ser aplicadas para o conjunto de *outputs*.

A adaptação mencionada está baseada na questão de considerar apenas as empresas DEA-eficientes, ou seja:

$$c_i v_1 x_{jt1} \leq v_i x_{jti} \leq d_i v_1 x_{jt1}, \quad \forall i > 1, \quad jt = 1, \dots, n \quad (3.7.2)$$

ou

$$v_i x_{jti} \leq k v_{i'} x_{ji'}, \quad \forall (i \neq i') > 1, \quad jt = 1, \dots, n \quad (3.7.3)$$

onde k é o coeficiente de regressão linear entre dois *inputs* quaisquer.

É relevante registrar que as restrições adicionadas, ao modelo, também emulam as regiões de segurança Tipo I e Tipo II (ARI e ARII), conforme Thanassoulis & Allen [87].

Por último, é pertinente destacar algumas vantagens desta adaptação:

- Melhor discriminação das empresas eficientes;
- Empresas ineficientes (de acordo com os modelos DEA – **M3.2** ou **M3.3**) não são consideradas nesta análise. Este procedimento proporciona uma significativa redução do esforço computacional;
- É possível obter resultados muito similares quando comparado com a técnica de DMU's artificiais (baseada na opinião do especialista). Esta constatação pode ser comprovada em Diallo et al [24].

4

Conceitos Básicos de Estatística Bayesiana e Simulação Estocástica

4.1

Elementos de inferência Bayesiana

Gamerman [34] define que tanto o modelo Bayesiano quanto o freqüentista trabalham na presença de observações x , cujo valor é inicialmente incerto e descrito através de uma distribuição de probabilidades $f(x|\theta)$. A quantidade θ^4 serve como indexador da família de distribuições das observações representando características de interesse que se deseja conhecer para poder ter uma descrição completa do processo. Continuando, Gamerman [34] alerta para o fato de que o principal interesse de estudo é a determinação do seu valor, não sendo portanto, um simples indexador.

Vale citar que para fazer inferência sobre θ dado x , é necessário determinar uma distribuição de probabilidade conjunta:

$$p(x, \theta) = f(x|\theta)p(\theta) \quad (4.1.1)$$

Após observar $X = x$, pode-se utilizar o Teorema de Bayes para determinar a distribuição a posteriori:

$$p(\theta|x) = \frac{f(x|\theta)p(\theta)}{\int f(x|\theta)p(\theta)d\theta} \quad (4.1.2)$$

ou

$$p(\theta|x) \propto f(x|\theta)p(\theta) \quad (4.1.3)$$

A razão por inserir a proporcionalidade em (4.1.2) resultando (4.1.3), é que o denominador (constante normalizadora) de (4.1.2) não é função de θ .

O objetivo dos tópicos seguintes é justificar estes preceitos de forma mais detalhada.

⁴ Em vários casos θ pode ser multi-dimensional.

4.2

Distribuição a priori

Medrano & Migon [62] reportam que a implementação da inferência segundo a escola Bayesiana requer que as distribuições a priori sejam especificadas para as quantidades de interesse. Neste contexto, Koop [54] sugere que o especialista colha diversas opiniões de profissionais do setor em estudo para, assim, tirar conclusões mais convincentes. No entanto, quando as prioris dos cientistas são fortes e divergentes, Gamerman & Migon [35] citam que é fundamental produzir uma análise neutra visando gerar um referencial.

Vale notar que em algumas situações, o pesquisador pode ter o sentimento de que a informação disponível para avaliar a distribuição a priori não existe. Este conhecimento pobre ou ignorância é tema de muita discussão na literatura. Afinal, é importante ressaltar que nunca se pode estar em completo estado de ignorância. Para expressar a idéia de pouco conhecimento a respeito do parâmetro, uma classe de prioris denominadas não-informativas tem se mostrado bastante útil. De acordo com Gamerman & Migon [35], inicialmente foram propostas prioris distribuídas uniformemente ($p(\theta) \propto k$, $-\infty < \theta < \infty$), o que implica não favorecer nenhum valor particular de θ . Neste caso, qualquer análise Bayesiana é baseada fundamentalmente na distribuição amostral. Entretanto, é importante sublinhar algumas dificuldades com relação a esta escolha:

- $p(\theta)$ é uma distribuição imprópria, ou seja: $\int p(\theta)d\theta \rightarrow \infty$, se o intervalo de θ for ilimitado;
- $p(\theta)$ não é invariante à reparametrização.

Um método alternativo para determinar prioris não-informativas, para um conjunto de parâmetros, é conhecido como Regra de Jeffrey. Esta regra é baseada na escolha de uma priori proporcional à raiz quadrada do determinante de uma matriz de informação esperada de Fisher. Ehlers [29] explica que o conceito de informação, aqui, está associada a uma espécie de curvatura média da função de verossimilhança no sentido de que quanto maior esta curvatura, mais precisa é a informação contida na verossimilhança.

É digno registrar que esta classe de prioris não-informativas é invariante, porém, eventualmente imprópria. Como, na prática, o objetivo é analisar a distribuição a posteriori, esta constatação é dispensável. No entanto, é de suma importância, antes de realizar qualquer inferência sobre a distribuição a posteriori, certificar que a mesma é uma distribuição própria. Mais uma vez, de acordo com Gamerman & Migon [35], um outro obstáculo dessa técnica é não satisfazer o Princípio da Verossimilhança (este princípio postula que toda informação contida em um determinado experimento, está contida na função de verossimilhança).

Vale notar que muitas vezes pode-se dividir a especificação de uma priori através de estágios. Ehlers [29] explica que, além de facilitar esta especificação, esta abordagem é natural em determinadas situações experimentais.

Para ilustrar, seja a distribuição a priori de θ a qual depende dos valores dos hiperparâmetros ϕ . Pode-se, então, escrever a distribuição condicional $p(\theta|\phi)$ ao invés da marginal $p(\theta)$. Além disso, ao invés de fixar valores para os hiperparâmetros, determina-se uma distribuição a priori $p(\phi)$. Neste ponto, completa-se o segundo estágio da hierarquia. Matematicamente, a distribuição marginal a priori de θ pode ser encontrada através da seguinte integral:

$$p(\theta) = \int p(\theta, \phi) d\phi = \int p(\theta|\phi)p(\phi) d\phi \quad (4.2.1)$$

Gamerman [34] cita que não há limitação quanto ao número de estágios. Contudo, quanto mais alto o estágio, mais complexo fica a especificação das distribuições. Assim sendo, geralmente são especificados dois (2) ou três (3) estágios.

4.3

Distribuição amostral (Função de Verossimilhança)

Gamerman & Migon [35] mencionam que a função de verossimilhança é a função que associa a cada valor θ , o valor $f(x|\theta)$. Portanto, ao fixar um valor para x e variar os valores de θ , pode ser observado a plausibilidade (ou verossimilhança) de cada um dos valores de θ . Vale lembrar que a inferência Bayesiana obedece ao Princípio da Verossimilhança. É através desta função que o conhecimento a priori sobre θ é modificado. Neste ponto, é importante lembrar

que em uma análise científica, é apropriado que esta função seja dominante com relação à distribuição a priori. Afinal, é coerente pensar que qualquer informação relevante sobre o parâmetro θ seja obtida através da realização de um experimento e, conseqüentemente, através dos dados.

4.4

Distribuição a posteriori

Gelman et al [38] apontam que a distribuição a posteriori contém todas as informações probabilísticas do parâmetro θ . Desta forma, a elaboração de um gráfico para investigar as características dessa distribuição pode ser considerado um procedimento bastante útil. Concomitante, pode-se resumir a informação (do parâmetro de interesse) contida na distribuição a posteriori através de valores numéricos (estimativas pontuais) tais como: média, mediana, moda. Ehlers [29] enfatiza que é importante reduzir a incerteza dessas estimativas através do intervalo de credibilidade (intervalo de confiança Bayesiano). Este intervalo de credibilidade expressa a probabilidade de θ estar em um intervalo pré-definido, condicional aos dados observados. Vale lembrar que é possível construir inúmeros intervalos. No entanto, aquele que contempla o menor comprimento possível é denominado de Máxima Densidade a Posteriori (MDP). De acordo com Cespedes [17], MDP é o intervalo em que a densidade para todo ponto pertencente ao intervalo é maior do que para todo não pertencente a ele.

4.5

Exemplo

Este tópico tem como finalidade elucidar as definições abordadas nos tópicos 4.2, 4.3 e 4.4, objetivando uma melhor compreensão das mesmas. Para tal, seja o exemplo registrado em Gamerman & Migon [35] onde dois físicos A e B, visando obter uma estimativa mais acurada de uma constante física θ , concordam em fazer um experimento em laboratório. O físico A é bastante experiente e determina a seguinte distribuição a priori para o parâmetro θ :

$$\theta_A \sim N(\mu_A, \tau_A^2) \quad (4.5.1)$$

Já o físico B tem pouca experiência (pouco conhecimento) e dessa forma opta por utilizar uma priori não- informativa:

$$\theta_B \sim N(\mu_B, \tau_B^2) \quad (4.5.2)$$

Vale salientar que estes parâmetros indexadores da família de distribuições a priori (μ_A , μ_B , τ_A^2 e τ_B^2) são denominados na literatura como hiperparâmetros para diferenciá-los dos parâmetros de interesse (θ_A, θ_B). Para este estudo, os valores adotados são os seguintes:

$$\mu_A = 900; \mu_B = 800; \tau_A^2 = (20)^2; \tau_B^2 = (80)^2$$

As curvas na Figura 4.1 mostram estas densidades a priori para A e B:

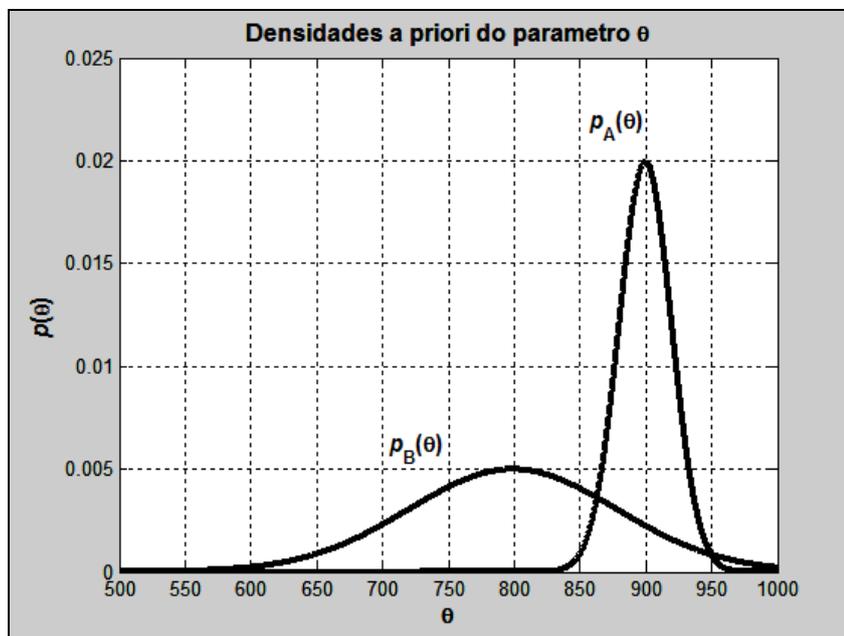


Figura 4.1: Densidades a priori do parâmetro θ .

Sabendo-se que a distribuição amostral $(X|\theta) \sim N(\theta, (40)^2)$, tem-se

$$f(\mathbf{x}|\theta, \sigma^2) = l(\theta, \sigma^2; \mathbf{x}) = \prod_{j=1}^n f(x_j|\theta, \sigma^2) \quad (4.5.3)$$

ou

$$l(\theta, \sigma^2; \mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \theta)^2 \right) \quad (4.5.4)$$

Suponha-se que o resultado de uma simples observação seja $x = 850$; então a função de verossimilhança é mostrada na Figura 4.2:

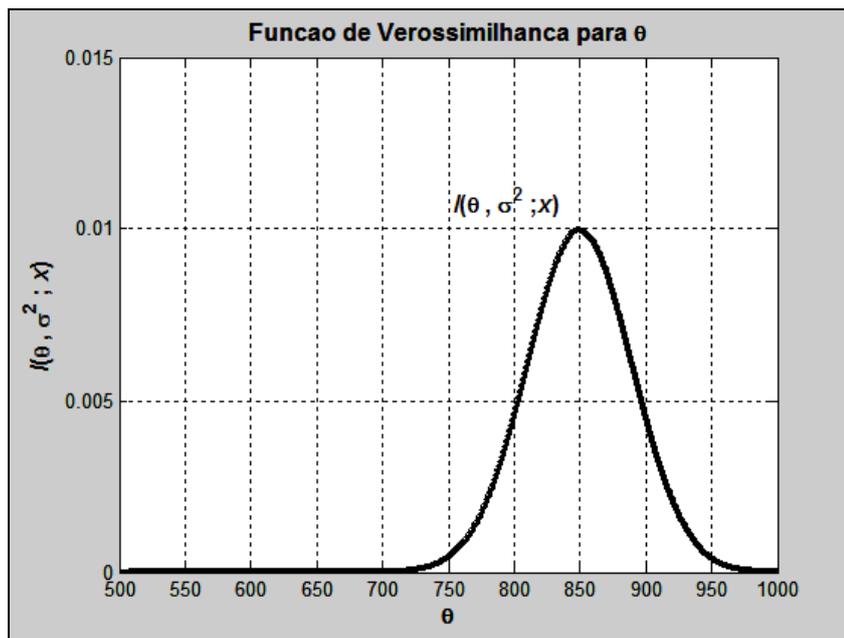


Figura 4.2: Função de Verossimilhança para θ .

A abordagem Bayesiana é adaptativa e permite revisar a distribuição a priori dos parâmetros com novas informações, obtendo-se ao final uma distribuição a posteriori (π). Por inspeção visual da Figura 4.3, pode-se perceber que a aquisição da amostra, e consequentemente a introdução da informação, modifica a distribuição a priori com uma considerável redução com relação à incerteza sobre os parâmetros (θ_A, θ_B) .

Dessa forma, tem-se:

- Para o físico A: $(\theta|X = 850) \sim N(890, (17,9)^2)$;
- Para o físico B: $(\theta|X = 850) \sim N(840, (35,7)^2)$.

Estas densidades são mostradas na Figura 4.3:

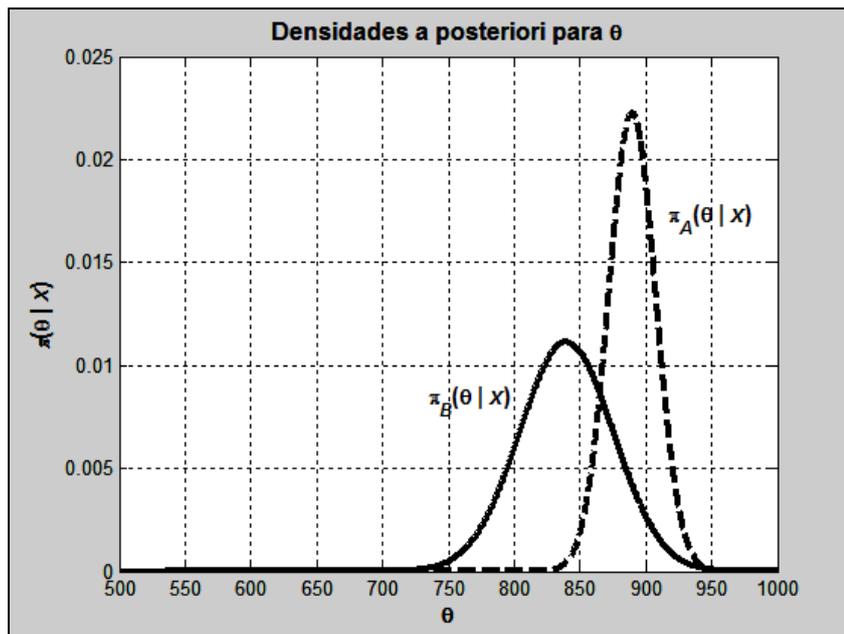


Figura 4.3: Densidades a posteriori do parâmetro θ .

Este exemplo ilustra uma situação importante onde, para um dado modelo observacional, as distribuições a priori e a posteriori pertencem a uma mesma classe de distribuições, no caso a distribuição normal. Esta propriedade de preservação da classe de distribuições define conjugação. Neste caso, a atualização do conhecimento que se tem sobre o parâmetro do modelo envolve apenas uma mudança nos hiperparâmetros.

Gamerman [34] alerta para o cuidado com a utilização indiscriminada de prioris conjugadas. Essencialmente, o problema é que a priori conjugada nem sempre é uma representação correta da incerteza a priori. Sua utilização está, muitas vezes, associada à tratabilidade analítica decorrente.

4.6

Obtenção de resumos de interesse através de simulação

Conforme citado no tópico 4.4, em diversas situações é interessante resumir a informação descrita na distribuição a posteriori em termos de esperanças de funções particulares do parâmetro θ , i.e [29]:

$$\mathbf{E}[g(\theta)|X] = \int g(\theta) p(\theta|x) d\theta \quad (4.6.1)$$

sendo $g(\cdot)$ uma função qualquer. Por exemplo, se $g(\theta) = \theta$, então $\mathbf{E}[g(\theta)|X]$ representa a média a posteriori.

É digno de registro que estes cálculos de integrais, em geral, não são analiticamente tratáveis. Esta questão, por muitos anos, atrasou o desenvolvimento de modelos Bayesianos. No entanto, com a evolução computacional, houve uma explosão de trabalhos utilizando métodos baseados em simulação. Os itens seguintes abordam alguns destes métodos: Integração de Monte Carlo, Amostragem por Importância e Monte Carlo via Cadeias de Markov (MCMC). É importante sublinhar que estes textos são de caráter introdutório. Para mais detalhes, o leitor é convidado a consultar as seguintes referências: Gamerman [34], Robert & Casella [76].

4.6.1

Integração via Monte Carlo

De acordo com as informações anteriores, em diversas situações a equação (4.6.1) é altamente complexa. Desta forma, os métodos estatísticos tradicionais falham e as inferências são feitas utilizando simulações. Basicamente, a simulação de Monte Carlo refere-se a qualquer simulação que envolve o uso de números aleatórios.

Em linhas gerais, o objetivo da Integração via Monte Carlo é calcular a seguinte integral:

$$\mathcal{G} = \mathbf{E}(X) = \int xf(x)dx. \quad (4.6.1.1)$$

Uma maneira plausível consiste em gerar aleatoriamente amostras $\{X_t, t = 1, \dots, T\}$ da distribuição de X e fazer a aproximação:

$$\mathbf{E}(X) \approx T^{-1} \sum_{t=1}^T x_t \quad (4.6.1.2)$$

Este conceito pode ser extrapolado considerando

$$\mathcal{G} = \mathbf{E}(\phi(X)) = \int \phi(x)f(x)dx. \quad (4.6.1.3)$$

Novamente, seja $\{X_t, t = 1, \dots, T\}$ uma amostra aleatória independente e identicamente distribuída (*iid*) da distribuição $f(x)$. Então,

$$\hat{g} = T^{-1} \sum_{t=1}^T \phi(x_t) \quad (4.6.1.4)$$

converge para $\mathbf{E}(\phi(X))$ com probabilidade um (1) quando $T \rightarrow \infty$. Como \hat{g} é uma média amostral de observações independentes, temos que $\mathbf{Var}(\hat{g}) = \frac{\mathbf{Var}(\phi(X))}{T}$. Mas $\mathbf{Var}(\phi(X))$ pode ser estimada através da variância amostral de $\phi(X_t)$, de modo que o erro padrão estimado para \hat{g} é dado por

$$se(\hat{g}) = \sqrt{\frac{1}{T(T-1)} \sum_{t=1}^T (\phi(X_t) - \hat{g})^2}. \quad (4.6.1.5)$$

Para ilustrar, seja o seguinte exemplo apresentado por Roberts [78]. Para este caso, X e Y têm distribuição conjunta (distribuição normal bivariada) dada por:

$$f(x, y) = \frac{1}{2\pi\sqrt{\sigma_x^2\sigma_y^2(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\} \quad (4.6.1.6)$$

sendo ρ o coeficiente de correlação entre X e Y , $-1 < \rho < 1$, $x \in (-\infty, \infty)$, $y \in (-\infty, \infty)$, $(\mu_x, \mu_y) \in \mathbf{R}^2$ e $(\sigma_x, \sigma_y) \in \mathbf{R}_+^2$.

O valor de $\mathbf{Pr}(X < 1, Y < 1)$ pode ser estimado considerando inicialmente

$$\int I_A(x, y) f(x, y) dx dy \quad (4.6.1.7)$$

e na seqüência calculando:

$$T^{-1} \sum_{t=1}^T I_A(x_t, y_t) \quad (4.6.1.8)$$

onde I_A corresponde a uma função indicadora: $A = \{(x, y) : x < 1, y < 1\}$. Há várias maneiras para simular uma distribuição normal bivariada. Conceitualmente, a forma mais fácil é utilizar os métodos baseados em misturas que exploram a questão da densidade conjunta das variáveis aleatórias (X, Y) , ou seja:

$$f(x, y) = f(y|x)f(x) \quad (4.6.1.9)$$

Dessa forma, o par (x, y) pode ser gerado em dois (2) passos: inicialmente gera-se $X=x$ a partir da sua distribuição marginal e, então, y é gerado condicionalmente do valor obtido de x , $(Y|X = x)$.

Utilizando $\rho = 0,5$, $\mu_x = \mu_y = 0$, $\sigma_x = \sigma_y = 1$, pode-se obter a seguinte distribuição condicional:

$$f(y|x) = N\left(\mu_y + \rho \frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y \sqrt{1 - \rho^2}\right) \quad (4.6.1.10)$$

A estimativa da probabilidade requerida após mil (1000) iterações ($T = 1.000$) é dada por 0,762. Vale registrar que quanto maior for T , mais acurada será a estimativa. A Figura 4.4 fornece o *scatterplot* dos valores simulados:

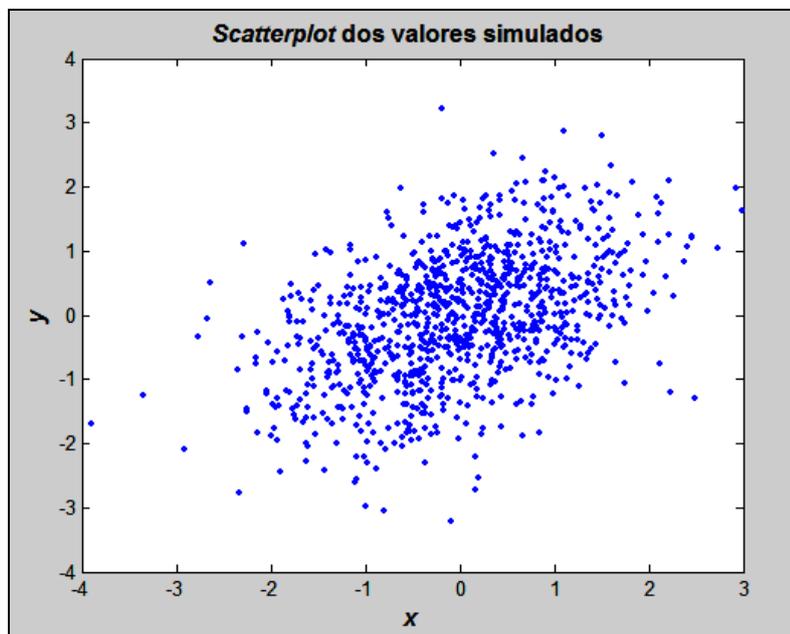


Figura 4.4: *Scatterplot* dos valores simulados.

Pode ser constatado que este método é bastante simples e fácil de se usar mesmo em problemas envolvendo altas dimensões. De acordo com Roberts [78], o custo por esta simplicidade é que a variância é alta.

Uma extensão da Integração via Monte Carlo é dada pela Amostragem por Importância. Esse método é útil quando não é possível gerar amostras de $f(x)$.

Matematicamente, podemos re-escrever a equação (4.6.1.3) como:

$$\mathcal{I} = \mathbf{E}(\phi(X)) = \int \psi(x)g(x)dx \quad (4.6.1.11)$$

onde $\psi(x) = \frac{\phi(x)f(x)}{g(x)}$ e $g(x)$, denominada de função de importância. A escolha desta função é baseada no fato de que a mesma seja fácil de amostrar. Gamerman [34] mostra que uma escolha ótima, em termos de minimizar a variância é usar $g \propto \phi \times f$.

Procedendo dessa maneira, o próximo passo consiste em estimar \mathcal{G} :

$$\hat{\mathcal{G}} = T^{-1} \sum_{t=1}^T \psi(X_t) \quad (4.6.1.12)$$

cuja variância é dada por:

$$\mathbf{Var}(\hat{\mathcal{G}}) = T^{-1} \int (\psi(x) - \mathcal{G})^2 g(x) dx \quad (4.6.1.13)$$

Tierney [91] comenta que, por várias décadas, o método Amostragem por Importância foi utilizado em um contexto Bayesiano para estimar características das distribuições a posteriori tais como a média ou o desvio-padrão de uma função de θ . Entretanto, em diversos casos é impossível amostrar diretamente da distribuição a posteriori ou mesmo descobrir uma função de importância plausível. Desta forma, os métodos baseados em simulação de Monte Carlo tornam-se praticamente inviáveis. Neste cenário, uma poderosa técnica denominada Monte Carlo via Cadeias de Markov (*Markov chain Monte Carlo* - MCMC) tem sido amplamente utilizada para resolver esta espécie de problema.

4.7

Monte Carlo via Cadeias de Markov (MCMC)

Monte Carlo via cadeias de Markov (MCMC) é uma poderosa ferramenta para gerar variáveis aleatórias.

De acordo com Casella & George [16], MCMC foi, inicialmente, proposto por Metropolis et al [65]. Este trabalho pioneiro resultou em desenvolvimentos futuros: Hastings [48], Geman & Geman [40] e Tanner & Wong [86]. Segundo Gilks et al [43], a apresentação da aplicabilidade desses métodos em uma gama de problemas estatísticos convencionais, especificamente em modelagem Bayesiana, é creditado aos artigos de Gelfand & Smith [36] e Gelfand et al [37].

Durante várias décadas, diversos artigos e livros têm sido publicados com a intenção de expor a aplicação dessa metodologia em diversos problemas reais

complexos, como exemplo: Spall [80], Doucet et al [25], Chib & Greenberg [20], Tierney [91], Besag et al [14], Casella & George [16], Gamerman [34], Robert & Casella [76], Gilks et al [41].

O protótipo do problema é como se segue. Considere uma seqüência de variáveis aleatórias $\{X_t, t=1, \dots, T\}$ tal que em cada tempo $t \geq 0$, o próximo estado X_{t+1} é gerado condicionalmente da distribuição $P(X_{t+1}|X_t)$, que é denominada kernel de transição da cadeia e não depende de t se a cadeia for homogênea no tempo. Isto é, a distribuição de X_{t+1} depende somente do estado atual da cadeia, X_t . Uma realização desta seqüência de variáveis aleatórias $\{X_t, t=1, \dots, T\}$ é conhecida na literatura como Cadeias de Markov. É importante relatar que X_0 representa alguma condição inicial.

Roberts [77] faz uma breve revisão para que a distribuição de X_t convirja para a distribuição estacionária. Assim sendo, ele explica que a cadeia precisa satisfazer três (3) importantes propriedades:

- i. irredutível;
- ii. aperiódica;
- iii. positiva recorrente.

A propriedade (iii) pode ser considerada a mais importante. Afinal, esta propriedade está relacionada com a existência de uma distribuição estacionária tal que se X_t é gerado desta distribuição, então todos os valores subsequentes também o serão.

Para ilustrar, reporte-se à equação (4.6.1.3) e considere $f(\cdot)$ como sendo a distribuição estacionária. Gilks et al [43] advogam que com o crescimento de t , os pontos amostrados X_t se parecerão mais e mais com amostras dependentes de $f(\cdot)$. Ignorando as primeiras l iterações da cadeia (período denominado “*burn-in*”), pode-se usar a cadeia de Markov para estimar $\mathbf{E}(\phi(X))$, onde $X \sim f(\cdot)$, da seguinte maneira:

$$\bar{\phi} = \frac{1}{T-l} \sum_{t=l+1}^T \phi(X_t) \quad (4.7.1)$$

Tal procedimento é chamado de média ergódica. É importante explicitar que a sua convergência, garantida pelo famoso teorema ergódico (vide Roberts [77]), é quase certa ou com probabilidade 1.

Neste ponto, é necessário descobrir como construir uma cadeia de Markov tal que a sua distribuição estacionária seja precisamente a distribuição de interesse⁵. Os tópicos seguintes têm por finalidade apresentar brevemente dois algoritmos MCMC mais utilizados para este fim: Algoritmo de Metropolis-Hastings e o Amostrador de Gibbs.

4.7.1

Algoritmo de Metropolis-Hastings

De acordo com Haykin, [49], o algoritmo de Metropolis-Hastings, introduzido nos primórdios da ciência da computação e que tem sido utilizado até nos dias atuais pela comunidade Física, baseia-se em um método de Monte Carlo para a simulação estocástica de uma coleção de átomos em equilíbrio a uma dada temperatura.

O algoritmo Metropolis-Hastings é um mecanismo que produz uma cadeia de Markov $\{X_t, t = 1, \dots, T\}$ com as seguintes características:

- a distribuição limite de interesse (π) pode ser conhecida apenas parcialmente. Assim, a constante normalizadora não precisa ser conhecida, pois será cancelada no quociente;
- nenhuma fatoração de (π) é necessária;
- os métodos são facilmente implementados, e
- a sequência de amostras é obtida através de uma cadeia de Markov.

A partir destas informações, seja a cadeia de Markov $\{X_t, t \geq 0\}$. Para o algoritmo de Metropolis-Hastings, a cada tempo $t \geq 0$ o próximo estado, X_{t+1} , é escolhido, primeiramente, amostrando um ponto candidato Y da distribuição proposta $q(\cdot | X_t)$.

⁵ É importante esclarecer que, em muitos textos, a distribuição de interesse (distribuição conjunta a posteriori) é denotada por π .

O ponto candidato Y é, então, aceito como o próximo estado da cadeia com probabilidade dada por:

$$\alpha(X_t, Y) = \min \left\{ 1, \frac{\pi(Y)q(X_t|Y)}{\pi(X_t)q(Y|X_t)} \right\} \quad (4.7.1.1)$$

Vale salientar que o uso de (4.7.1.1) é fundamental para garantir que se construa uma cadeia ergódica e que, portanto, se obtenha a distribuição estacionária. Então, se o ponto candidato Y for aceito, o próximo estado será $X_{t+1} = Y$. Caso contrário, ou seja, se o candidato for rejeitado, a cadeia não se moverá, isto é, $X_{t+1} = X_t$.

Vale explicar que quando a distribuição proposta for simétrica ($q(Y|X) = q(X|Y)$), tem-se um caso especial do algoritmo Metropolis-Hastings nomeado de algoritmo Metropolis. Neste caso, a equação (4.7.1.1) é sensivelmente modificada transformando-se em:

$$\alpha(X_t, Y) = \min \left\{ 1, \frac{\pi(Y)}{\pi(X_t)} \right\} \quad (4.7.1.2)$$

Em termos práticos, os seguintes passos podem ser utilizados para gerar uma seqüência de amostras aleatórias através do algoritmo Metropolis-Hastings [41]:

- i. inicie a cadeia em X_t e ajuste $t = 0$;
- ii. gere um ponto candidato $Y \sim q(\cdot|X_t)$;
- iii. gere $U \sim Unif(0,1)$;
- iv. se $U \leq \alpha(X_t, Y)$ então ajuste $X_{t+1} = Y$, senão ajuste $X_{t+1} = X_t$;
- v. incremente $t = t+1$ e repita os passos (ii) até (v).

Para esclarecer este algoritmo, suponha-se que a distribuição de Cauchy padrão, sem a constante normalizadora, seja distribuição limite de interesse:

$$\pi(x) \propto \frac{1}{(1+x^2)}; \quad -\infty < x < \infty \quad (4.7.1.3)$$

A distribuição proposta, q , para este exemplo é uma normal com a média dada pelo valor prévio e o desvio-padrão atribuído arbitrariamente como $\sigma = 2$.

Os resultados obtidos pela simulação são apresentados nas Figuras 4.5 e 4.6. Vale informar que a Figura 4.5 fornece um gráfico construído com o valor da

variável e o número de iterações. Pode ser verificado que este gráfico está de acordo com Melo & Ehlers [64] que afirmam que o mesmo deve apresentar um caráter aleatório.

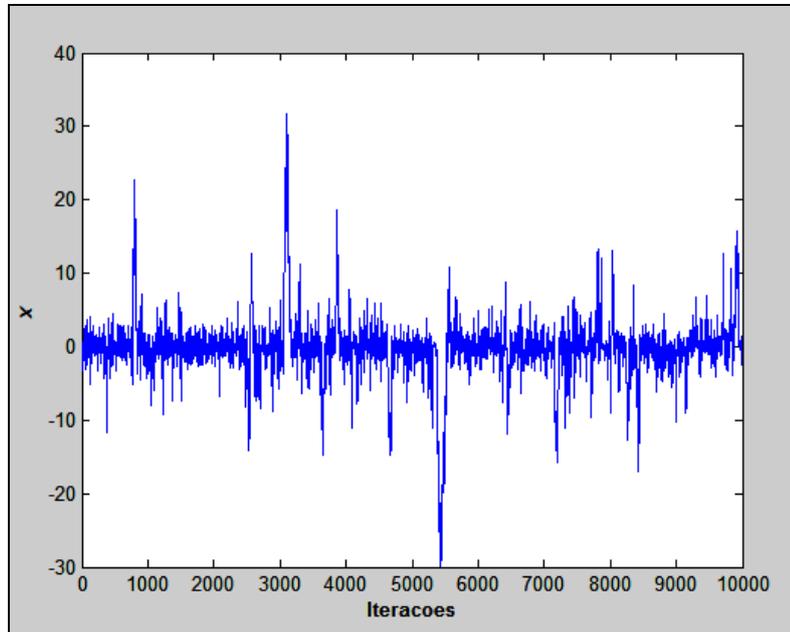


Figura 4.5: Valores simulados através do algoritmo Metropolis-Hastings.

A seguir, a Figura 4.6 mostra o histograma da densidade estimada (aproximada) com a curva correspondente à função densidade de probabilidade verdadeira (equação 4.7.1.3).

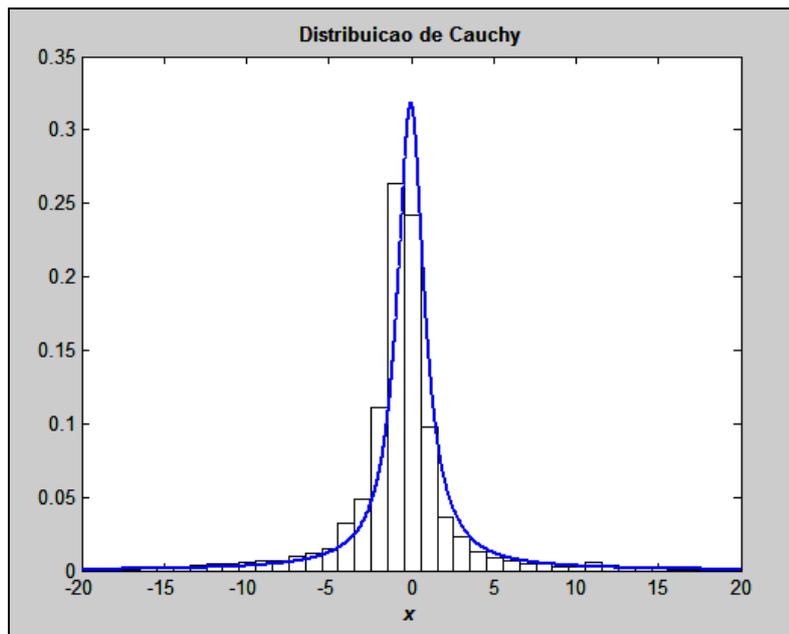


Figura 4.6: Algoritmo Metropolis-Hastings na geração da distribuição de Cauchy.

Não é difícil perceber que, neste exemplo, as amostras poderiam ter sido geradas diretamente da distribuição de Cauchy, sem a necessidade de aplicação do algoritmo Metropolis-Hastings.

4.7.2

Amostrador de Gibbs

O Amostrador de Gibbs foi originalmente concebido dentro de um contexto de reconstrução de imagens por Geman & Geman [40] e está contido em uma grande classe de esquemas de simulação estocástica que utilizam cadeias de Markov [34].

Embora ele seja um caso especial do algoritmo Metropolis-Hastings, o mesmo apresenta duas particularidades, a saber:

- Todos os pontos gerados são aceitos;
- Existe a necessidade de conhecermos a distribuição condicional completa.

Entenda por distribuição condicional completa, a distribuição da d -ésima componente de um determinado parâmetro condicionada a todas as outras componentes.

Segundo Gamerman [34], o amostrador de Gibbs é, essencialmente, um esquema iterativo de amostragem de uma cadeia de Markov, cujo *kernel* de transição é formado pelas distribuições condicionais completas.

Para descrever o funcionamento deste algoritmo, suponha que o objetivo seja determinar algumas características da distribuição marginal de uma variável aleatória θ_1 em um problema constituído de mais duas variáveis aleatórias: θ_2 e θ_3 . Este tratamento é abordado em Casella & George [16]. Matematicamente, tem-se:

$$p(\theta_1) = \iint p(\theta_1, \theta_2, \theta_3) d\theta_2 d\theta_3 \quad (4.7.2.1)$$

Analisando a equação (4.7.2.1), percebe-se que para obter a distribuição marginal, é necessário o cálculo de uma integral sobre todas as outras variáveis. Extrapolando e considerando que em muitas aplicações θ seja um vetor de parâmetros particionado como $\theta = (\theta_1, \dots, \theta_d)$, a integral (4.7.2.1) torna muito complicada (e algumas vezes impossível) de se avaliar. O amostrador de Gibbs é uma forma de estimar densidades marginais através de simulação. Uma questão que é importante frisar refere-se ao fato de que a partição escolhida seja fácil de amostrar. Neste ponto, é interessante apresentar os procedimentos deste algoritmo [34]:

- i. inicie o contador de iterações da cadeia $t=1$ e arbitre valores iniciais $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})$;
- ii. obtenha um novo valor de $\theta^{(t)} = (\theta_1^{(t)}, \dots, \theta_d^{(t)})$ a partir de $\theta^{(t-1)}$ através de sucessivas gerações de valores:

$$\begin{aligned} \theta_1^{(t)} &\sim p\left(\theta_1 \mid \theta_2^{(t-1)}, \dots, \theta_d^{(t-1)}\right) \\ \theta_2^{(t)} &\sim p\left(\theta_2 \mid \theta_1^{(t)}, \theta_3^{(t-1)}, \dots, \theta_d^{(t-1)}\right) \\ &\vdots \\ \theta_d^{(t)} &\sim p\left(\theta_d \mid \theta_1^{(t)}, \dots, \theta_{d-1}^{(t)}\right) \end{aligned}$$

- iii. mude o contador de t para $t+1$ e retorne a (ii) até a convergência.

Assim, cada iteração se completa após d movimentos ao longo dos eixos coordenados das componentes de θ . Após a convergência, os valores resultantes

formam uma amostra de $p(\theta)$ ⁶. Vale notar que, mesmo em problemas envolvendo grandes dimensões, são utilizadas simulações univariadas ou em bloco o que, em geral, é uma vantagem computacional [29]. Este fato tem contribuído de forma significativa para a aplicação desta metodologia, principalmente na área de econometria aplicada com ênfase Bayesiana (vide, por exemplo, Koop et al [57]).

Para ilustrar, seja um processo de Poisson com priori hierárquica proposto por Ehlers [30]:

$$Y_j \sim \text{Poi}(\lambda), \quad j = 1, \dots, n \quad (4.7.2.2)$$

$$\lambda \sim \text{Exp}(\beta) \quad (4.7.2.3)$$

$$\beta \sim \text{Gama}(C, D) \quad (4.7.2.4)$$

A partir dessas informações, pode-se escrever a equação da distribuição conjunta de Y , λ e β como:

$$p(\mathbf{y}, \lambda, \beta) = \prod_{j=1}^n p(y_j | \lambda) p(\lambda | \beta) p(\beta) \quad (4.7.2.5)$$

Ao observar $Y_j = y_j$ e fazendo $z = \sum_{j=1}^n y_j$, tem-se as seguintes

distribuições condicionais completas:

$$\pi(\lambda | \mathbf{y}, \beta) \propto \exp[-\lambda(\beta + n)] \lambda^z \quad (4.7.2.6)$$

$$\pi(\beta | \mathbf{y}, \lambda) \propto \exp[-\beta(\lambda + D)] \beta^C \quad (4.7.2.7)$$

Então, as distribuições condicionais completas são dadas por:

$$\pi(\lambda | \mathbf{y}, \beta) \sim \text{Gama}(z + 1, \beta + n) \quad (4.7.2.8)$$

$$\pi(\beta | \mathbf{y}, \lambda) \sim \text{Gama}(1 + C, \lambda + D) \quad (4.7.2.9)$$

Em seguida, é realizada uma simulação com cem (100) dados ($n = 100$) de um processo com média aproximadamente quatro (4), $C = D = 0,1$. Por fim, rodou-se mil (1000) iterações ($T = 1.000$). A Figura 4.7 fornece os resultados da simulação:

⁶ a distribuição p não precisa, necessariamente, ser uma distribuição a posteriori [16].

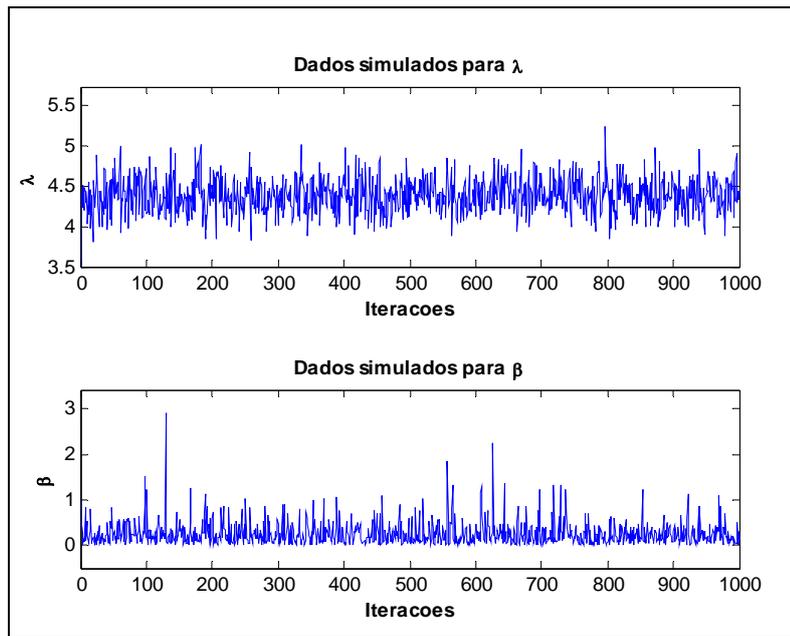


Figura 4.7: Dados simulados para λ e β utilizando o algoritmo de Gibbs.

Descartando-se as duzentas (200) primeiras iterações ($l = 200$), é possível calcular as médias amostrais do parâmetro λ ao longo das iterações. A Figura 4.8 representa graficamente os resultados obtidos. Por inspeção visual da mesma, verifica-se que é notória a convergência do mesmo.

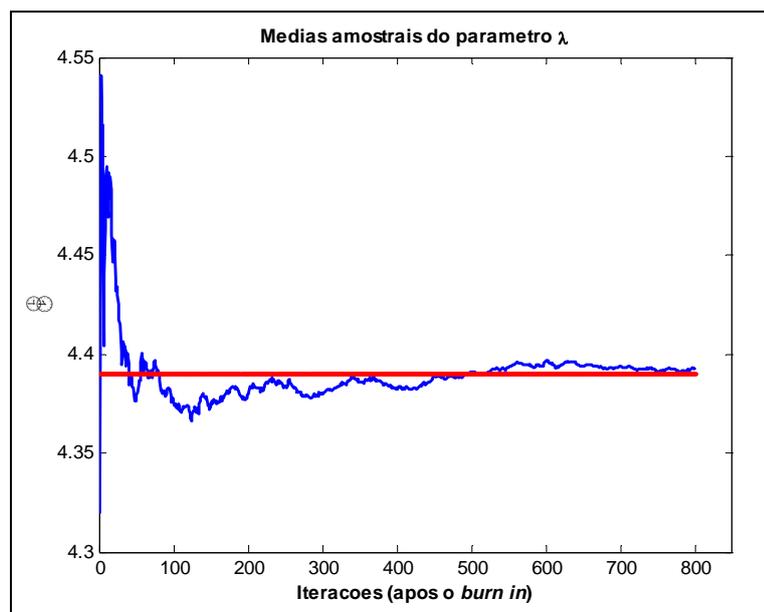


Figura 4.8: Resultado das médias amostrais para o parâmetro λ .

Agora, suponha que o objetivo seja estimar a distribuição marginal da variável aleatória X . Para tal, seja a seguinte distribuição conjunta de X , Y e N [16]:

$$p(x, y, n) \propto \binom{n}{x} y^{x+\alpha-1} (1-y)^{n-x+\beta-1} \exp(-\lambda) \frac{\lambda^n}{n!} \quad (4.7.2.10)$$

$$x = 0, 1, \dots, n; \quad 0 \leq y \leq 1; \quad n = x, x+1, \dots$$

Como a distribuição marginal de X não tem forma fechada, faz-se necessário o uso de simulação para a determinação da mesma. Utilizando o mesmo raciocínio do primeiro exemplo, têm-se as seguintes distribuições condicionais completas:

$$\pi(x|y, n) \sim \text{Bin}(n, y) \quad (4.7.2.11)$$

$$\pi(y|x, n) \sim \text{Beta}(x+\alpha, n-x+\beta) \quad (4.7.2.12)$$

$$\pi(n|x, y) \propto \exp[-\lambda(1-y)] \frac{[\lambda(1-y)]^{n-x}}{(n-x)!} \quad (4.7.2.13)$$

A Figura 4.9 mostra a distribuição estimada considerando que os valores utilizados foram: $\lambda = 16$, $\alpha = 2$, $\beta = 4$, $T = 5.000$ e $l = 200$.

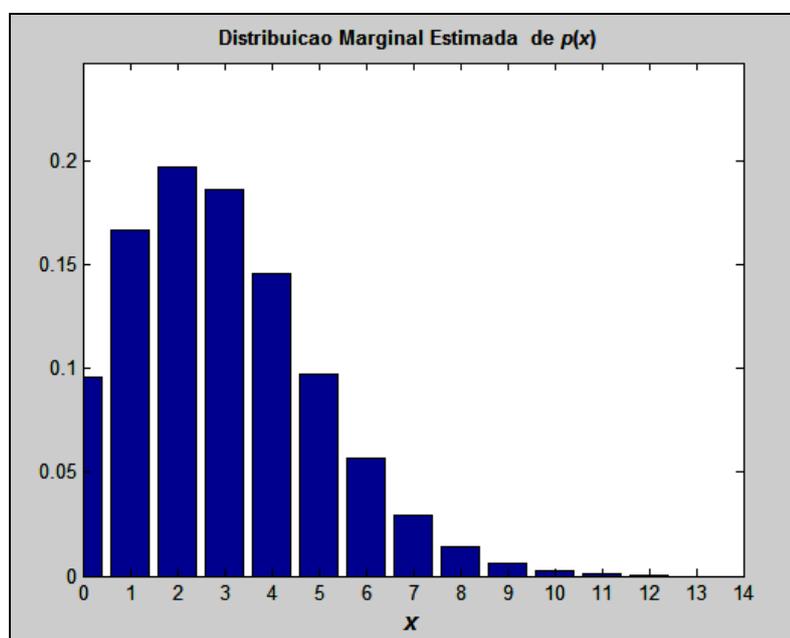


Figura 4.9: Distribuição marginal estimada de $p(x)$.

Até o momento, pôde ser constatado que a implementação dos algoritmos MCMC é relativamente simples. Esta facilidade constantemente tem resultado em erros crassos, dentre os quais Gelman [38] cita:

1. modelagem inapropriada;
2. erros de programação;
3. baixa convergência dos parâmetros do modelo.

Embora todos os três (3) itens citados sejam de suma importância, o terceiro tópico é um assunto que tem despertado muito interesse, principalmente na comunidade estatística com enfoque Bayesiano, onde esses algoritmos têm sido exaustivamente utilizados.

4.7.3

Monitoração da convergência

Uma vez que a simulação é realizada, é necessário verificar a convergência das seqüências simuladas. O excelente trabalho de Cowles & Carlin [22] faz uma revisão e também compara diversos métodos propostos na literatura para este fim. Ainda de acordo com Cowles & Carlin [22], os mesmos podem ser concentrados em duas (2) grandes áreas:

- i. Teórica: esta área envolve o conhecimento de uma matemática sofisticada como também uma infinidade de cálculos laboriosos que precisam ser repetidos para cada modelo sob certas considerações;
- ii. Aplicada: esta área foca nos resultados produzidos pela simulação.

Um método bastante simples, para monitorar a convergência, é o gráfico de autocorrelação amostral. Neste gráfico é interessante que o decaimento das barras seja o mais rápido possível. No entanto, muitas das vezes, devido à natureza Markoviana das amostras colhidas, isso não acontece. Nestes casos, algumas soluções propostas são [36,42, 74]:

- a) Reparametrização;
- b) *Thinning*;
- c) Cadeias Paralelas;

- d) Aumentar o número de iterações da cadeia e, conseqüentemente, o *burn-in*.

A Figura 4.10, gerada a partir das simulações realizadas no primeiro exemplo do tópico 4.7.2, mostra os gráficos de autocorrelação amostral para os parâmetros λ e β .

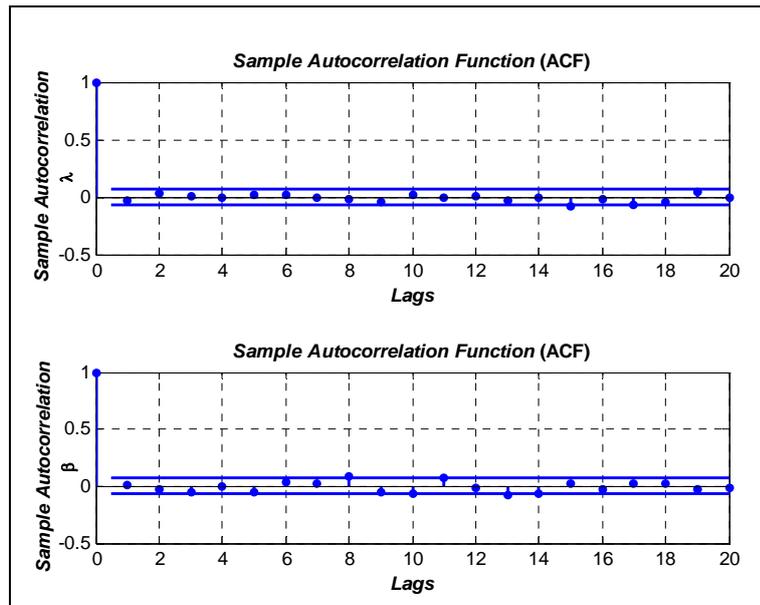


Figura 4.10: Gráficos de autocorrelação amostral dos parâmetros λ e β .

Pela característica exposta, ou seja, decaimento rápido, não é difícil conjecturar a convergência dos parâmetros e, neste caso, que as suas estimativas são confiáveis.

5

Modelos Econométricos

5.1

Fronteira Estocástica de Produção

Segundo Bauer [13], os primeiros modelos de fronteira estocástica (também conhecida na literatura como modelos de erro-composto) foram introduzidos por Meeusen & van den Broeck [63], Aigner et al [1] e Battese & Corra [12]. O modelo básico é dado por:

$$y_j = h(\mathbf{x}_j; \boldsymbol{\beta}) \exp(v_j - u_j) \quad (5.1.1)$$

Admitindo que $h(\mathbf{x}_j, \boldsymbol{\beta})$ seja linear nos logaritmos, pode-se obter, após a aplicação de uma transformação logarítmica em (5.1.1), a seguinte função de produção [21]:

$$\ln y_j = \beta_0 + \sum_{i=1}^m \beta_i \ln x_{ji} + \sum_{i \leq k=1}^m \sum_{k=1}^m \beta_{ik} \ln x_{ji} \ln x_{jk} + v_j - u_j \quad (5.1.2)$$

A equação (5.1.2) é denominada na literatura por função Translog. Quando os produtos cruzados são nulos, tem-se um caso particular denominado função Cobb-Douglas. De posse dessas informações, pode-se definir a parte determinística da fronteira:

- $\ln y_j$ - logaritmo natural do *output* da j -ésima DMU ($j = 1, \dots, n$);
- $\ln x_{ji}$ - logaritmo natural do i -ésimo *input* da j -ésima DMU (incluindo o intercepto);
- $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \dots \ \beta_m]^T$ - um vetor de parâmetros desconhecidos a serem estimados.

Já o desvio entre o nível de produção observado e a parte determinística da fronteira de produção é dado pela combinação de duas componentes:

- u_j - um erro que assume apenas valores não negativos e que captura o efeito da ineficiência técnica;
- v_j - um erro simétrico que captura qualquer choque aleatório não gerenciável.

Fazendo $\varepsilon_j = v_j - u_j$, tem-se a seguinte equação de regressão:

$$\ln y_j = \beta_0 + \sum_{i=1}^m \beta_i \ln x_{ji} + \sum_{i \leq k=1}^m \sum_{k=1}^m \beta_{ik} \ln x_{ji} \ln x_{jk} + \varepsilon_j \quad (5.1.3)$$

Em geral, admite-se que v_j seja uma variável aleatória *i.i.d.* (independente e identicamente distribuída) com distribuição simétrica, usualmente uma distribuição normal. A hipótese de simetria da distribuição de v_j está traduzida pelo fato de que as condições ambientais desfavoráveis e favoráveis são igualmente prováveis. Por sua vez, a distribuição de probabilidade para a componente u_j ($u_j \geq 0$) não é evidente, a priori, e por isso pode ser especificada segundo diferentes formas, por exemplo, Meeusen & van den Broeck [63] utilizaram a distribuição exponencial, Aigner et al [1] recomendaram a distribuição *Half-Normal*, Stevenson [84] propôs a Normal-Truncada e, finalmente, Greene [45] sugeriu a distribuição Gama. Mais recentemente Medrano & Migon [62] utilizaram a distribuição log Normal. É importante sublinhar que v_j é independente de u_j . Outra questão que merece ser registrada, refere-se ao fato da distribuição do erro composto, ε_j , ser assimétrica. Portanto, modelos de regressão com erros simetricamente distribuídos e com média zero não são apropriados para estimar a função fronteira estocástica em (5.1.3).

Para estimar a eficiência técnica de cada produtor é necessário extrair da estimativa de ε_j as estimativas de v_j e u_j . Conforme explicado, isto requer alguma suposição sobre as distribuições das duas componentes de erro. Estas suposições permitem a utilização do método de máxima verossimilhança na estimação dos parâmetros do modelo de regressão e também são importantes no procedimento de estimação da eficiência técnica dos produtores, baseado na distribuição condicional de u_j em ε_j .

Para ilustrar, seja a formulação básica dos modelos de fronteira estocástica, onde as seguintes hipóteses acerca dos termos aleatórios v_j e u_j são admitidas:

$$(H1) \quad v_j \sim iid N(0, \sigma_v^2) \Rightarrow f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left(-\frac{v^2}{2\sigma_v^2}\right)$$

$$(H2) \quad u_j \sim iid N^+\left(\eta_j, \tau^2\right) \Rightarrow f(u) = \frac{1}{\sqrt{2\pi}\tau\Phi\left(\frac{\eta}{\tau}\right)} \exp\left(-\frac{(u-\eta)^2}{2\tau^2}\right) \quad (\text{Normal-Truncada}$$

não negativa)

$$(H3) \quad u_j \text{ é independente de } v_j$$

$$(H4) \quad v_j \text{ e } u_j \text{ são independentes de } x_j.$$

Ressalta-se que, na hipótese **(H2)**, se $\eta_j = 0$, a distribuição de $f(u)$ é denominada *Half-Normal*. A partir da hipótese **(H3)** tem-se que a distribuição conjunta de v_j e u_j é o produto das densidades marginais em **(H1)** e **(H2)**. Esta formulação, Normal-truncada, foi inicialmente proposta por Stevenson [84], visando amenizar as críticas que envolviam as distribuições *Half Normal* e Exponencial. Afinal, conforme reportado em [84], estas distribuições têm a moda em zero tendo assim, uma alta probabilidade de que as ineficiências sejam nulas. A partir dessas considerações e dado que $\varepsilon_j = v_j - u_j$, tem-se a distribuição conjunta de u_j e ε_j :

$$f(u, \varepsilon) = \frac{1}{2\pi\tau\sigma_v\Phi\left(\frac{\eta}{\tau}\right)} \exp\left(-\frac{1}{2}\left[\frac{(u-\eta)^2}{\tau^2} + \frac{(\varepsilon+u)^2}{\sigma_v^2}\right]\right) \quad (5.1.4)$$

Integrando a função densidade (5.1.4) em relação a u , obtém-se a distribuição marginal de ε , dada por:

$$f(\varepsilon) = \frac{1}{\sqrt{\sigma_v^2 + \tau^2}} \cdot \phi\left(\frac{\varepsilon + \eta}{\sqrt{\sigma_v^2 + \tau^2}}\right) \cdot \Phi\left(\frac{-\tau(\varepsilon + \eta)}{\sigma_v\sqrt{\sigma_v^2 + \tau^2}} + \frac{\eta\sqrt{\sigma_v^2 + \tau^2}}{\sigma_v\tau}\right) \cdot \left[\Phi\left(\frac{\eta}{\tau}\right)\right]^{-1} \quad (5.1.5)$$

onde $\phi(\cdot)$ é a função de densidade da Normal-padrão e $\Phi(\cdot)$ é a função de distribuição acumulada da Normal-padrão.

A função densidade (5.1.5) pode ser parametrizada como:

$$f(\varepsilon) = \frac{1}{\sigma} \cdot \phi\left(\frac{\varepsilon + \eta}{\sigma}\right) \cdot \Phi\left(\frac{-\lambda(\varepsilon + \eta)}{\sigma} + \eta^*\right) \cdot \left[\Phi\left(\frac{\eta}{\tau}\right)\right]^{-1} \quad (5.1.6)$$

onde $\sigma = \sqrt{\sigma_v^2 + \tau^2}$, $\lambda = \tau/\sigma_v$ e $\eta^* = \frac{\eta\sqrt{\sigma_v^2 + \tau^2}}{\sigma_v\tau}$.

A distribuição, $f(\varepsilon)$, é negativamente assimétrica com média e variância dadas por:

$$\mathbf{E}(\varepsilon) = -\mathbf{E}(u) = \left[\Phi\left(\frac{\eta}{\tau}\right)\right]^{-1} \left(-\frac{\eta}{2} - \frac{\tau}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\eta}{\tau}\right)^2\right)\right) \quad (5.1.7)$$

$$\mathbf{Var}(\varepsilon) = \eta^2 \frac{a}{2} \left(1 - \frac{a}{2}\right) + \frac{a}{2} \left(\frac{\pi - a}{2}\right) \tau^2 + \sigma_v^2 \quad (5.1.8)$$

onde $a = \left[\Phi\left(\frac{\eta}{\tau}\right)\right]^{-1}$

Dado que $\varepsilon_j = y_j - \left(\beta_0 + \sum_{i=1}^m \beta_i \ln x_{ji} + \sum_{i \leq k=1}^m \sum_{j=1}^m \beta_{ik} \ln x_{ji} \ln x_{jk}\right)$ e usando a

função densidade em (5.1.6), o logaritmo da função de verossimilhança para uma amostra com n DMU's é escrita como:

$$L = \text{constante} - n \ln \sigma - n \ln \left[\Phi\left(\frac{\eta}{\tau}\right)\right] + \sum_j \ln \left[\Phi\left(\frac{-\lambda(\varepsilon_j + \eta)}{\sigma} + \eta^*\right)\right] - \frac{1}{2} \sum_j \left(\frac{\varepsilon_j + \eta}{\sigma}\right)^2 \quad (5.1.9)$$

As estimativas de máxima verossimilhança, são os valores de τ , σ_v e β_i ($i=0, \dots, m$) que maximizam a equação (5.1.9). Estas estimativas são assintoticamente consistentes.

A parametrização do logaritmo da verossimilhança em termos de $\sigma^2 = \tau^2 + \sigma_v^2$ e $\gamma = \tau^2 / (\tau^2 + \sigma_v^2) \in [0, 1]$ ajuda na maximização da função de verossimilhança e facilita a interpretação do modelo.

Quando $\gamma=0$, a ineficiência é inexistente, pois σ_v^2 domina a variância do erro composto e os desvios em relação à fronteira devem-se unicamente a ruídos aleatórios. Por outro lado, quando $\gamma=1$, τ^2 domina a variância do erro composto e

o desvio em relação à fronteira deve-se inteiramente à ineficiência. Portanto, pode-se utilizar o teste da razão da verossimilhança (*LR test*) para testar a hipótese nula $\mathbf{H0}$: $\gamma=0$. Porém, como γ pode cair nos limites do intervalo $[0,1]$, a distribuição da razão de verossimilhança é uma mistura de qui-quadrados, i.e., aproximadamente uma χ_1^2 .

Para estimar a eficiência técnica de cada produtor é necessário ter uma estimativa de u_j , a componente do erro que captura o efeito da ineficiência técnica. Esta estimativa pode ser obtida a partir dos resíduos $\hat{\varepsilon}_j$, pois $\varepsilon_j = v_j - u_j$. Usando as funções densidade em (5.1.4) e (5.1.6), Jondrow et al [52] estimaram u_j como sendo a média ou a moda da distribuição condicionada, $f(u|\varepsilon) = f(u, \varepsilon)/f(\varepsilon)$. A equação (5.1.10) fornece os resultados encontrados:

$$f(u|\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma_* \left[1 - \Phi\left(-\frac{\tilde{\eta}}{\sigma_*}\right)\right]} \exp\left\{-\frac{(u - \tilde{\eta})^2}{2\sigma_*^2}\right\} \Rightarrow u_i|\varepsilon_i \sim N^+(\tilde{\eta}, \sigma_*^2) \quad (5.1.10)$$

onde $\tilde{\eta}$ é função dos valores teóricos τ , σ^2 , ou seja: $\tilde{\eta} = (-\tau^2\varepsilon + \eta\sigma_v^2)/\sigma^2$ e $\sigma_*^2 = \tau^2\sigma_v^2/\sigma^2$.

Então, a estimativa pontual de u_j , também definida em função dos valores teóricos τ , σ^2 , é calculada pela média de $f(u|\varepsilon)$:

$$\hat{u}_j = E(u_j|\varepsilon_j) = \sigma_* \left[\frac{\tilde{\eta}_j}{\sigma_*} + \frac{\phi\left(\frac{\tilde{\eta}_j}{\sigma_*}\right)}{1 - \Phi\left(-\frac{\tilde{\eta}_j}{\sigma_*}\right)} \right] \quad (5.1.11)$$

ou pela moda de $f(u|\varepsilon)$

$$\hat{u}_i = M(u_i|\varepsilon_i) = \tilde{\eta}_j \quad , \text{ se } \tilde{\eta}_j \geq 0 \quad (5.1.12)$$

$\hat{u}_i = 0$, caso contrário

Finalmente, tem-se a estimativa da eficiência técnica da j -ésima DMU:

$$\mathbf{ET}_j = \exp(-\hat{u}_j) \quad (5.1.13)$$

5.2

Fronteira Estocástica de Custos

Da mesma forma que na fronteira de produção, v_j é irrestrita em sinal e captura choques aleatórios não gerenciáveis e u_j , uma variável não negativa que reflete o grau de ineficiência do produtor. Entretanto, diferentemente do modelo de fronteira estocástica de produção, na fronteira de custo o erro u_j captura o efeito da ineficiência econômica, i.e., reflete as parcelas alocativa e técnica da ineficiência.

A medida de eficiência econômica da j -ésima DMU é a razão entre o mínimo custo teórico especificado pela função fronteira estocástica ($C(y_j, \mathbf{w}_j, \boldsymbol{\beta})$) e o custo efetivamente observado na j -ésima DMU (\mathbf{E}_j):

$$\mathbf{CE}_j = \frac{\mathbf{E}_j}{C(y_j, \mathbf{w}_j, \boldsymbol{\beta}) \exp(v_j)} = \exp(u_j) \quad (5.2.1)$$

onde, $\mathbf{w}_j = [w_{j1} \ \dots \ w_{jm}]^T$ e os demais termos como definidos anteriormente;

$\mathbf{E}_j = \mathbf{x}_j \mathbf{w}_j$ é o custo total incorrido pela j -ésima DMU; $C(y_j, \mathbf{w}_j, \boldsymbol{\beta}) \exp(v_j)$ é o custo estabelecido pela fronteira estocástica.

Arranjando os termos da igualdade mostrada em (5.2.1) encontra-se a seguinte equação:

$$\mathbf{E}_j = C(y_j, \mathbf{w}_j, \boldsymbol{\beta}) \exp(v_j + u_j), \quad j=1, \dots, n \quad (5.2.2)$$

Nesta tese são considerados somente modelos de fronteira com equação única, baseados no custo total, na quantidade de *outputs* e nos preços dos *inputs*. Tais modelos não permitem decompor a medida de eficiência econômica nas componentes alocativa e técnica.

Vale sublinhar que só é possível obter a decomposição da medida de eficiência, por meio de modelos de equações simultâneas, que são mais

sofisticados e necessitam de dados acerca do custo total, quantidades de *outputs*, preços e quantidades dos *inputs* de cada produtor.

Assumindo a hipótese de eficiência alocativa, o erro u_j é relacionado com a ineficiência técnica e, portanto, os modelos de equação única estimam medidas de eficiência técnica, segundo uma orientação ao insumo (*input*). Porém, se esta suposição não é feita, a interpretação de u_j não é clara, pois ele representa uma mistura de ineficiência técnica e alocativa. Por esta razão, a medida de eficiência obtida a partir de uma fronteira de custo é denominada eficiência econômica.

Admitindo que a parcela determinística da fronteira estocástica de custo, $C(y_j, w_j, \beta)$, tenha apenas um *output*, pode-se adotar uma forma funcional, por exemplo Cobb-Douglas, resultando na seguinte equação:

$$\ln E_j = \beta_0 + \beta_j \ln y_j + \beta_1 \ln w_{j1} \dots + \beta_m \ln w_{jm} + v_j + u_j \quad (5.2.3)$$

onde $u_j = -\ln(\text{CE}_j) \geq 0$.

Comparando as equações (5.2.2) e (5.2.3) (à luz da função Cobb-Douglas), pode-se perceber que na fronteira de custo o termo relacionado à ineficiência, u_j , é adicionado à fronteira. A razão para tal deve-se ao fato de que a função custo representa o custo mínimo enquanto a função produção representa o máximo nível de produção. Apesar destas diferenças, a estimação dos parâmetros da equação (5.2.3) e dos índices de eficiência das DMU's são realizadas da mesma forma que na fronteira de produção.

A seguir, na Figura 5.1, são apresentadas três densidades de probabilidade resultantes da combinação $v_j + u_j$, onde $u_j \sim$ Normal Truncada e $v_j \sim$ Normal, geradas considerando $\sigma = \tau = 1$, em um contexto de fronteira de custos.

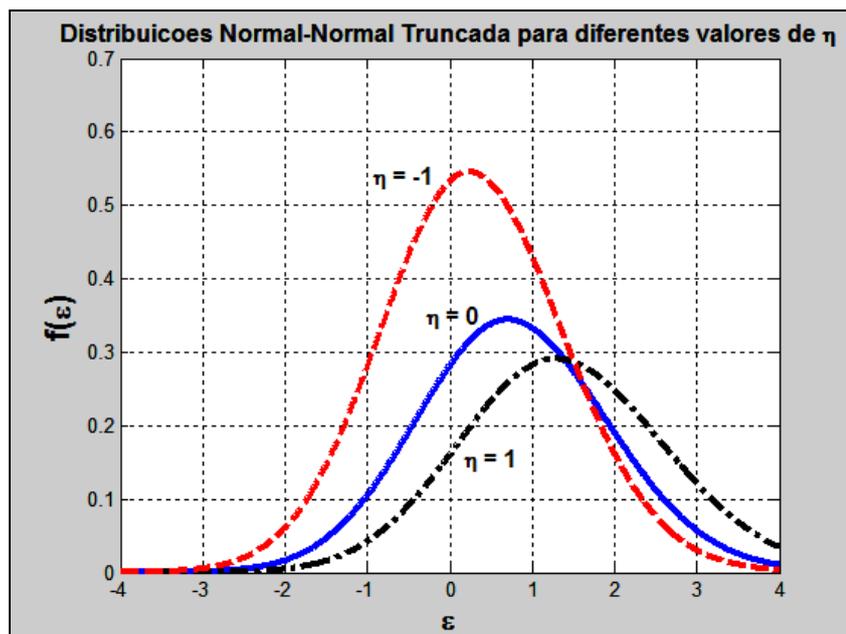


Figura 5.1: Densidades de probabilidade da combinação $v_j + u_j$.

Vale mencionar que a fronteira estocástica é alvo de várias críticas, principalmente, com relação ao termo aleatório u_j . Uma primeira crítica refere-se ao fato de que, geralmente, não há nenhuma justificativa a priori para a seleção de qualquer forma de distribuição deste termo. Outro ponto relevante abordado por Koop et al [55], baseia-se no fato de que u_j é muito sensível à escolha da forma funcional da fronteira. Neste contexto, é importante destacar que há na literatura outras formas funcionais mais flexíveis, conforme apresentado em [55, 62].

De qualquer modo, este trabalho calcula as ineficiências técnicas usando as funções Translog e Cobb-Douglas.

5.3

Fronteira Estocástica sob o paradigma Bayesiano

O modelo de fronteira estocástica sob o paradigma Bayesiano foi proposto no trabalho seminal de van den Broeck, Koop, Osiewalski e Steel – BKOS - [94]. Nesta mesma publicação, os autores citam, com bastante propriedade, duas razões para adotar a técnica Bayesiana em modelos de fronteira estocástica:

- i. a técnica Bayesiana pode ser considerada mais robusta por fornecer resultados mais acurados, mesmo para pequenas amostras;

- ii. os modelos Bayesianos permitem que o especialista incorpore alguma informação que ele detém, a priori, sobre os parâmetros do modelo.

Vale registrar que, na maioria das vezes, os modelos Bayesianos são bastante complexos. Portanto, é inevitável a aplicação de métodos baseados em simulação. Neste cenário, vale lembrar que van den Broeck et al et al [94] empregaram os métodos de Monte Carlo via função de importância. Em seguida, Koop, Osiewalski e Steel – KOS - [56] investigaram a utilização do Amostrador de Gibbs para fazer inferências na distribuição a posteriori. Griffin & Steel [46] comentam que este trabalho pode ser considerado o responsável por diversos outros estudos envolvendo fronteiras estocásticas e algoritmos MCMC. De fato, posteriormente Koot et al [57] apresentaram modelos Bayesianos com dados em painel, Fernández et al [32] examinaram uma questão fundamental em inferência Bayesiana que é a existência da distribuição a posteriori quando a distribuição a priori é imprópria. Em Osiewalski & Steel [69] pode-se encontrar uma descrição clara e objetiva sobre os métodos numéricos. Finalizando, Tsionas [93] propõe uma alternativa para modelos cuja ineficiência se distribui como uma distribuição Gama ($\Gamma(\cdot)$), não sendo necessário que o parâmetro da forma seja inteiro (distribuições da família Erlang). O uso de fronteiras estocásticas com ênfase Bayesiana também pode ser encontrado em, por exemplo, [17,61].

5.3.1

Modelos Bayesianos para Fronteira Estocástica

Conforme já registrado, em sua versão básica, o modelo econométrico com erro composto para estimação de uma fronteira estocástica de custos pode ser expressa matematicamente como:

$$\ln y_j = \beta_0 + \sum_{i=1}^m \beta_i \ln x_{ji} + \sum_{i=1}^m \sum_{k=1}^m \beta_{ik} \ln x_{ji} \ln x_{jk} + v_j + u_j \quad (5.3.1.1)$$

Em um contexto Bayesiano, inicialmente formula-se a distribuição amostral e então faz-se a especificação das distribuições de probabilidade assumidas a priori para os parâmetros do modelo. Neste estudo são considerados

os coeficientes β , da função fronteira, e os parâmetros das distribuições de probabilidade associadas com os termos aleatórios u e v . Também é assumido que o termo aleatório $v_j \sim iid N(0, \sigma^2)$, i.e. distribuição Normal com média 0 e variância constante σ^2 . É relevante mencionar que para u_j , adotou-se três distribuições muito utilizadas na literatura:

- i. $u_j \sim iid \Gamma(\alpha, \lambda^{-1})^7$ - Exponencial com média (desvio-padrão) λ ;
- ii. $u_j \sim iid N^+(0, \lambda^{-1})$ - *Half Normal*;
- iii. $u_j \sim iid N^+(\eta_j, \tau^2)$ - Normal-Truncada.

5.3.2

Varying Efficiency Distribution (VED) models

Em diversas situações, é razoável assumir que as eficiências das DMU's com características similares sejam relacionadas no modelo. Para tal, atribuiu a cada DMU um vetor constituído de variáveis exógenas s_{jl} ($l=2, \dots, k$), em geral representadas por *dummies*. Para estes casos, Osiewalski & Steel [69] propuseram, para a eficiência média, a seguinte parametrização:

$$\lambda_j = \prod_{l=1}^k \phi_l^{-s_{jl}} \quad (5.3.2.1)$$

Vale informar que $\phi_l > 0$ são os parâmetros desconhecidos e $s_{j1} \equiv 1$. Se $k > 1$, as distribuições de u_j podem diferir para diferentes j e então Koop et al [57] denominaram esta especificação como modelo *Varying Efficiency Distribution (VED)*. Se $k = 1$, então $\lambda_j = \phi_1^{-1}$ e todos os termos relacionados às ineficiências constituem amostras independentes da mesma distribuição. Segundo Osiewalski & Steel [69], este é um caso especial chamado *Common Efficiency Distribution (CED)*.

⁷ Para o caso da distribuição exponencial, considere $\alpha = 1$.

Aqui, é interessante apresentar uma outra parametrização, análoga à equação (5.3.2.1), proposta por Koop et al [57]:

$$\lambda_j = \mathbf{exp}(-s_j \gamma) \quad (5.3.2.2)$$

Igualando as expressões (5.3.2.1) e (5.3.2.2), não é difícil concluir que $\phi_l = \mathbf{exp}(\gamma_l)$.

Para elucidar esses conceitos, seja a distribuição conjunta de y_j e $u_j \sim iid \Gamma(1, \lambda^{-1})$ dado \mathbf{x}_j , $s_j = [1 \text{ dummy}_1 \dots \text{dummy}_k]$ e $\theta = [\boldsymbol{\beta}^T \ \sigma^{-2} \ \phi_1 \ \dots \ \phi_k]^T$:

$$p(y_j, u_j | s_j, \mathbf{x}_j, \theta) = f(y_j | h(\mathbf{x}_j; \boldsymbol{\beta}) + u_j, \sigma^2) \Gamma\left(u_j \mid 1, \prod_{l=1}^k \phi_l^{s_{jl}}\right) \quad (5.3.2.3)$$

A distribuição amostral é dada por:

$$p(y_j | s_j, \mathbf{x}_j, \theta) = \int_0^\infty f(y_j | h(\mathbf{x}_j; \boldsymbol{\beta}) + u_j, \sigma^2) \Gamma\left(u_j \mid 1, \prod_{l=1}^k \phi_l^{s_{jl}}\right) du_j \quad (5.3.2.4)$$

Esta distribuição também pode ser representada da seguinte forma:

$$p(y_j | s_j, \mathbf{x}_j, \theta) = \lambda_j^{-1} \cdot \mathbf{exp}\left[-\lambda_j^{-1} \left(m_j + \frac{1}{2} \sigma^2 \lambda_j^{-1}\right)\right] \Phi\left(\frac{m_j}{\sigma}\right) \quad (5.3.2.5)$$

onde:

$$m_j = y_j - h(\mathbf{x}_j; \boldsymbol{\beta}) - \sigma^2 \lambda_j^{-1}$$

A função de verossimilhança, $l(\theta; \text{dados})$, bastante complexa, é o produto das densidades mostradas em (5.3.2.5) em função do indexador j .

A abordagem Bayesiana requer a especificação das distribuições a priori para os parâmetros. A princípio, a distribuição a priori de θ pode ser qualquer, mas é usualmente não recomendável incorporar muita informação subjetiva sobre os mesmos. Dessa forma, tem-se as seguintes distribuições a priori:

$$\boldsymbol{\beta} \sim N^+(0, \sigma_{\boldsymbol{\beta}}^2); \quad (5.3.2.6)$$

$$\sigma^{-2} \sim \Gamma\left(\frac{n_0}{2}, \frac{c_0}{2}\right); \quad (5.3.2.7)$$

$$\phi_l \sim \Gamma(a_l, g_l) \quad (5.3.2.8)$$

$$\exp(-\gamma) \sim \Gamma\left(1, -\ln(r^*)\right)^8 \quad (5.3.2.9)$$

Segundo Fernández et al [32], é fundamental que as prioris de σ sejam informativas ($n_0 > 0$ e $c_0 > 0$) para assegurar a existência da distribuição a posteriori em modelos de fronteira estocástica com amostra *cross-section*. Com relação aos k parâmetros da distribuição da eficiência, Koop et al [57] sugerem utilizar $a_l = g_l = 1$ para $l = 2, \dots, k$, $a_1 = 1$ e $g_1 = -\ln(r^*)$, onde $r^* \in (0,1)$ é o hiperparâmetro a ser determinado. De acordo com van den Broeck et al [94], no modelo CED, r^* pode ser interpretado como a eficiência mediana a priori, pois ele é exatamente a mediana da distribuição marginal do termo relacionado com a eficiência, $r_j = \exp(-u_j)$. Procedendo desta maneira, pode-se assegurar que o modelo VED é consistente com o modelo CED.

Até o momento, foi citado apenas o caso onde $u_j \sim iid \Gamma(1, \lambda^{-1})$ (item (i) retrocitado). Quando $u_j \sim iid N^+(0, \lambda^{-1})$, item (ii), tem-se um modelo parcimonioso. De acordo com van den Broeck et al [94], a distribuição a priori para o parâmetro λ é dado por

$$\lambda \sim \Gamma\left(1, -\ln(r^*)\right) \quad (5.3.2.10)$$

Com relação ao item (iii), Normal-Truncada, este será apresentado oportunamente.

5.4

Inferência Bayesiana usando Amostrador de Gibbs

Conforme já foi reportado, a técnica Bayesiana combina toda a informação sobre as quantidades de interesse em sua distribuição a posteriori. Vale citar que como o objetivo de uma análise envolvendo fronteira estocástica é fazer não somente inferências sobre os parâmetros da fronteira de tecnologia, mas também sobre as eficiências individuais de cada empresa, é interessante expressar a densidade a posteriori da seguinte forma:

⁸ Vide Griffin & Steel [46] para maiores detalhes a respeito desta parametrização.

$$\pi(\mathbf{u}, \theta | \mathbf{y}, \mathbf{X}, \mathbf{S}) = p(\mathbf{u} | \mathbf{y}, \mathbf{X}, \mathbf{S}) p(\theta | \mathbf{y}, \mathbf{X}, \mathbf{S}) \quad (5.4.1)$$

ou

$$\pi(\mathbf{u}, \theta | \mathbf{y}, \mathbf{X}, \mathbf{S}) \propto p(\mathbf{u} | \mathbf{y}, \mathbf{X}, \mathbf{S}) p(\theta) l(\theta; \mathbf{y}, \mathbf{X}, \mathbf{S}) \quad (5.4.2)$$

Agora, ao invés de integrar a expressão (5.4.2) em função de \mathbf{u} (conforme foi feito em (5.3.2.4)), Osiewalski & Steel [69] apresentam uma forma mais fácil, que é considerar θ dado \mathbf{u} e os dados. Pode ser verificado que a inclusão do termo aleatório \mathbf{u} , no Amostrador de Gibbs, faz com que a dimensão do modelo seja $(n+m+k+1)$ e, conseqüentemente, maior que o número de unidades observadas. Apesar desta alta dimensionalidade, os passos necessários para implementar o Amostrador de Gibbs são relativamente fáceis conforme mostrado em Koop et al [56].

Neste ponto, é digno descrever as distribuições condicionais completas:

$$\pi(\sigma^{-2} | \mathbf{y}, \mathbf{X}, \mathbf{S}, \mathbf{u}, \boldsymbol{\beta}, \phi) = \pi(\sigma^{-2} | \mathbf{y}, \mathbf{X}, \mathbf{u}, \boldsymbol{\beta}) = \Gamma \left(\frac{n+n_0}{2}, \frac{c_0 + \sum_j (y_j - h(\mathbf{x}_j; \boldsymbol{\beta}) - u_j)^2}{2} \right) \quad (5.4.3)$$

$$\pi(\boldsymbol{\beta} | \mathbf{y}, \mathbf{X}, \mathbf{S}, \mathbf{u}, \sigma^{-2}, \phi) = \pi(\boldsymbol{\beta} | \mathbf{y}, \mathbf{X}, \mathbf{u}, \sigma^{-2}) \propto N^+ \left(\boldsymbol{\beta} | 0, \sigma_{\boldsymbol{\beta}}^2 \right) \times \exp \left(-\frac{1}{2} \sigma^{-2} \sum_j (y_j - h(\mathbf{x}_j; \boldsymbol{\beta}) - u_j)^2 \right) \quad (5.4.4)$$

A distribuição condicional completa de ϕ_l ($l = 1, \dots, k$) apresenta a seguinte forma geral:

$$\begin{aligned} \pi(\phi_l | y, \mathbf{X}, \mathbf{S}, \mathbf{u}, \boldsymbol{\beta}, \sigma^{-2}, \phi_{(-l)}) &= \pi(\phi_l | \mathbf{S}, \phi_{(-l)}) \propto \\ \exp\left(-\phi_l \sum_j u_j D_{jl}\right) &\times \Gamma\left(\phi_l \left| 1 + \sum_j s_{jl}, g_l \right.\right) \end{aligned} \quad (5.4.5)$$

onde:

$$D_{jl} = \prod_{j \neq l}^k \phi_j^{s_{jl}} \quad (5.4.6)$$

para $l=1, \dots, k$ ($D_{j1} = 1$ para $k=1$) e $\phi_{(-l)}$ denota ϕ sem o seu l -ésimo elemento.

Por fim, o vetor relacionado às ineficiências não-observadas condicionado aos parâmetros e dados tem uma densidade Normal-Truncada, ou seja:

$$\pi(\mathbf{u} | y, \mathbf{X}, \mathbf{S}, \theta) = \prod_{j=1}^n \frac{f_N(u_j | \eta^*, \sigma_*^2)}{\Phi\left(\frac{\eta^*}{\sqrt{\sigma_*^2}}\right)} \quad (5.4.7)$$

sendo:

$$\eta^* = \frac{(\eta\sigma^2 + y_j\tau^2 - h(\mathbf{x}_j; \boldsymbol{\beta})\tau^2)}{(\sigma^2 + \tau^2)} \quad (5.4.8)$$

e

$$\sigma_*^2 = \frac{(\sigma^2 + \tau^2)}{\sigma^2\tau^2} \quad (5.4.9)$$

De posse desses resultados, pode-se verificar que as amostras de u_j são geradas a partir de (5.4.7) e imediatamente convertidas em indicadores de eficiência definida como $r_j = \exp(-u_j)$.

6

Clustering: Definindo Grupos Estratégicos

6.1

Introdução

Não é difícil deduzir que as 60 empresas distribuidoras de energia elétrica, operantes no SEB, apresentam características heterogêneas. Entretanto, é indispensável que o pesquisador garanta, antes de aplicar os métodos da empresa de referência, DEA e SFA, que as distribuidoras analisadas sejam comparáveis [83].

Para atender a este requisito, os dados das 60 concessionárias foram processados por uma Rede Neural auto-organizável. É interessante registrar que a razão da escolha da Rede Neural de Kohonen deve-se, principalmente, aos seguintes fatores:

- a) Obtenção de um mapa topologicamente ordenado das distribuidoras, o que facilita a visualização e a interpretação dos agrupamentos;
- b) Usualmente, é mais fácil interpretar a operação das redes não-supervisionadas através da análise dos pesos.

À luz deste cenário, é importante mencionar que as variáveis (atributos) escolhidas para o desenvolvimento deste estudo estão em concordância com o trabalho de Zanini [97]. São elas:

- i. Custo operacional (OPEX);
- ii. Quantidade de energia distribuída (MWh);
- iii. Total de unidades consumidoras atendidas;
- iv. Extensão da rede de distribuição (Km).

6.2

Análise de dados

A análise de dados é requerida em praticamente todas as aplicações computacionais.

Martin-del-Brio & Serrano-Cinca [59] mencionam que o pré-processamento dos dados de entrada é um importante aspecto que tem demonstrado apresentar bastante influência tanto no desenvolvimento quanto na configuração do mapa de Kohonen. Dessa forma, a escolha de alguma estratégia para tratamento dos dados está diretamente relacionada com o objetivo que se queira alcançar e, também, com a forma como os dados vão estar dispostos.

Quando as variáveis de entrada têm dimensões muito diferentes, apresentar os dados em uma escala linear é uma maneira de inicializar cada variável por igual, ou seja, todas têm o mesmo grau de importância. Desta forma, cada componente do vetor de entradas é colocado em uma escala de mesma dimensão, por exemplo $[-1,1]$.

Ainda de acordo com Martin-del-Brio & Serrano-Cinca [59], é interessante normalizar o vetor de dados de maneira que ele apresente comprimento unitário. Este procedimento, bastante comum, é muito utilizado em mapeamento auto-organizável quando a medida usada é a correlação. No entanto, quando é utilizada a medida de distância euclidiana, este procedimento não é estritamente recomendado, exceto se a norma dos vetores for muito alta.

6.3

Resultados obtidos via *cluster analysis*

Os quatro (4) atributos (padronizados) foram processados em uma Rede Neural de Kohonen cujo mapeamento usado consiste de dezesseis (16) unidades arranjadas em uma grade quadrada 4×4 e a medida de similaridade utilizada foi a distância euclidiana. Como o número de *clusters* não é previamente conhecido, tipicamente este número é sobre-estimado, usando-se mais unidades do que o necessário [67].

Inicialmente, atribuiu-se aos pesos sinápticos pequenos valores aleatórios no intervalo $[-1,1]$. Posteriormente, a rede foi treinada por 2.500 iterações sendo que, em cada iteração, os 60 padrões são apresentados à rede em seqüência

aleatoriamente definida para a iteração em vigor. A função de decaimento exponencial foi usada para atualizar o raio da distância da unidade vencedora às suas unidades vizinhas [82]:

$$\text{dist} = \text{round} \left(0,35 + \text{dist0} \cdot \exp \left(- \left(\frac{\text{dec} \cdot \text{iter}}{\text{Niter}} \right) \right) \right) \quad (6.3.1)$$

onde:

- round = função que executa o arredondamento para o número inteiro mais próximo;
- dist0 = 3, distância inicial;
- exp = função exponencial;
- dec = constante de decaimento;
- iter = iteração atual;
- Niter = número total de iterações, 2.500.

A taxa de aprendizado η é atualizada em cada iteração da seguinte forma:

$$\eta = \eta_0 \cdot \exp \left(- \left(\frac{\text{dec} \cdot \text{iter}}{\text{Niter}} \right) \right) \quad (6.3.2)$$

onde:

η = valor da taxa de aprendizado em cada iteração;

η_0 = valor inicial da taxa de aprendizado: 0,0058.

A Figura 6.1 mostra os padrões similares descobertos pela Rede Neural de Kohonen:

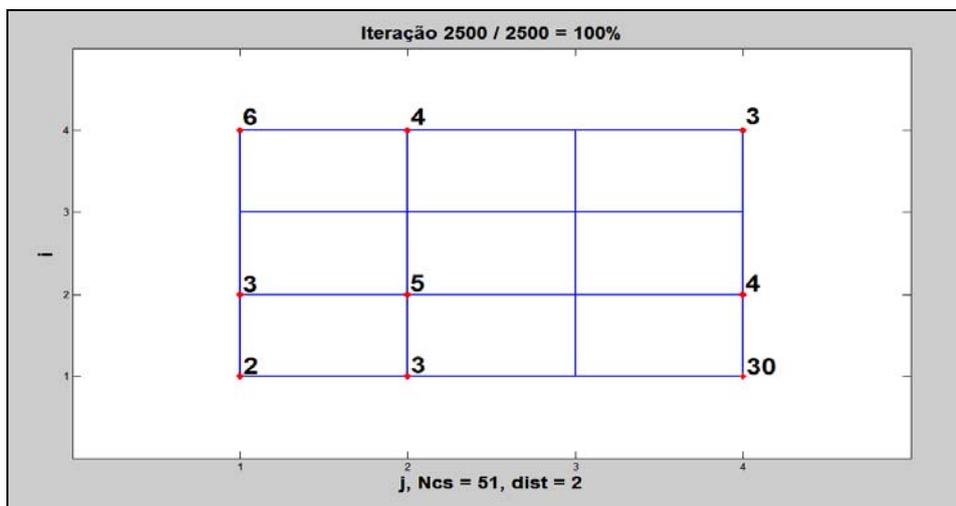


Figura 6.1: Resultado de um processo de *clustering* em uma simulação.

Os resultados de cada *cluster*, mostrados na Figura 6.1, são decodificados na Tabela 6.2:

Tabela 6.2 – *Clusters* obtidos pela Rede de Kohonen em uma simulação.

<i>Cluster</i>	Concessionárias
C1	[1 9 17 41 53]
C2	[2 15 22 28]
C3	[3 4 5 6 7 11 16 20 24 25 26 31 33 34 36 37 42 43 46 47 48 49 50 52 55 56 57 58 59 60]
C4	[8 30 45 54]
C5	[10 14 51]
C6	[12 19 27 32 38 44]
C7	[13 29 35]
C8	[18 40]
C9	[21 23 39]

Neste ponto, é pertinente apresentar os resultados obtidos através da Rede Neural de Kohonen via simulação de Monte Carlo (MCRNK). Para tal, realizou-se 5.000 simulações resultando em 48 *clusters* distintos. A Figura 6.2 mostra o histograma dos *clusters* encontrados:

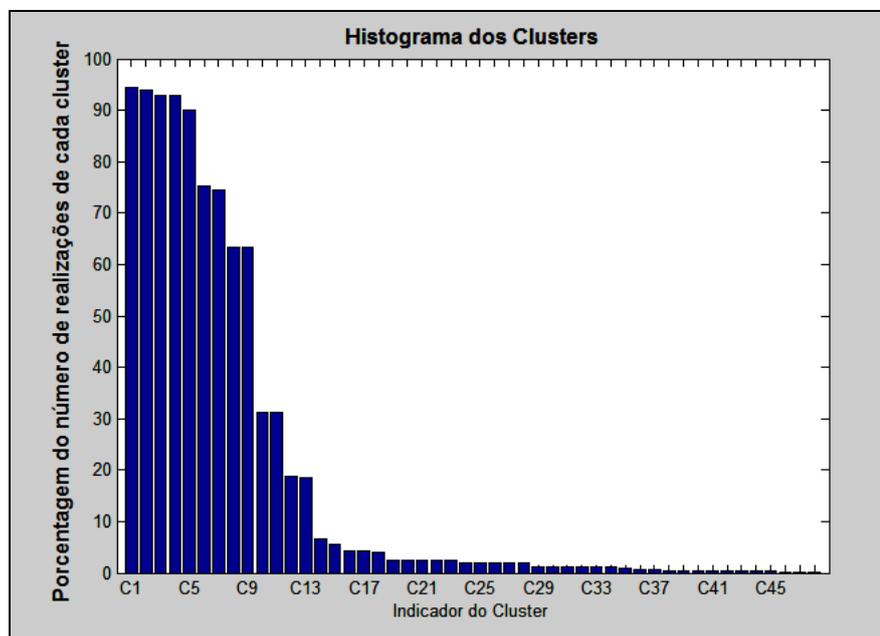


Figura 6.2: Histograma dos *clusters*.

Com base em uma análise pictórica do histograma (Figura 6.2) buscou-se elaborar a Tabela 6.3 cujo objetivo é apresentar a frequência e também a composição de cada *cluster*.

Tabela 6.3 – Sequência ordenada dos *clusters* mais freqüentes.

<i>Cluster</i>	Porcentagem	Concessionárias
C1	94,48 %	[13 15 22 28 35]
C2	94,04 %	[30 54]
C3	93,00 %	[12 19 27 32 38 44]
C4	92,96 %	[9 17 18 40]
C5	89,96 %	[2 29]
C6	75,38 %	[10 14 53]
C7	74,36 %	[1 41 51]
C8	63,38 %	[8 16 21 23 39 45]
C9	63,38 %	[3 4 5 6 7 11 20 24 25 26 31 33 34 36 37 42 43 46 47 48 49 50 52 55 56 57 58 59 60]
C10	31,10 %	[6 8 16 21 23 39 45]
C11	31,10 %	[3 4 5 7 11 20 24 25 26 31 33 34 36 37 42 43 46 47 48 49 50 52 55 56 57 58 59 60]
C12	18,66 %	[10 53]
C13	18,60 %	[1 14 41 51]
C14	6,56 %	[12 19 32 38 44]
C15	5,52 %	[3 4 5 6 7 11 16 20 24 25 26 31 33 34 36 37 42 43 46 47 48 49 50 52 55 56 57 58 59 60]
C16	4,40 %	[23 39 45]
C17	4,40 %	[8 21 30 40 54]
C18	4,08 %	[2 27 29]
C19	2,48 %	[2 35]
C20	2,48 %	[10 15 28]
C21	2,48 %	[1 41 51 53]
C22	2,48 %	[9 14 17 18]
C23	2,48 %	[13 22 27 29]
C24	1,92 %	[2]
C25	1,92 %	[18 28]
C26	1,92 %	[1 17 41 53]
C27	1,92 %	[9 10 14 51]
C28	1,92 %	[13 15 22 29 35]
C29	1,12 %	[10 14 51]
C30	1,12 %	[13 29 35]
C31	1,12 %	[2 15 22 28]
C32	1,10 %	[18 40]
C33	1,10 %	[1 9 17 41 53]
C34	1,08 %	[9 18 40]
C35	1,02 %	[1 17 41 51]
C36	0,72 %	[21 23]
C37	0,72 %	[8 30 39 45 54]
C38	0,44 %	[10]
C39	0,44 %	[27 29]
C40	0,44 %	[9 14 18]
C41	0,44 %	[30 40 54]
C42	0,44 %	[1 17 41 51 53]
C43	0,44 %	[2 12 19 32 38 44]
C44	0,40 %	[21 23 29]
C45	0,40 %	[8 30 45 54]
C46	0,06 %	[1 14 17 41 51]
C47	0,02 %	[18 40 53]
C48	0,02 %	[1 9 17 41]

O entendimento da Tabela 6.3 é direto. Como exemplo, seja a concessionária número 6; em 63,38% das realizações (simulações), ela pertence ao *cluster* 9 (C9). No entanto, ela também pode, com 31,10 %, pertencer ao *cluster* 10 (C10).

Para uma melhor inspeção visual, o histograma é truncado nos 13 *clusters* mais freqüentes. A Figura 6.3 mostra este processo:

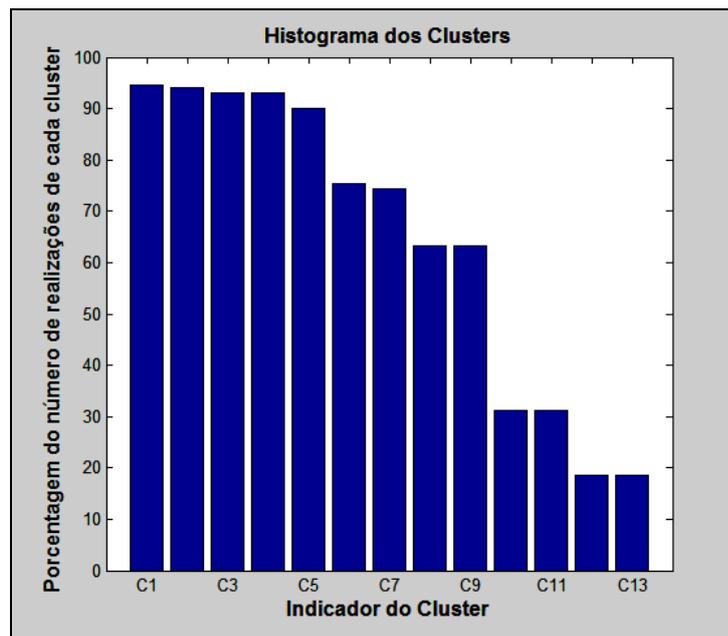


Figura 6.3: Histograma dos 13 *clusters* mais freqüentes.

Pela Tabela 6.3, percebe-se que, com exceção do *cluster* 9 (que contém um número expressivo de empresas), há vários *clusters* constituídos de poucas empresas. Baseado na análise dos pesos sinápticos, foi possível obter três (3) agrupamentos. Com este procedimento, pode-se verificar que cada *cluster* passa a ser composto de um número significativo de empresas com características similares, conforme informado na Tabela 6.4:

Tabela 6.4 – Agrupamentos finais.

<i>Cluster</i>	Concessionárias
C1	[1 8 10 14 16 21 23 30 39 41 45 51 53 54]
C2	[2 9 12 13 15 17 18 19 22 27 28 29 32 35 38 40 44]
C3	[3 4 5 6 7 11 20 24 25 26 31 33 34 36 37 42 43 46 47 48 49 50 52 55 56 57 58 59 60]

Vale informar que estes agrupamentos representam uma importante etapa na obtenção dos próximos resultados.

7

Mensuração da Eficiência Relativa Utilizando Modelos DEA

7.1

Introdução

O propósito deste Capítulo é calcular os *scores* de eficiência através de modelos DEA. Antes, porém, é importante determinar as variáveis consideradas na análise, com a definição dos *inputs* e *outputs*. Nesse contexto, a Tabela 7.1 cumpre este papel:

Tabela 7.1 – Variáveis consideradas.

Tipo da Variável	Variável	Descrição
<i>Input</i>	OPEX	Custo operacional (R\$ 1.000);
<i>outputs</i>	MWh	quantidade de energia distribuída (MWh);
	NC	total de unidades consumidoras atendidas
	KM	extensão da rede de distribuição (Km)

Vale frisar que os *outputs* são os *cost drivers* dos custos operacionais. O montante de energia distribuída (MWh) é uma *proxy* do produto total, o número de unidades consumidoras (NC) é uma *proxy* da quantidade de serviço e a extensão da rede de distribuição (KM) reflete a dispersão dos consumidores na área de concessão, um fator determinante dos custos operacionais.

O próximo tópico publica e discute os resultados obtidos através dos modelos DEA.

7.2

Resultados DEA e discussões

As medidas de eficiência obtidas pelos modelos DEA devem expressar o potencial de redução dos custos operacionais, i.e., trata-se de uma medida orientada ao insumo. Denotando as medidas de eficiência por θ , o potencial de redução dos custos operacionais para a j -ésima distribuidora, ou seja, o custo

operacional reconhecido pelo regulador seria igual a $OPEX_j(1 - \theta_j)$. Assim, o modelo DEA é orientado ao *input*, sendo que o OPEX é o único atributo com característica de insumo conforme indicado na Tabela 7.1.

Com relação ao rendimento de escala do modelo DEA foram avaliadas duas opções: rendimentos constantes (RCE) e rendimentos variáveis (RVE). O modelo DEA/RCE foi aplicado separadamente em cada agrupamento de DMU's, pois desta forma pode-se garantir que as DMUs avaliadas sejam comparáveis, i.e., homogêneas. Já no modelo DEA/RVE, a restrição de convexidade, vide **(M3.4)**, garante que as DMU's ineficientes sejam comparadas somente com as DMU's com mesmas características ou nível de atividade [21]. Desta forma, o modelo DEA/RVE pode ser aplicado ao conjunto formado por todas as distribuidoras sem a necessidade de classificá-las previamente. No entanto, a aplicação do modelo RVE ao conjunto de todas as distribuidoras não evita que as DMU's ineficientes, em um determinado agrupamento, sejam comparadas com as distribuidoras dos demais agrupamentos.

Em função da hipótese de convexidade, a fronteira de eficiência estimada pelo modelo RVE é sempre caracterizada por rendimentos crescentes de escala nos menores níveis de atividade e por rendimentos decrescentes de escala nos maiores níveis de atividade. Ressalta-se que esta propriedade do modelo RVE independe dos dados contidos na amostra de DMU's. Em resumo, na fronteira gerada pelo modelo RVE os rendimentos não são crescentes ao longo de toda fronteira e a produtividade diminui à medida que o nível de atividade aumenta para além da escala ótima. Esta característica da fronteira RVE não atende aos objetivos da regulação econômica, em particular o estímulo aos ganhos de produtividade. Para fins de um esquema de regulação por incentivo, a alternativa ao modelo VRS consiste em estimar a fronteira pelo modelo NDRS (*Non-Decreasing Return to Scale*) uma variante do modelo VRS que não admite rendimentos decrescentes de escala [99].

Por fim, vale citar que as medidas de eficiência obtidas através do modelo DEA clássico (sem restrições aos pesos) foram calculados com o auxílio dos *software* DEAP (www.uq.edu.au/economics/cepa/software.htm); para o modelo *Adjusted Contingent Restrictions on Weight*, utilizou-se o LINGO 9

(www.lindo.com); para o modelo NDRS usou-se DEA Excel Solver [99]. Para simplificar, adotou-se as seguintes nomenclaturas:

- **R1** - Resultados do modelo DEA/RCE Sem Restrições;
- **R2** - Resultados do modelo DEA/RCE Com Restrições;
- **R3** - Resultados do modelo DEA/NDRS.

Os mesmos são mostrados na Tabela 7.2:

Tabela 7.2 – Índices de eficiência θ_j^{DEA} .

Concessionária	R1	R2	R3
CLUSTER 1			
AES-SUL	1,0000	1,0000	1,0000
CEAL	0,6650	0,3450	0,6048
CEEE	0,2840	0,2000	0,2731
CELPA	0,3940	0,2080	0,3625
CELTINS	0,5090	0,5080	0,3792
CEPISA	0,7210	0,4800	0,6587
CERON	0,5450	0,5450	0,4328
COSERN	0,8980	0,5630	0,8336
ENERGIPE	0,7440	0,4510	0,7007
ESCELSA	0,6830	0,6830	0,6806
MANAUS	0,3820	0,3800	0,3826
PIRATININGA	0,9130	0,9130	0,9134
RGE	1,0000	1,0000	0,9974
SAELPA	0,9700	0,5060	0,8818
CLUSTER 2			
BANDEIRANTES	1,0000	0,9440	0,7987
CEB	0,3390	0,3130	0,2877
CELESC	0,5920	0,5700	0,4387
CELG	0,5180	0,5180	0,5072
CELPE	1,0000	0,9000	1,0000
CEMAR	0,6750	0,5270	0,6754
CEMAT	0,4930	0,4930	0,4581
CEMIG	0,8110	0,8110	0,6540
CERJ	0,7690	0,6790	0,7439
COELBA	0,7350	0,6400	0,7349
COELCE	0,7950	0,7170	0,7952
COPEL	1,0000	1,0000	1,0000
CPFL	1,0000	0,9370	0,8095
ELEKTRO	1,0000	1,0000	0,8344
ELETRIPAULO	0,7420	0,7080	0,6149
ENERSUL	1,0000	1,0000	1,0000
LIGHT	0,7690	0,6980	0,6310
CLUSTER 3			
BOA VISTA	0,2970	0,2740	0,1955
BRAGANTINA	0,7590	0,6960	0,4400
CAUIÁ	0,6260	0,6080	0,4527
CAT-LEO	0,8310	0,8310	0,6165
CEA	0,4500	0,4280	0,3200
CELB	0,9420	0,9100	0,7157
CENF	0,6630	0,6310	0,5162
CFLO	0,8030	0,7420	0,5429
CHESP	1,0000	1,0000	0,8968
COCEL	0,8510	0,7950	0,5433
CPEE	0,8440	0,7990	0,5473
CSPE	1,0000	0,9510	0,6451
DEMEI	0,8170	0,7680	0,6680
ELETRACRE	0,7410	0,6890	0,5783
ELETRICAR	0,7240	0,7000	0,5260
JAGUARI	1,0000	0,8290	0,6280
JOÃO CESA	0,8880	0,7500	1,0000
MOCOCA	0,7530	0,7140	0,5262
MUXFELDT	1,0000	0,9640	1,0000
NACIONAL	0,8350	0,7820	0,5998
NOVA PALMA	1,0000	1,0000	0,8885
PANAMBI	0,5690	0,5310	0,4298
POÇOS DE CALDAS	1,0000	0,9270	0,6843
SANTA CRUZ	0,7280	0,7050	0,4883
SANTA MARIA	0,8180	0,7900	0,5958
SULGIPE	1,0000	1,0000	0,8310
URUSSANGA	0,4600	0,3820	0,3821
V. PARANAPANEMA	0,5700	0,5540	0,4032
XANXERÊ	0,5340	0,5000	0,3523

Muitas vezes, é pertinente analisar uma série de dados através de medidas de posição central (média e/ou mediana). Não obstante, é também interessante investigar a variabilidade dos dados. A Tabela 7.3 fornece os resultados obtidos:

Tabela 7.3 – Estatísticas descritivas dos modelos DEA.

	R1	R2	R3
Média	0,7575	0,6914	0,6350
Mediana	0,7690	0,7025	0,6223
Desvio-padrão	0,2085	0,2224	0,2147

Interpretando os 3 modelos, **R1**, **R2** e **R3** na Tabela 7.3, com relação à média e mediana, verifica-se que não há valores discrepantes. Nesta mesma análise, pode ser corroborado que **R2** apresenta valores intermediários a **R1** e **R3**. Com relação ao desvio-padrão, conclui-se que **R2** discrimina melhor as eficiências das empresas (maior variabilidade).

Ocorre, porém, que com apenas os 3 valores mencionados, não é possível ter uma idéia da assimetria da distribuição dos dados. Para tal, foi realizado um *scatterplot matrix*, conforme mostrado na Figura 7.1:

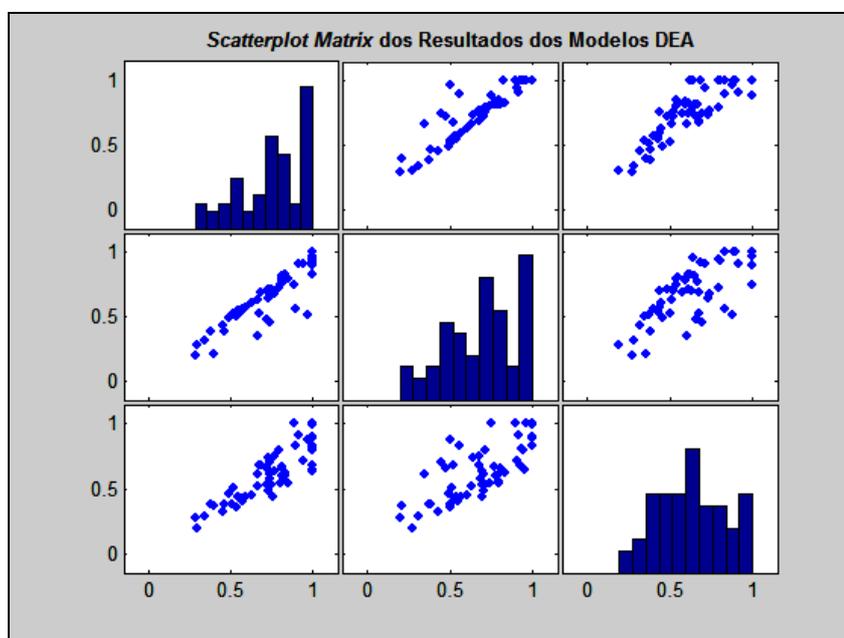


Figura 7.1: *Scatterplot Matrix* dos índices de eficiência (Modelos DEA).

Analisando a Figura 7.1, salta aos olhos a assimetria das distribuições relacionadas com os modelos **R1** e **R2**. Vale lembrar que estes modelos consideram as empresas em seus respectivos *clusters*. A Figura 7.1 também revela uma alta correlação entre os modelos **R1** e **R2**, moderada entre os modelos **R1** e **R3** e baixa para os modelos **R2** e **R3**.

Finalizando, será que estes índices de eficiência mensurados são diferentes para cada um dos modelos? Caso seja verdadeira esta hipótese, ela é significativa?

Para responder a essas indagações, foi efetuado um Teste de Kruskal-Wallis (também chamado teste H). Segundo Triola [92], este teste de hipótese, não-paramétrico, usa postos de dados amostrais de três ou mais populações independentes. É usado para testar a Hipótese nula de que as amostras independentes provêm de populações com a mesma mediana; a Hipótese alternativa é a afirmativa de que as populações têm medianas que não são todas iguais.

- *H0*: As amostras provêm de populações com a mesma mediana;
- *H1*: As amostras provêm de populações com medianas que não são todas iguais.

Ao aplicar o teste, conclui-se que a Hipótese nula (*H0*) não deve ser aceita, isto é, os modelos têm características diferentes e, portanto, os parâmetros estimados não são provenientes de uma mesma população.

O objetivo desta investigação é alertar ao especialista para o fato de que embora os 3 modelos sejam variantes de uma mesma metodologia, é legítimo que se faça uma reflexão mais elaborada antes de tomar qualquer decisão.

8

Estimação da Eficiência Técnica Usando Fronteiras Estocásticas

8.1

Introdução

A estimação da eficiência técnica relativa foi apresentada no Capítulo 7 por meio de três modelos DEA *input-oriented*. Neste Capítulo, os índices de eficiência são calculados à luz dos modelos SFA. É digno de registro que todas estas análises têm como propósito mitigar a assimetria da informação e aprimorar a habilidade do agente regulador em comparar os desempenhos das distribuidoras.

Nesse contexto, a organização deste Capítulo é como se segue:

- (i) tópico 8.2 apresenta os modelos SFA com enfoque Clássico;
- (ii) tópico 8.3 apresenta os modelos SFA com enfoque Bayesiano.

8.2

Resultados dos modelos econométricos com enfoque clássico

Foram testadas quatro especificações distintas para a fronteira estocástica de custos, resultantes das combinações de duas formas funcionais (Cobb-Douglas e Translog) com duas distribuições para o termo aleatório u_j (*Half-Normal* e *Normal truncada*):

Cobb-Douglas:

$$\ln \text{OPEX}_j = \beta_0 + \beta_1 \ln \text{MWh}_j + \beta_2 \ln \text{NC}_j + \beta_3 \ln \text{KM}_j + v_j + u_j$$

(8.2.1)

Translog:

$$\begin{aligned} \ln \text{OPEX}_j &= \beta_0 + \beta_1 \ln \text{MWh}_j + \beta_2 \ln \text{NC}_j + \beta_3 \ln \text{KM}_j + \\ &\frac{1}{2} \cdot (\beta_{11} (\ln \text{MWh}_j)^2 + \beta_{22} (\ln \text{NC}_j)^2 + \beta_{33} (\ln \text{KM}_j)^2) + \\ &\frac{1}{2} \cdot (\beta_{12} \ln \text{MWh}_j \ln \text{NC}_j + \beta_{13} \ln \text{MWh}_j \ln \text{KM}_j + \beta_{23} \ln \text{NC}_j \ln \text{KM}_j) + v_j + u_j \end{aligned}$$

(8.2.2)

Com relação à distribuição Normal-Truncada, $u_j \sim iid N^+(\eta_j, \tau^2)$, considerou-se a média definida como sendo $\eta_j = \delta_0 + \delta_1 \ln(\text{OPEX}_j / \text{MWh}_j)$, onde a razão OPEX/MWh é um índice tradicionalmente considerado na avaliação da produtividade das concessionárias de distribuição de energia elétrica.

Nas distintas versões para o modelo SFA, estimadas com o FRONTIER 4.1 (www.uq.edu.au/economics/cepa/software.htm), foram adicionadas duas variáveis *dummies* na forma aditiva com o objetivo de informar em que *cluster* se encontra cada distribuidora.

De posse destas informações, obteve-se os seguintes resultados:

- **R4** - Resultados do modelo *Half-Normal/Cobb-Douglas* considerando os *clusters*;
- **R5** - Resultados do modelo *Half-Normal/Cobb-Douglas* desconsiderando os *clusters*;
- **R6** - Resultados do modelo *Half-Normal/Translog* considerando os *clusters*;
- **R7** - Resultados do modelo *Half-Normal/Translog* desconsiderando os *clusters*;
- **R8** - Resultados do modelo Normal-Truncada/Cobb-Douglas considerando os *clusters*;
- **R9** - Resultados do modelo Normal-Truncada/Cobb-Douglas desconsiderando os *clusters*;
- **R10** - Resultados do modelo Normal-Truncada/Translog considerando os *clusters*;
- **R11** - Resultados do modelo Normal-Truncada/Translog desconsiderando os *clusters*;

Vale observar que em todos os modelos SFA editados nas Tabelas 8.1 e 8.2, a estatística γ é significativa indicando a presença da ineficiência técnica (teste da razão da verossimilhança – LR *test*). Outro comentário que merece ser enfatizado diz respeito ao fato dos coeficientes δ_0 e δ_1 (modelos SFA com Normal-Truncada), serem significativos.

Tabela 8.1 – Índices de eficiência $\theta_j^{\text{SFA-C}/\text{HN}}$.

	Concessionária	R4	R5	R6	R7
CLUSTER 1	AES-SUL	0,9304	0,9216	0,9806	0,9999
	CEAL	0,5816	0,6902	0,6519	0,7181
	CEEE	0,3103	0,3864	0,2504	0,2709
	CELPA	0,3713	0,4562	0,3454	0,4130
	CELTINS	0,3803	0,4872	0,4908	0,4491
	CEPISA	0,5584	0,6696	0,6690	0,7327
	CERON	0,4548	0,5680	0,5471	0,5232
	COSERN	0,8084	0,8523	0,8858	0,9364
	ENERGIPE	0,7380	0,8167	0,8794	0,8690
	ESCELSA	0,7684	0,8273	0,6694	0,6964
	MANAUS	0,4676	0,5704	0,6902	0,6764
	PIRATININGA	0,8872	0,8900	0,9998	0,9988
	RGE	0,9176	0,9129	0,9448	0,9999
	SAELPA	0,7357	0,8085	0,8226	0,9999
CLUSTER 2	BANDEIRANTES	0,6977	0,7075	0,7735	0,4524
	CEB	0,3814	0,4245	0,4725	0,3498
	CELESC	0,5776	0,6044	0,4481	0,3611
	CELG	0,4971	0,5339	0,4773	0,4245
	CELPE	0,8967	0,8652	0,9845	0,9982
	CEMAR	0,6310	0,6626	0,8028	0,7383
	CEMAT	0,4606	0,5105	0,5196	0,4657
	CEMIG	0,7187	0,7134	0,3367	0,3375
	CERJ	0,7756	0,7684	0,7875	0,8285
	COELBA	0,7321	0,7294	0,6619	0,7881
	COELCE	0,7884	0,7818	0,8288	0,7750
	COPEL	0,8888	0,8585	0,8896	0,8391
	CPFL	0,8811	0,8469	0,7378	0,6427
	ELEKTRO	0,8968	0,8679	0,8481	0,7118
ELETROPAULO	0,6888	0,6766	0,5172	0,7652	
ENERSUL	0,7150	0,7480	0,8849	0,9358	
LIGHT	0,7292	0,7185	0,6189	0,6955	
CLUSTER 3	BOA VISTA	0,3113	0,3457	0,3177	0,2957
	BRAGANTINA	0,6884	0,6973	0,7064	0,6931
	CAUIÁ	0,6257	0,6344	0,5987	0,5920
	CAT-LEO	0,7402	0,7275	0,7083	0,7394
	CEA	0,4653	0,4907	0,4619	0,4492
	CELB	0,8740	0,8407	0,9693	0,9815
	CENF	0,6868	0,6968	0,6946	0,7034
	CFLO	0,7838	0,7812	0,8801	0,8234
	CHESP	0,8065	0,8126	0,8696	0,8047
	COCEL	0,8047	0,8040	0,8663	0,8649
	CPEE	0,7959	0,7932	0,8486	0,8582
	CSPE	0,8758	0,8512	0,9965	0,9999
	DEMEI	0,8333	0,8222	0,8571	0,9993
	ELETROACRE	0,7131	0,7127	0,7035	0,7237
	ELETROCAR	0,7281	0,7491	0,7420	0,7491
	JAGUARI	0,6997	0,7221	0,9627	0,9759
	JOÃO CESA	0,7810	0,8112	0,8460	0,8436
	MOCOCA	0,7628	0,7701	0,7980	0,7367
	MUXFELDT	0,9279	0,9082	0,8621	0,8215
NACIONAL	0,8066	0,7885	0,9052	0,9574	
NOVA PALMA	0,9128	0,8934	0,9949	0,9662	
PANAMBI	0,6148	0,6620	0,5782	0,5312	
POÇOS DE CALDAS	0,6925	0,7157	0,7220	0,9856	
SANTA CRUZ	0,6959	0,6986	0,6735	0,6803	
SANTA MARIA	0,7268	0,7394	0,7596	0,8028	
SULGIPE	0,8555	0,8341	0,9758	0,8561	
URUSSANGA	0,3779	0,4396	0,4152	0,5478	
V. PARANAPANEMA	0,5731	0,5914	0,5556	0,5420	
XANXERÉ	0,5340	0,5791	0,5374	0,5575	

Tabela 8.2 – Índices de eficiência $\theta_j^{\text{SFA-C/NT}}$.

Concessionária	R8	R9	R10	R11
CLUSTER 1				
AES-SUL	0,9109	1,0000	0,9054	0,9582
CEAL	0,3880	0,5486	0,6225	0,2540
CEEE	0,2037	0,3840	0,2469	0,1826
CELPA	0,2362	0,3840	0,3420	0,1802
CELTINS	0,2605	0,3153	0,4415	0,1442
CEPISA	0,3576	0,4283	0,5628	0,1963
CERON	0,3201	0,4379	0,5215	0,2023
COSERN	0,5819	0,9003	0,8680	0,4250
ENERGIPE	0,5391	0,8608	0,8853	0,4018
ESCELSA	0,5863	0,9782	0,6428	0,5605
MANAUS	0,3453	0,7678	0,6827	0,3638
PIRATININGA	0,7330	1,0000	0,9074	0,8752
RGE	0,8223	0,9907	0,9182	0,7550
SAELPA	0,4769	0,6174	0,7242	0,2849
CLUSTER 2				
BANDEIRANTES	0,7508	0,9904	0,5487	0,7652
CEB	0,3591	0,4416	0,5005	0,2086
CELESC	0,5638	0,8137	0,4419	0,3926
CELG	0,4501	0,4598	0,4858	0,2172
CELPE	0,8863	0,9117	0,9995	0,4364
CEMAR	0,5360	0,4435	0,7258	0,2049
CEMAT	0,4598	0,4950	0,5465	0,2333
CEMIG	0,6773	0,9477	0,3408	0,4871
CERJ	0,6909	0,7745	0,8316	0,3658
COELBA	0,6194	0,6717	0,6847	0,3175
COELCE	0,7037	0,6936	0,8201	0,3256
COPEL	0,8734	0,8661	0,8726	0,4102
CPFL	0,9121	0,9869	0,7398	0,6627
ELEKTRO	0,9999	0,9848	0,8688	0,6227
ELETROPAULO	0,5898	0,9387	0,5471	0,4749
ENERSUL	0,7465	0,7476	0,9417	0,3510
LIGHT	0,6550	0,9577	0,6424	0,4999
CLUSTER 3				
BOA VISTA	0,2315	0,2926	0,3328	0,1340
BRAGANTINA	0,5571	0,7695	0,6949	0,3564
CAUIÁ	0,4547	0,5472	0,5840	0,2520
CAT-LEO	0,5302	0,5762	0,6439	0,2652
CEA	0,3380	0,4135	0,4692	0,1898
CELB	0,7025	0,8221	0,9757	0,3762
CENF	0,5056	0,5545	0,7089	0,2516
CFLO	0,6433	0,7930	0,9451	0,3610
CHESP	0,6616	0,5135	0,7335	0,2273
COCEL	0,7102	0,8461	0,8892	0,3855
CPEE	0,6821	0,8142	0,8568	0,3719
CSPE	0,7905	0,9286	0,9980	0,4370
DEMEI	0,6750	0,7132	0,9991	0,3190
ELETROACRE	0,5038	0,5174	0,6634	0,2349
ELETROCAR	0,6140	0,6565	0,7510	0,2968
JAGUARI	0,6721	0,9810	0,7778	0,5694
JOÃO CESA	0,8007	0,9628	0,9989	0,4722
MOCOCA	0,6307	0,7133	0,8267	0,3233
MUXFELDT	0,9944	0,9011	0,9990	0,4006
NACIONAL	0,6244	0,7758	0,9717	0,3546
NOVA PALMA	0,9295	0,8455	0,9983	0,3763
PANAMBI	0,5199	0,5604	0,6433	0,2510
POÇOS DE CALDAS	0,5808	0,6358	0,7168	0,2902
SANTA CRUZ	0,5315	0,6533	0,6562	0,3015
SANTA MARIA	0,5847	0,6136	0,7352	0,2790
SULGIPE	0,6526	0,5285	0,7654	0,2359
URUSSANGA	0,3792	0,5632	0,3923	0,2566
V. PARANAPANEMA	0,4241	0,5025	0,5426	0,2310
XANXERÉ	0,4539	0,5330	0,5540	0,2422

Uma comparação visual, dos dados editados nas Tabelas 8.1 e 8.2, pode revelar características importantes possibilitando uma melhor compreensão dos mesmos. Para tal, foi utilizado os *boxplots*, que são úteis para revelar o centro, a dispersão e a distribuição dos dados. Os mesmos são apresentados na Figura 8.1:

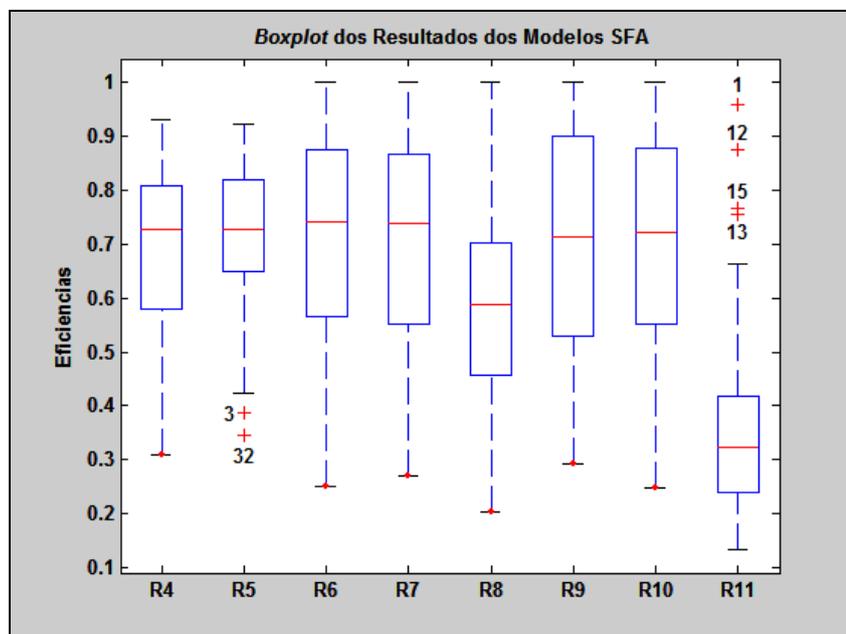


Figura 8.1: *Boxplot* dos índices de eficiência (Modelos SFA Clássica).

Examinando a Figura 8.1, percebe-se nitidamente que:

- Os resultados obtidos por *Half-Normal/Translog* (**R6** e **R7**) são muito semelhantes dando a entender desnecessária, nestes casos, a classificação das empresas por *clusters*;
- Os resultados obtidos por *Normal-Truncada/Cobb-Douglas* sem considerar a classificação das empresas por *clusters*; (**R9**), são mais dispersos;
- Identificação de *outliers* nos modelos **R5**(3, 32) e **R11** (1, 12, 13, 15);
- O modelo **R11** está totalmente destoante, fazendo-se necessária uma análise mais criteriosa.

8.3

Resultados dos modelos econométricos com enfoque Bayesiano

Esta seção, considerada a mais relevante desta tese, tem como objetivos estabelecer a metodologia de fronteira estocástica sob o paradigma Bayesiano e, também, sua aplicação para medir os custos operacionais. Em virtude da natureza econométrica, é admitido que u_j tenha distribuição Exponencial, *Half-Normal* e Normal-Truncada. Para esta última, tem-se a seguinte formulação:

$$u_j \sim iid N^+(\eta_j, \tau^2) \quad (8.3.1)$$

onde:

$$\eta_j = [1 \quad dummy_1 \quad dummy_2 \quad \ln(OPEX_j / MWh_j)]_x [\phi_1 \quad \phi_2 \quad \phi_3 \quad \phi_4]^T \quad (8.3.2)$$

É importante esclarecer que foi usado duas parametrizações:

a) **Parametrização 1** proposta por van den Broeck et al [94]:

$$\psi_l = \phi_l / \tau; \quad (l = 1, 4) \quad (8.3.3)$$

b) **Parametrização 2** - extensão de (a), ou seja:

$$\psi_l = \phi_l / \tau, \quad (l = 1, 2, 3, 4) \quad (8.3.4)$$

Vale lembrar que a variável $\ln(OPEX / MWh)$ foi padronizada passando a ter média zero (0) e desvio-padrão um (1). As distribuições a priori para os parâmetros do modelo (8.3.1) sugeridas por van den Broeck et al [94], são:

$$\phi_l \sim N(0,1); \quad (l = 1, 4) \quad (8.3.5)$$

$$\phi_l \sim \Gamma(2,1); \quad (l = 2, 3) \quad (8.3.6)$$

$$\tau^{-2} \sim \Gamma(5; 5 \ln^2(r^*)) \quad (8.3.7)$$

Os modelos Bayesianos foram implementados usando o *software* livre WinBUGS (*Bayesian inference Using Gibbs Sampling for Windows*), que pode ser obtido em www.mrc-bsu.cam.ac.uk/bugs/Welcome.htm. Para todas as situações, inicialmente foram realizadas 20.000 iterações (*burn-in*). Em seguida,

rodou-se mais 400.000 iterações. Utilizando a ferramenta *Sample Monitor Tool* (opção disponível no WinBUGS) atribuiu-se um *thin* de 40 amostras). Com relação à análise de convergência dos parâmetros, esta foi realizada de duas maneiras:

- Critério Raftery & Lewis [73] via *Bayesian Output Analysis Program* (BOA) *version* 1.1 disponível em (<http://www.public-health.uiowa.edu/boa>);
- Gráficos de autocorrelação serial disponibilizados no WinBUGS.

Assim sendo, de forma análoga aos modelos de fronteira estocástica com enfoque clássico, tem-se (ver Tabelas 8.3 e 8.4):

- **R12** - Resultados do modelo Bayesiano *Half-Normal/Cobb-Douglas*;
- **R13** - Resultados do modelo Bayesiano *Half-Normal/Translog*;
- **R14** - Resultados do modelo Bayesiano Exponencial/Cobb-Douglas considerando os *clusters*;
- **R15** - Resultados do modelo Bayesiano Exponencial/Cobb-Douglas desconsiderando os *clusters*;
- **R16** - Resultados do modelo Bayesiano Exponencial/Translog considerando os *clusters*;
- **R17** - Resultados do modelo Bayesiano Exponencial/Translog desconsiderando os *clusters*;
- **R18** - Resultados do modelo Bayesiano Normal-Truncada/Cobb-Douglas considerando os *clusters*;
- **R19** - Resultados do modelo Bayesiano Normal-Truncada/Cobb-Douglas desconsiderando os *clusters*;
- **R20** - Resultados do modelo Bayesiano Normal-Truncada/Translog considerando os *clusters*;
- **R21** - Resultados do modelo Bayesiano Normal-Truncada/Translog desconsiderando os *clusters*;

Tabela 8.3 – Índices de eficiência r_j^{SFA-B} .

Concessionária	R12	R13	R14	R15	R16
CLUSTER 1					
AES-SUL	0,9181	0,9058	0,9256	0,9352	0,9173
CEAL	0,6864	0,6683	0,8165	0,8163	0,8049
CEEE	0,4489	0,4638	0,6132	0,5926	0,6262
CELPA	0,4930	0,4970	0,6613	0,6451	0,6674
CELTINS	0,5293	0,5264	0,7054	0,6819	0,7051
CEPISA	0,6553	0,6506	0,7978	0,7948	0,7935
CERON	0,6117	0,6015	0,7711	0,7612	0,7643
COSERN	0,8399	0,8140	0,8873	0,8972	0,8760
ENERGIPE	0,8104	0,7794	0,8777	0,8837	0,8636
ESCELSA	0,8460	0,8322	0,8901	0,9015	0,8824
MANAUS	0,6189	0,6141	0,7795	0,7668	0,7794
PIRATININGA	0,8844	0,8722	0,9065	0,9181	0,9025
RGE	0,9077	0,8919	0,9193	0,9303	0,9106
SAELPA	0,7719	0,7581	0,8583	0,8667	0,8501
CLUSTER 2					
BANDEIRANTES	0,7761	0,8082	0,8696	0,8697	0,8900
CEB	0,4848	0,4870	0,6925	0,6356	0,7154
CELESC	0,6804	0,6978	0,8260	0,8142	0,8404
CELG	0,5963	0,6171	0,7739	0,7472	0,7938
CELPE	0,8399	0,8315	0,8935	0,8985	0,8937
CEMAR	0,6458	0,6516	0,8090	0,7912	0,8233
CEMAT	0,6003	0,6000	0,7824	0,7513	0,7914
CEMIG	0,7697	0,7986	0,8598	0,8660	0,8755
CERJ	0,7489	0,7426	0,8578	0,8551	0,8616
COELBA	0,7089	0,7191	0,8385	0,8312	0,8489
COELCE	0,7643	0,7647	0,8630	0,8635	0,8693
COPEL	0,8441	0,8481	0,8939	0,9006	0,8977
CPFL	0,8464	0,8454	0,8960	0,9026	0,9006
ELEKTRO	0,8735	0,8638	0,9091	0,9128	0,9078
ELETROPAULO	0,6798	0,7131	0,8180	0,8128	0,8447
ENERSUL	0,7980	0,7820	0,8811	0,8782	0,8767
LIGHT	0,7224	0,7408	0,8440	0,8400	0,8586
CLUSTER 3					
BOA VISTA	0,4057	0,4125	0,8519	0,5344	0,8699
BRAGANTINA	0,7502	0,7282	0,9441	0,8544	0,9443
CAUIÁ	0,6673	0,6428	0,9297	0,8020	0,9325
CAT-LEO	0,7456	0,7244	0,9414	0,8512	0,9425
CEA	0,5389	0,5269	0,8998	0,6923	0,9065
CELB	0,8244	0,7863	0,9536	0,8897	0,9518
CENF	0,6968	0,6660	0,9360	0,8221	0,9373
CFLO	0,7867	0,7620	0,9494	0,8722	0,9495
CHESP	0,8098	0,8024	0,9532	0,8817	0,9545
COCEL	0,8284	0,8116	0,9542	0,8916	0,9553
CPEE	0,8183	0,7976	0,9538	0,8867	0,9541
CSPE	0,8604	0,8359	0,9584	0,9069	0,9573
DEMEI	0,7877	0,7595	0,9502	0,8718	0,9501
ELETROACRE	0,7018	0,6727	0,9369	0,8256	0,9368
ELETROCAR	0,7835	0,7689	0,9492	0,8686	0,9500
JAGUARI	0,8060	0,8290	0,9521	0,8810	0,9575
JOÃO CESA	0,8315	0,8772	0,9575	0,8908	0,9628
MOCOCA	0,7860	0,7650	0,9494	0,8703	0,9509
MUXFELDT	0,8858	0,8810	0,9631	0,9175	0,9634
NACIONAL	0,7738	0,7395	0,9480	0,8647	0,9474
NOVA PALMA	0,8847	0,8760	0,9628	0,9183	0,9626
PANAMBI	0,6946	0,7063	0,9387	0,8185	0,9426
POÇOS DE CALDAS	0,7776	0,7571	0,9472	0,8664	0,9488
SANTA CRUZ	0,7395	0,7112	0,9413	0,8478	0,9421
SANTA MARIA	0,7747	0,7508	0,9474	0,8644	0,9480
SULGIPE	0,8071	0,7859	0,9518	0,8811	0,9515
URUSSANGA	0,5664	0,6807	0,9150	0,7130	0,9397
V. PARANAPANEMA	0,6386	0,6201	0,9241	0,7806	0,9278
XANXERÊ	0,6582	0,6618	0,9308	0,7942	0,9371

Tabela 8.4 – Índices de eficiência r_j^{SFA-B} .

Concessionária	R17	R18	R19	R20	R21		
CLUSTER 1	AES-SUL	0,9297	0,9785	0,9822	0,9828	0,9809	
	CEAL	0,8109	0,3283	0,4789	0,3076	0,4416	
	CEEE	0,6263	0,2373	0,3128	0,2275	0,2909	
	CELPA	0,6666	0,2366	0,3237	0,2240	0,2998	
	CELTINS	0,6961	0,1785	0,2731	0,1705	0,2513	
	CEPISA	0,8011	0,2577	0,3829	0,2497	0,3619	
	CERON	0,7626	0,2517	0,3743	0,2380	0,3431	
	COSERN	0,8880	0,5491	0,7836	0,5189	0,7368	
	ENERGIPE	0,8724	0,5033	0,7381	0,4639	0,6757	
	ESCELSA	0,8960	0,7136	0,9095	0,6900	0,8850	
	MANAUS	0,7744	0,4400	0,6204	0,4154	0,5616	
	PIRATININGA	0,9140	0,9690	0,9786	0,9732	0,9765	
	RGE	0,9233	0,9280	0,9723	0,9207	0,9689	
	SAELPA	0,9619	0,3798	0,5518	0,3674	0,5259	
	CLUSTER 2	BANDEIRANTES	0,8864	0,2630	0,9611	0,9656	0,9622
		CEB	0,6536	0,5144	0,3603	0,2553	0,3257
CELESC		0,8284	0,2915	0,6482	0,5395	0,6374	
CELG		0,7717	0,6042	0,3836	0,3149	0,3747	
CELPE		0,8960	0,2777	0,8028	0,6272	0,7941	
CEMAR		0,8006	0,2939	0,3948	0,2858	0,3788	
CEMAT		0,7613	0,6764	0,4037	0,3028	0,3783	
CEMIG		0,8794	0,4997	0,7953	0,7836	0,8466	
CERJ		0,8554	0,4441	0,6652	0,5064	0,6436	
COELBA		0,8416	0,4470	0,5808	0,4625	0,5754	
COELCE		0,8666	0,5736	0,6016	0,4657	0,5912	
COPEL		0,9023	0,8871	0,7518	0,6376	0,7715	
CPFL		0,9023	0,8234	0,9534	0,9277	0,9557	
ELEKTRO		0,9091	0,6672	0,9436	0,8638	0,9402	
ELETROPAULO		0,8360	0,4432	0,7940	0,7326	0,8326	
ENERSUL		0,8730	0,6861	0,6205	0,4737	0,5967	
CLUSTER 3	LIGHT	0,8529	0,1389	0,8356	0,7276	0,8463	
	BOA VISTA	0,5611	0,2799	0,2417	0,1245	0,2143	
	BRAGANTINA	0,8480	0,3043	0,6357	0,3369	0,5763	
	CAUIÁ	0,7961	0,2052	0,4661	0,2495	0,4195	
	CAT-LEO	0,8439	0,4143	0,4985	0,2827	0,4631	
	CEA	0,6931	0,2721	0,3488	0,1816	0,3102	
	CELB	0,8768	0,3721	0,7168	0,3676	0,6509	
	CENF	0,8092	0,2353	0,4849	0,2425	0,4395	
	CFLO	0,8645	0,3887	0,6767	0,3371	0,6217	
	CHESP	0,8816	0,3829	0,4622	0,2316	0,4683	
	COCEL	0,8867	0,4559	0,7149	0,3558	0,6719	
	CPEE	0,8806	0,3271	0,6860	0,3472	0,6393	
	CSPE	0,8974	0,2618	0,8051	0,4116	0,7483	
	DEMEI	0,8627	0,3012	0,6347	0,3077	0,5996	
	ELETROACRE	0,8134	0,5471	0,4584	0,2349	0,4181	
	ELETROCAR	0,8655	0,4143	0,5603	0,2776	0,5316	
JAGUARI	0,8947	0,3315	0,9038	0,5296	0,8776		
JOÃO CESA	0,9142	0,3852	0,8593	0,4739	0,9098		
MOCOCA	0,8654	0,3819	0,6106	0,3012	0,5691		
MUXFELDT	0,9154	0,3723	0,8140	0,3901	0,8487		
NACIONAL	0,8527	0,2446	0,6703	0,3402	0,6038		
NOVA PALMA	0,9136	0,3013	0,7471	0,3585	0,7577		
PANAMBI	0,8325	0,3301	0,4807	0,2330	0,4669		
POÇOS DE CALDAS	0,8626	0,2948	0,5346	0,2825	0,5047		
SANTA CRUZ	0,8379	0,2610	0,5507	0,2968	0,5004		
SANTA MARIA	0,8574	0,2257	0,5259	0,2714	0,4928		
SULGIPE	0,8749	0,2526	0,4832	0,2477	0,4680		
URUSSANGA	0,8137	0,2405	0,4389	0,2393	0,4617		
V. PARANAPANEMA	0,7770	0,2262	0,4255	0,2262	0,3839		
XANXERÊ	0,8068	0,2220	0,4433	0,2220	0,4160		

A Figura 8.2 reproduz as Tabelas 8.3 e 8.4 do estudo em que os resultados são descritos. É importante esclarecer que esta Figura 8.2 é uma representação visual das eficiências dos modelos Bayesianos, em função das formas funcionais Cobb-Douglas e Translog.

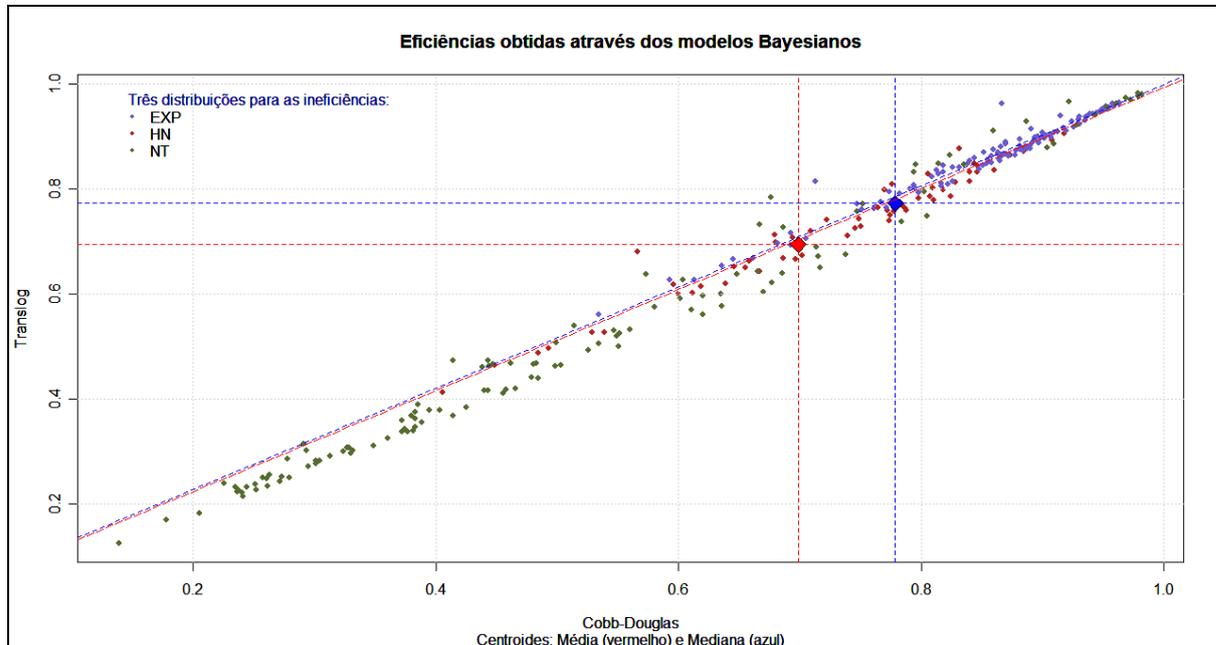


Figura 8.2: Representação gráfica dos índices de eficiência (Modelos Bayesianos).

Por inspeção visual da Figura 8.2, percebe-se que as estimativas, dos índices de eficiência, calculadas pelos modelos com distribuição Normal-Truncada são nitidamente distintas das eficiências estimadas a partir das distribuições Exponencial e *Half-Normal*. Também pode ser observado que a distribuição Exponencial é a que pior discrimina os resultados

Finalmente, foi selecionado dentre todos os modelos de natureza Bayesiana, aquele que pode ser considerado o mais apropriado. Esta seleção foi feita segundo o critério “*Deviance Information Criterion (DIC)*”, onde o menor valor indica um bom ajuste do modelo relativo ao número de parâmetros do mesmo [62]. Desta forma, assumiu-se que o modelo **R20** é o mais adequado. As estimativas deste modelo, obtidas a partir da distribuição a posteriori, via métodos MCMC, estão publicadas nas Tabela 8.5 e 8.6. Vale, ainda, ressaltar uma outra vantagem da técnica Bayesiana que é a concepção do intervalo de credibilidade para cada estimativa.

Tabela 8.5 – Estimativa pontual dos coeficientes de regressão e parâmetros do modelo de fronteira estocástica (enfoque Bayesiano).

Parâmetro	Variável	E(. dados)	D(. dados)
β_0	Constante	9,308	0,0748
β_1	lnMWh	0,1718	0,1084
β_2	lnNC	0,1433	0,0990
β_3	lnKM	0,4257	0,0965
β_{11}	(lnMWh)²	0,1292	0,0931
β_{22}	(lnNC)²	0,0369	0,0163
β_{33}	(lnKM)²	0,0071	0,0024
β_{12}	lnMWhlnNC	0,0774	0,0440
β_{13}	lnMWhlnKM	0,0113	0,0055
β_{23}	lnNClnKM	0,0126	0,0046
ϕ_1	Constante ($s_{j1} \equiv 1$)	0,4420	0,0236
ϕ_2	s_{j2}	0,4297	0,0680
ϕ_3	s_{j3}	0,2908	0,0460
ϕ_4	ln(OPEX / MWh)	0,4420	0,0236
σ^2	-	0,0055	0,0020

Tabela 8.6 – Índices de eficiência $r_j^{\text{SFA-B/NT-TL}}$; ($r^* = 0,95$).

Concessionária	r_j	d.p.	2,5%	mediana	97,5%
CLUSTER 1					
AES-SUL	0,9828	0,0169	0,9373	0,9878	0,9996
CEAL	0,3076	0,0306	0,2510	0,3063	0,3723
CEEE	0,2275	0,0194	0,1918	0,2265	0,2694
CELPA	0,2240	0,0209	0,1860	0,2226	0,2683
CELTINS	0,1705	0,0194	0,1356	0,1695	0,2118
CEPISA	0,2497	0,0264	0,2021	0,2485	0,3051
CERON	0,2380	0,0252	0,1921	0,2367	0,2907
COSERN	0,5189	0,0478	0,4323	0,5168	0,6195
ENERGIPE	0,4639	0,0454	0,3809	0,4618	0,5587
ESCELSA	0,6900	0,0563	0,5877	0,6866	0,8093
MANAUS	0,4154	0,0414	0,3397	0,4133	0,5038
PIRATININGA	0,9732	0,0250	0,9086	0,9807	0,9992
RGE	0,9207	0,0524	0,8045	0,9278	0,9962
SAELPA	0,3674	0,0358	0,3011	0,3663	0,4408
CLUSTER 2					
BANDEIRANTES	0,9656	0,0299	0,8895	0,9735	0,9990
CEB	0,2553	0,0247	0,2117	0,2536	0,3079
CELESC	0,5395	0,0407	0,4651	0,5373	0,6247
CELG	0,3149	0,0263	0,2663	0,3135	0,3704
CELPE	0,6272	0,0484	0,5382	0,6250	0,7280
CEMAR	0,2858	0,0277	0,2349	0,2843	0,3441
CEMAT	0,3028	0,0281	0,2527	0,3013	0,3617
CEMIG	0,7836	0,0616	0,6706	0,7812	0,9126
CERJ	0,5064	0,0409	0,4319	0,5050	0,5924
COELBA	0,4625	0,0367	0,3961	0,4608	0,5389
COELCE	0,4657	0,0379	0,3959	0,4643	0,5448
COPEL	0,6376	0,0493	0,5457	0,6355	0,7399
CPFL	0,9277	0,0469	0,8251	0,9335	0,9968
ELEKTRO	0,8638	0,0601	0,7483	0,8633	0,9790
ELETROPAULO	0,7326	0,0553	0,6308	0,7297	0,8487
ENERSUL	0,4737	0,0448	0,3915	0,4718	0,5672
LIGHT	0,7276	0,0534	0,6290	0,7253	0,8387
CLUSTER 3					
BOA VISTA	0,1245	0,0153	0,0970	0,1235	0,1576
BRAGANTINA	0,3369	0,0373	0,2690	0,3351	0,4162
CAUIÁ	0,2495	0,0276	0,1992	0,2482	0,3068
CAT-LEO	0,2827	0,0301	0,2276	0,2817	0,3453
CEA	0,1816	0,0212	0,1431	0,1807	0,2271
CELB	0,3676	0,0418	0,2919	0,3650	0,4572
CENF	0,2425	0,0293	0,1890	0,2409	0,3048
CFLO	0,3371	0,0402	0,2638	0,3349	0,4214
CHESP	0,2316	0,0302	0,1775	0,2297	0,2958
COCEL	0,3558	0,0429	0,2774	0,3533	0,4463
CPEE	0,3472	0,0412	0,2738	0,3448	0,4347
CSPE	0,4116	0,0478	0,3255	0,4090	0,5142
DEMEI	0,3077	0,0392	0,2359	0,3058	0,3901
ELETROACRE	0,2349	0,0277	0,1849	0,2332	0,2934
ELETROCAR	0,2776	0,0343	0,2159	0,2757	0,3510
JAGUARI	0,5296	0,0610	0,4200	0,5266	0,6595
JOÃO CESA	0,4739	0,0605	0,3660	0,4706	0,6014
MOCOCA	0,3012	0,0365	0,2356	0,2989	0,3782
MUXFELDT	0,3901	0,0519	0,2978	0,3871	0,4981
NACIONAL	0,3402	0,0393	0,2699	0,3376	0,4238
NOVA PALMA	0,3585	0,0466	0,2750	0,3561	0,4563
PANAMBI	0,2330	0,0297	0,1801	0,2314	0,2959
POÇOS DE CALDAS	0,2825	0,0336	0,2210	0,2808	0,3532
SANTA CRUZ	0,2968	0,0232	0,2381	0,2952	0,3638
SANTA MARIA	0,2714	0,0323	0,2136	0,2699	0,3398
SULGIPE	0,2477	0,0306	0,1934	0,2461	0,3124
URUSSANGA	0,2393	0,0305	0,1855	0,2373	0,3061
V. PARANAPANEMA	0,2262	0,0254	0,1796	0,2250	0,2802
XANXERÊ	0,2220	0,0277	0,1722	0,2205	0,2804

Nota: $r_j = \exp(-u_j)$.

Examinando as estimativas pontuais dos parâmetros da fronteira de eficiência (β), pôde-se concluir que estes não são sensíveis à escolha de r^* , diferente do que acontece com as estimativas dos índices de eficiência, σ^2 e ϕ .

A Figura 8.3 mostra o comportamento dos valores dos coeficientes de regressão ϕ para diferentes valores de r^* , ou seja:

1. Valores obtidos após análise de convergência segundo o Critério Raftery & Lewis [73] e considerando a **Parametrização 1**; $r^* = 0,60; 0,75; 0,80; 0,85; 0,95$;
2. Idem para a **Parametrização 2** e $r^* = 0,5; 0,55; 0,70; 0,75; 0,80; 0,90; 0,95$.

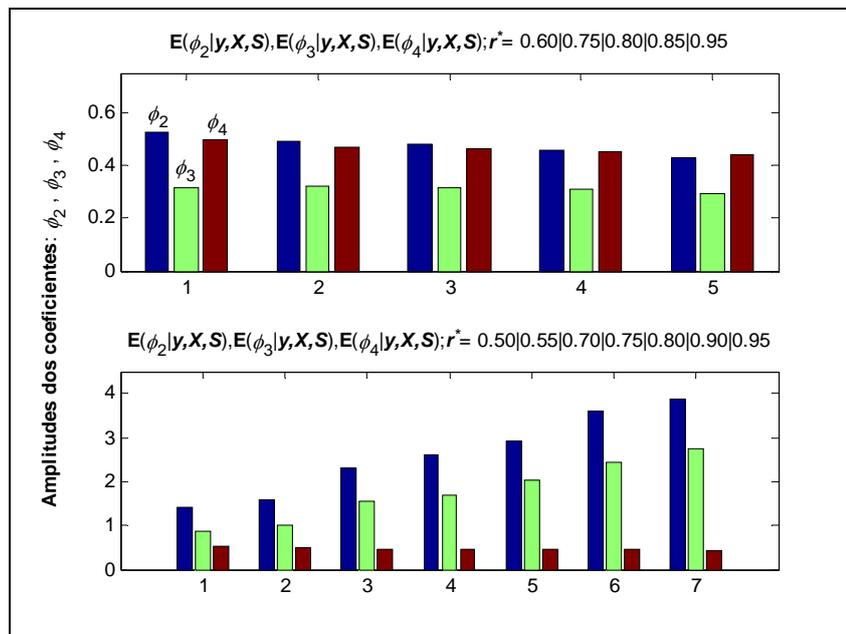


Figura 8.3: Estimativa pontual dos coeficientes de regressão para diferentes valores de r^* .

Nítidamente, pode ser visto que a **Parametrização 1** fornece resultados mais estáveis para ϕ_2, ϕ_3, ϕ_4 . Em contra-partida, a **Parametrização 2** contempla um maior número de modelos levando-se em conta a convergência de ϕ_2, ϕ_3, ϕ_4 .

9

Análise Multivariada de Dados

9.1

Introdução

Uma das dificuldades inerentes em estatística multivariada é a visualização dos dados, principalmente em dimensões maiores que três. No entanto, é importante citar que, muitas vezes, estas medidas apresentam informações redundantes.

A finalidade deste Capítulo é usar a técnica Análise Fatorial para identificar fatores (indicadores). De acordo com Pereira [71], a identificação de novas dimensões na análise fatorial vai recorrer à Análise de Componentes Principais (ACP). Para tal, inicialmente, é preciso calcular as correlações entre as variáveis.

Em concordância com o exposto, a próxima seção reproduz as matrizes de correlação amostral de Pearson e Spearman dos 21 modelos publicados.

9.2

Matriz de correlação de Pearson e Spearman

Muitas vezes é pertinente investigar a existência de uma associação entre dois modelos. A literatura especializada menciona que há diversos métodos para quantificar essa associabilidade. Para fins de análise, aqui são examinados os coeficientes de correlação de Pearson e Spearman.

É importante informar que o coeficiente de correlação de Pearson mede a relação linear dos dados amostrais. Já o coeficiente de correlação de Spearman usa postos de dados amostrais para medir a força da correlação entre dois modelos quaisquer sendo, portanto, desnecessária a suposição de linearidade entre os modelos.

Concomitante a estes conceitos, é interessante realizar um teste de hipótese onde as hipóteses, nula e alternativa, são as seguintes [92]:

- $H_0: \rho = 0$, isto é, não há correlação entre as duas variáveis.
- $H_1: \rho \neq 0$, isto é, há correlação entre as duas variáveis.

As Tabelas 9.1 e 9.2 registram os resultados dos coeficientes de correlação de Pearson e Spearman, respectivamente.

Tabela 9.1 – Coeficientes de correlação de Pearson.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21		
R1	1,000																						
R2	0,908	1,000																					
R3	0,870	0,760	1,000																				
R4	0,911	0,869	0,823	1,000																			
R5	0,908	0,820	0,849	0,982	1,000																		
R6	0,843	0,751	0,762	0,861	0,883	1,000																	
R7	0,791	0,649	0,717	0,800	0,834	0,908	1,000																
R8	0,863	0,886	0,820	0,908	0,866	0,751	0,631	1,000															
R9	0,696	0,663	0,642	0,756	0,749	0,595	0,543	0,812	1,000														
R10	0,764	0,697	0,716	0,854	0,865	0,939	0,889	0,775	0,612	1,000													
R11	0,589	0,596	0,600	0,618	0,617	0,496	0,407	0,693	0,864	0,432	1,000												
R12	0,918	0,859	0,820	0,962	0,976	0,861	0,798	0,890	0,808	0,836	0,677	1,000											
R13	0,906	0,850	0,836	0,927	0,938	0,808	0,748	0,909	0,849	0,780	0,720	0,980	1,000										
R14	0,640	0,668	0,392	0,699	0,656	0,634	0,607	0,596	0,443	0,643	0,287	0,706	0,680	1,000									
R15	0,898	0,831	0,786	0,939	0,958	0,834	0,783	0,848	0,777	0,805	0,622	0,981	0,960	0,707	1,000								
R16	0,613	0,659	0,359	0,669	0,611	0,588	0,558	0,593	0,435	0,600	0,277	0,668	0,655	0,994	0,671	1,000							
R17	0,889	0,784	0,801	0,894	0,919	0,788	0,764	0,827	0,781	0,748	0,629	0,951	0,965	0,671	0,971	0,644	1,000						
R18	0,545	0,538	0,599	0,575	0,573	0,401	0,319	0,653	0,843	0,337	0,958	0,611	0,654	0,136	0,571	0,128	0,580	1,000					
R19	0,724	0,691	0,687	0,776	0,772	0,624	0,551	0,828	0,986	0,615	0,918	0,823	0,860	0,437	0,781	0,425	0,783	0,891	1,000				
R20	0,529	0,529	0,607	0,546	0,538	0,348	0,267	0,653	0,827	0,291	0,936	0,576	0,635	0,096	0,540	0,095	0,561	0,992	0,871	1,000			
R21	0,718	0,692	0,711	0,758	0,752	0,581	0,511	0,838	0,976	0,573	0,916	0,803	0,858	0,407	0,759	0,402	0,775	0,896	0,992	0,889	1,000		

Tabela 9.2 – Coeficientes de correlação de Spearman.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21			
R1	1,000																							
R2	0,911	1,000																						
R3	0,860	0,735	1,000																					
R4	0,859	0,822	0,822	1,000																				
R5	0,863	0,786	0,855	0,982	1,000																			
R6	0,837	0,741	0,756	0,860	0,875	1,000																		
R7	0,764	0,615	0,723	0,774	0,808	0,892	1,000																	
R8	0,872	0,884	0,810	0,894	0,864	0,761	0,636	1,000																
R9	0,695	0,655	0,647	0,730	0,737	0,591	0,524	0,824	1,000															
R10	0,758	0,681	0,716	0,865	0,870	0,928	0,872	0,784	0,597	1,000														
R11	0,674	0,631	0,633	0,705	0,713	0,564	0,499	0,804	0,997	0,564	1,000													
R12	0,884	0,829	0,812	0,938	0,967	0,857	0,761	0,890	0,803	0,842	0,778	1,000												
R13	0,888	0,835	0,832	0,904	0,927	0,798	0,700	0,921	0,853	0,772	0,830	0,973	1,000											
R14	0,570	0,625	0,287	0,590	0,562	0,603	0,554	0,531	0,344	0,645	0,299	0,625	0,580	1,000										
R15	0,886	0,827	0,818	0,939	0,967	0,858	0,759	0,892	0,809	0,835	0,784	0,999	0,976	0,611	1,000									
R16	0,554	0,617	0,264	0,550	0,519	0,565	0,513	0,520	0,334	0,604	0,288	0,594	0,566	0,994	0,579	1,000								
R17	0,891	0,787	0,845	0,888	0,923	0,791	0,731	0,877	0,824	0,756	0,802	0,950	0,977	0,543	0,957	0,524	1,000							
R18	0,617	0,553	0,659	0,654	0,664	0,510	0,461	0,745	0,933	0,497	0,949	0,690	0,734	0,070	0,701	0,048	0,722	1,000						
R19	0,736	0,690	0,693	0,776	0,784	0,640	0,564	0,851	0,994	0,643	0,989	0,840	0,878	0,374	0,845	0,359	0,851	0,929	1,000					
R20	0,607	0,537	0,687	0,628	0,638	0,486	0,436	0,746	0,914	0,475	0,930	0,660	0,724	0,019	0,673	0,006	0,716	0,985	0,908	1,000				
R21	0,747	0,703	0,721	0,781	0,788	0,630	0,560	0,871	0,990	0,635	0,984	0,841	0,893	0,366	0,847	0,355	0,865	0,922	0,994	0,913	1,000			

Ao examinar as Tabelas 9.1 e 9.2, considerando um nível de significância habitualmente adotado de 5%, conclui-se que a maioria dos modelos apresenta coeficientes de correlação significativos, exceto: **R14-R18**, **R14-R20**, **R16-R18** e **R16-R20**.

As Figuras 9.1, 9.2 e 9.3 reproduzem as Tabelas 9.1 e 9.2 do estudo em que os coeficientes de correlação de Pearson e Spearman são descritos.

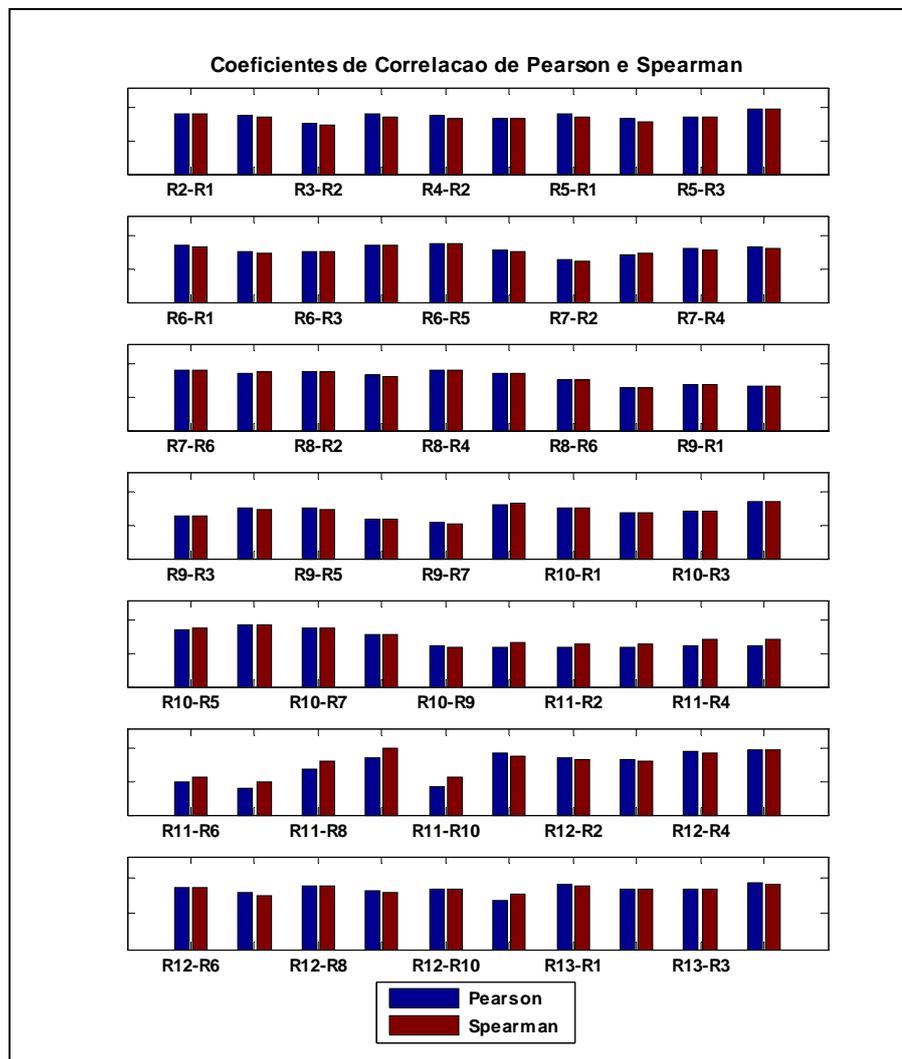


Figura 9.1 Coeficientes de correlação de Pearson e Spearman.

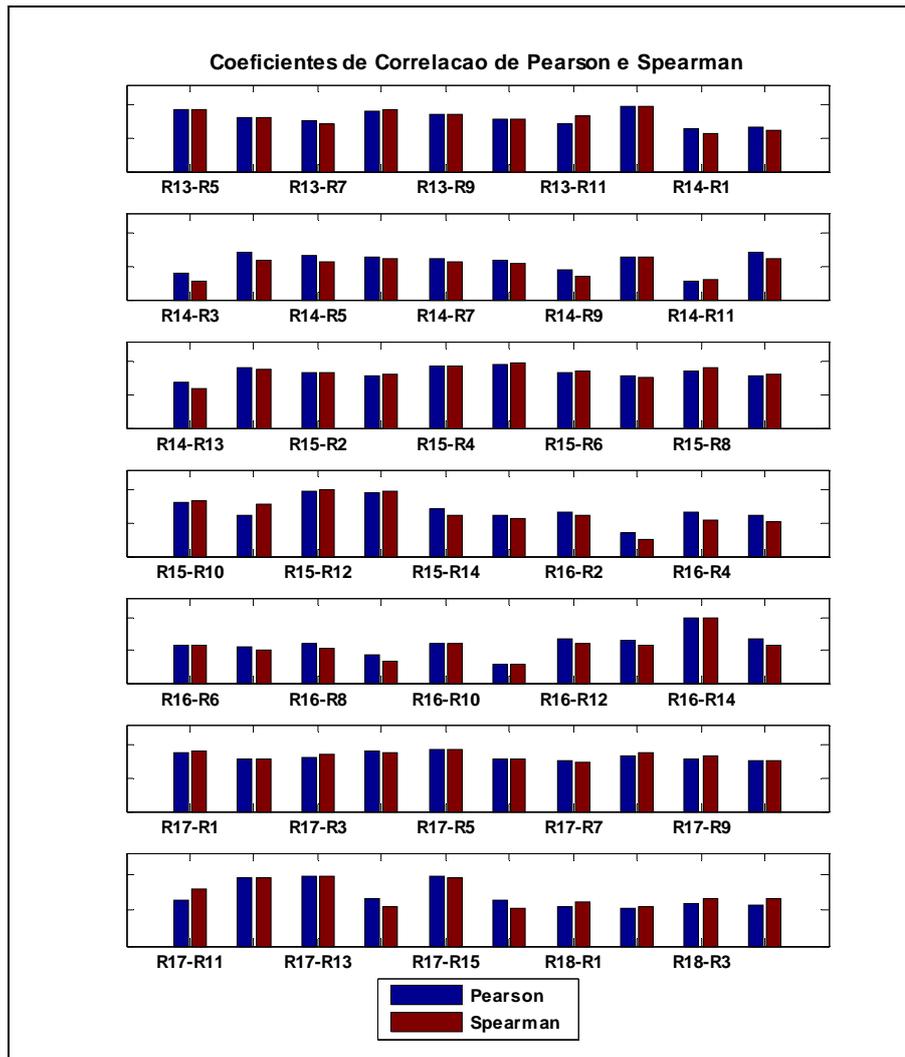


Figura 9.2: Coeficientes de correlação de Pearson e Spearman.

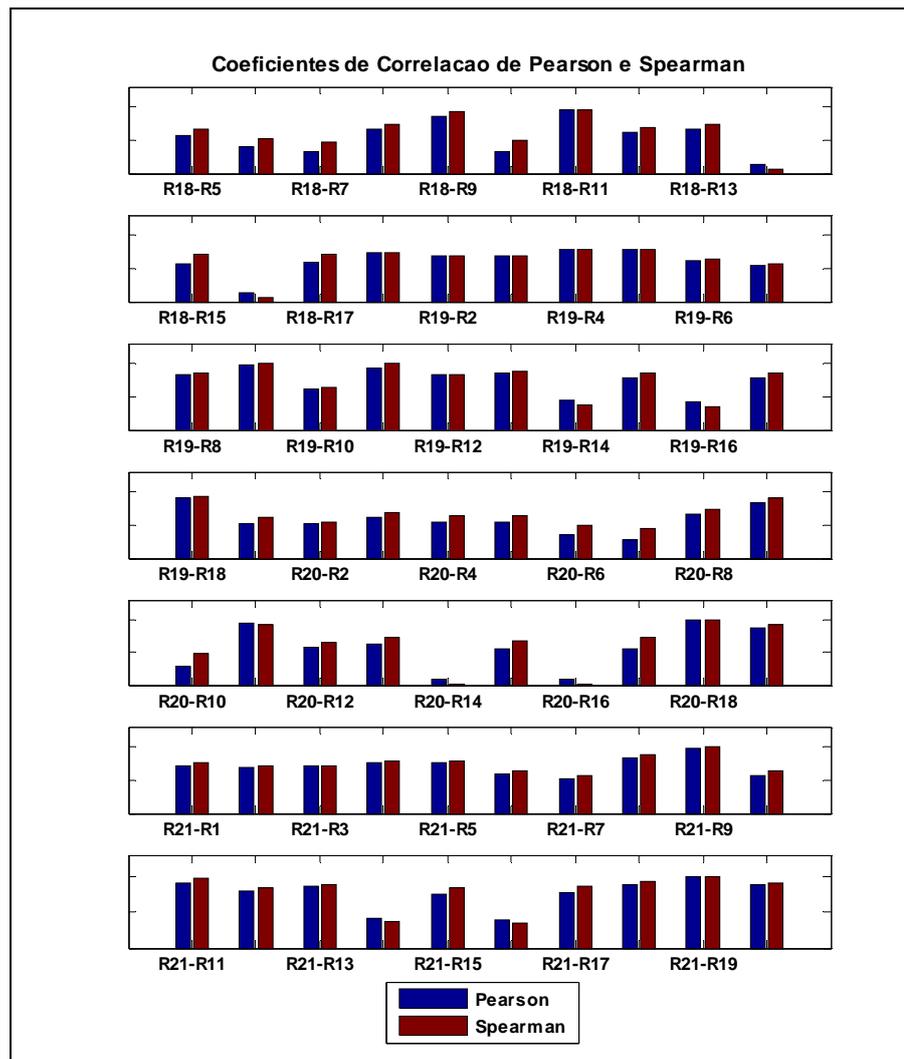


Figura 9.3: Coeficientes de correlação de Pearson e Spearman.

9.3

Análise Fatorial

Segundo Pereira [71], a Análise de Componentes Principais, popularmente chamada de PCA (*Principal Component Analysis*), é aplicada como recurso para a identificação de dimensões abstratas sobre as quais os objetos estudados possam ser projetados satisfatoriamente e, então, estudados em relação a essas dimensões ao invés das dimensões originais das medidas realizadas. Esta metodologia, considerada por muitos autores um dos tipos de análise fatorial, gera um novo conjunto de variáveis, denominado de Componentes Principais. Cada componente principal é uma combinação linear das variáveis originais e são ortogonais entre si, ignorando, então, a redundância de informação. Mingoti [66] cita que uma

outra particularidade desta técnica é explicar a variância e covariância de um vetor aleatório, composto de p -variáveis aleatórias.

Porém, antes de aplicar esses conceitos, é importante conceber algumas premissas da análise fatorial. Considere, inicialmente, o problema de analisar a normalidade do vetor aleatório. Há na literatura vários testes com esta finalidade. No entanto, aqui é realizada uma verificação baseada no gráfico de probabilidade qui-quadrado conforme mostrado na Figura 9.4:

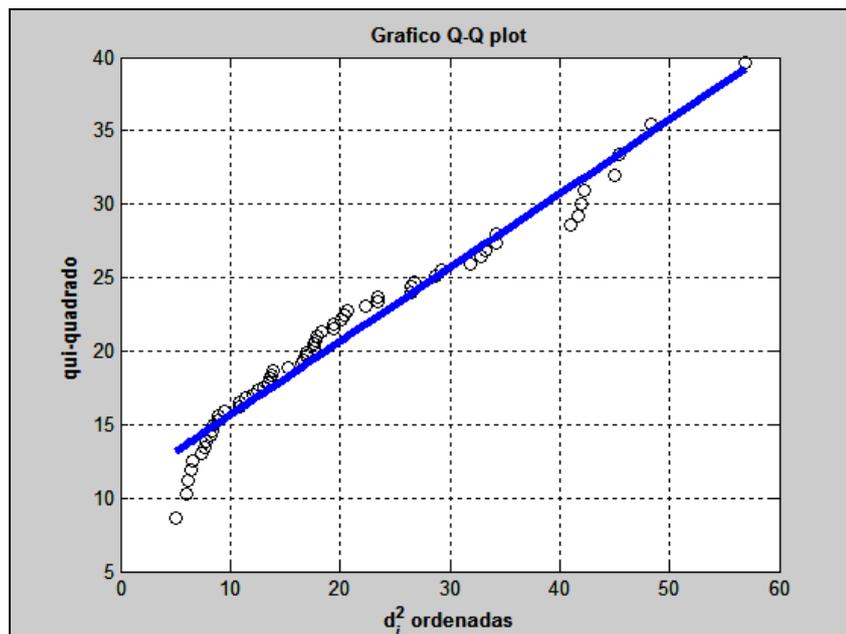


Figura 9.4: Gráfico qui-quadrado (Q-Q plot).

Por inspeção visual da Figura 9.4, não é difícil perceber que a normalidade p -variada pode ser coerente com os dados amostrais [66].

Uma outra questão que também merece destaque refere-se à determinação da medida de adequação dos dados, *Kaiser-Meyer-Olkin Measure of Sampling Adequacy* (KMO). Conforme reportado por Pereira [71], para uma boa adequação dos dados à análise fatorial, é necessário que esta medida assuma valores iguais ou superiores a 0,80. Assim sendo, vale informar que o KMO encontrado para este estudo foi de 0,833.

De posse dessas informações, tem-se que as variáveis latentes não observáveis F_1 e F_2 correspondem, respectivamente, a 73,60 % e 15,29% da variância total. Analisando estes fatores (F_1, F_2) , pode-se gerar um modelo que

explica, aproximadamente, 90% das variações das medidas originais. Neste ponto, é interessante registrar dois conceitos [47]:

1. **Comunalidades** (\hat{h}_i^2): quantia total de variância que uma variável original compartilha com todas as outras variáveis incluídas na análise;
2. **Variância específica** ($\hat{\psi}_i$): variância de cada variável, única àquela variável e que não é explicada ou associada com as outras variáveis na análise fatorial.

Para evitar ambigüidade, na interpretação dos resultados, muitas vezes é desejável empregar um método rotacional para simplificar a estrutura fatorial, e, dessa forma, conseguir soluções fatoriais mais simples e teoricamente mais significativas. Neste contexto, é importante esclarecer que foi utilizado, para este estudo, o método de rotação fatorial ortogonal VARIMAX [47].

A título de ilustração, a Tabela 9.3 apresenta os resultados finais.

Tabela 9.3 – Comparação entre os *factor loadings* não rotacionados e rotacionados.

Modelo (R _i)	F ₁	F ₂	F ₁	F ₂	\hat{h}_i^2	$\hat{\psi}_i$
R1	0,9251	0,1561	0,8180	0,4600	0,8802	0,1198
R2	0,8765	0,1277	0,7620	0,4520	0,7845	0,2155
R3	0,8505	-0,0197	0,6490	0,5500	0,7238	0,2762
R4	0,9579	0,1605	0,8460	0,4770	0,9433	0,0567
R5	0,9560	0,1552	0,8410	0,4800	0,9379	0,0621
R6	0,8542	0,3238	0,8680	0,2850	0,8344	0,1656
R7	0,7884	0,3750	0,8490	0,2040	0,7622	0,2378
R8	0,9295	-0,0290	0,7050	0,6070	0,8647	0,1353
R9	0,8706	-0,3683	0,4460	0,8340	0,8936	0,1064
R10	0,8273	0,3562	0,8670	0,2430	0,8114	0,1886
R11	0,7657	-0,5670	0,2390	0,9220	0,9079	0,0921
R12	0,9788	0,1152	0,8340	0,5260	0,9713	0,0287
R13	0,9788	0,0290	0,7800	0,5930	0,9589	0,0411
R14	0,6701	0,5385	0,8600	0,0020	0,7391	0,2609
R15	0,9526	0,1494	0,8350	0,4820	0,9297	0,0703
R16	0,6437	0,5187	0,8270	0,0010	0,6834	0,3166
R17	0,9358	0,1021	0,7920	0,5090	0,8862	0,1138
R18	0,7086	-0,6806	0,1230	0,9750	0,9653	0,0347
R19	0,8910	-0,3969	0,4440	0,8690	0,9515	0,0485
R20	0,6826	-0,7118	0,0840	0,9830	0,9726	0,0274
R21	0,8785	-0,4310	0,4120	0,8870	0,9575	0,0425

Examinando a Tabela 9.3, é possível verificar que o fator 1 reúne melhor as informações relativas às formas funcionais dos modelos (**R6** e **R10**). Por outro lado, o fator 2 refere-se mais à questão dos *clusters* (**R20** e **R18**). Não é difícil perceber que esses fatores são coerentes com a natureza deste estudo. Para auxiliar a interpretação dos fatores, pode-se ainda recorrer a um gráfico de dispersão que examine a localização dos modelos em um sistema de coordenadas criado pelos fatores. (ver Figuras 9.5 e 9.6).

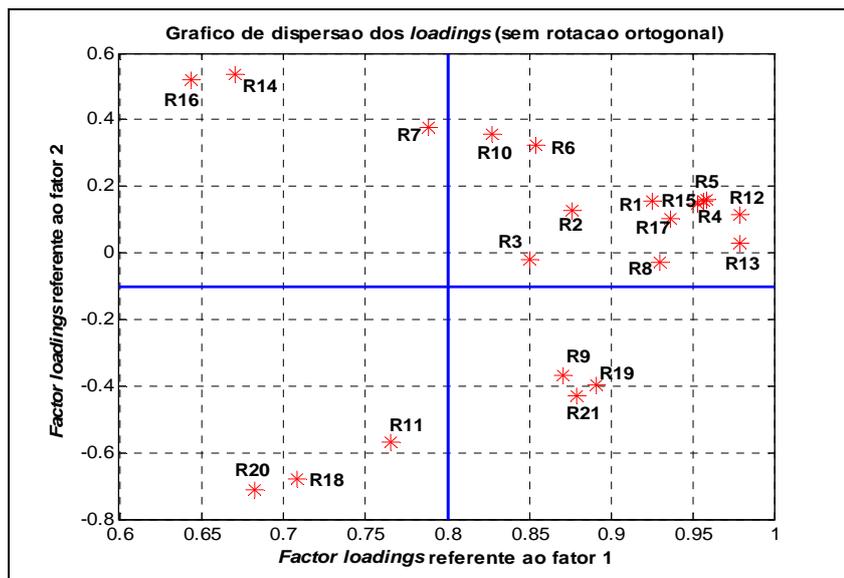


Figura 9.5: Gráfico de dispersão dos *loadings* (sem rotação ortogonal).

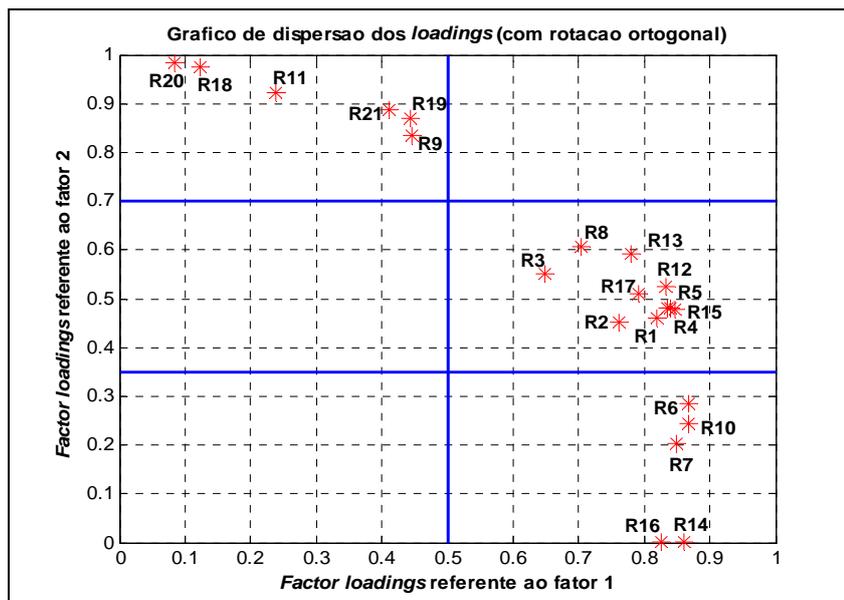


Figura 9.6: Gráfico de dispersão dos *loadings* (com rotação ortogonal).

Após rotacionar os eixos fatoriais, nota-se como os modelos se agrupam:

1. Grupo 1: **R9, R11, R18, R19, R20 e R21**;
2. Grupo 2: **R1, R2, R3, R4, R5, R8, R12, R13, R15 e R17**;
3. Grupo 3: **R6, R7, R10, R14 e R16**.

Um outro exercício que pode ser feito é combinar os *factor scores* das empresas em uma representação gráfica como a da Figura 9.7:

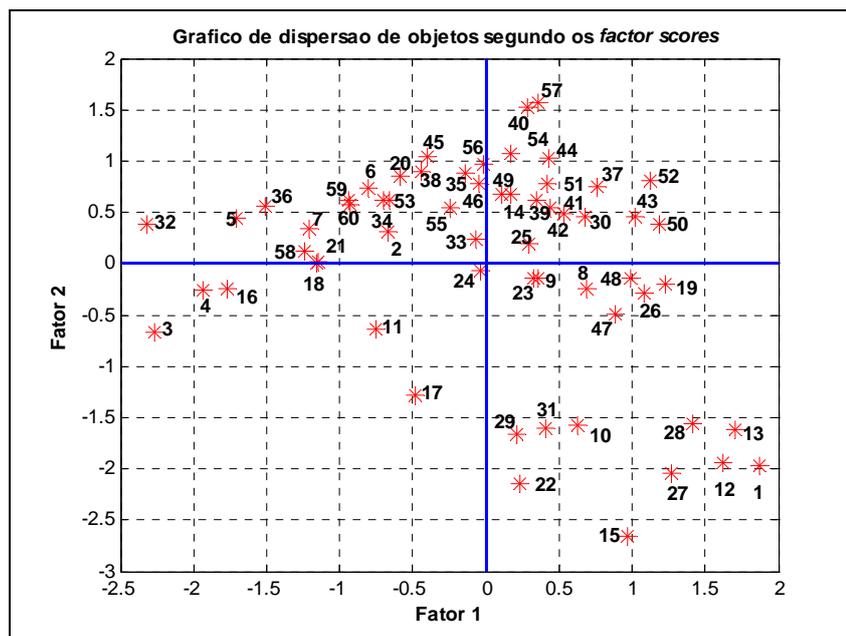


Figura 9.7: Gráfico de dispersão de objetos segundo os *factor scores*.

Inspecionando a Figura 9.4, percebe-se que a maioria das concessionárias (vide primeiro e segundo quadrantes) se destaca com relação à análise *cluster*. Conclusão análoga pode ser feita com relação à forma funcional (vide primeiro e quarto quadrantes). Já as empresas 3, 4, 11, 16, 17 e 24 (vide terceiro quadrante) dão uma contribuição menor a estes 2 fatores.

Por último, pode-se utilizar o primeiro fator para uma simples ordenação (*ranking*) das concessionárias em estudo. Mingoti [66] discorre que existem alguns questionamentos em relação a esta aplicação. Entretanto, vale informar que este procedimento tem sido amplamente utilizado para este fim.

Tabela 9.4 – *Ranking* das concessionárias.

	Concessionária	Posição
CLUSTER 1	AES-SUL	1
	CEAL	8
	CEEE	14
	CELPA	13
	CELTINS	12
	CEPISA	10
	CERON	11
	COSERN	4
	ENERGIPE	6
	ESCELSA	5
	MANAUS	9
	PIRATININGA	3
	RGE	2
	SAELPA	7
CLUSTER 2	BANDEIRANTES	5
	CEB	17
	CELESC	13
	CELG	16
	CELPE	3
	CEMAR	14
	CEMAT	15
	CEMIG	10
	CERJ	8
	COELBA	12
	COELCE	9
	COPEL	4
	CPFL	2
	ELEKTRO	1
	ELETROPAULO	11
	ENERSUL	6
	LIGHT	7
CLUSTER 3	BOA VISTA	29
	BRAGANTINA	18
	CAUIÁ	24
	CAT-LEO	19
	CEA	28
	CELB	6
	CENF	22
	CFLO	12
	CHESP	13
	COCEL	7
	CPEE	8
	CSPE	3
	DEMEI	9
	ELETROACRE	21
	ELETROCAR	17
	JAGUARI	5
	JOÃO CESA	4
	MOCOCA	15
	MUXFELDT	1
	NACIONAL	10
	NOVA PALMA	2
	PANAMBI	23
	POÇOS DE CALDAS	14
	SANTA CRUZ	20
	SANTA MARIA	16
	SULGIPE	11
	URUSSANGA	27
V. PARANAPANEMA	26	
XANXERÉ	25	

De posse das informações apresentadas na Tabelas 9.4, conclui-se o seguinte:

	<i>CLUSTER 1</i>	<i>CLUSTER 2</i>	<i>CLUSTER 3</i>
Melhor Empresa	AES-SUL	ELEKTRO	MUXFELDT
Pior Empresa	CEEE	CEB	BOA VISTA

10

Conclusões e Perspectivas

10.1

Conclusões

Para os próximos estágios de revisão tarifária, a ANEEL sinaliza a possibilidade de utilizar modelos DEA e SFA na determinação da eficiência dos custos operacionais, um elemento fundamental no cálculo do Fator X das distribuidoras. Para tal, um expressivo número de modelos e métodos foi apresentado nos Capítulos precedentes. De fato, com o auxílio destes indicadores específicos gerados, o órgão regulador (ANEEL) poderá avaliar melhor as concessionárias de energia elétrica antes de qualquer tomada de decisões.

Neste ponto é registrado contribuições de grande relevância desta tese:

1. Rede Neural de Kohonen: consultando a literatura especializada, é possível encontrar um significativo número de trabalhos envolvendo esta metodologia. No entanto, é categórico afirmar que na maioria das vezes, percebe-se claramente a falta de um embasamento teórico sobre o funcionamento da mesma. Para sanar esta carência metodológica, no Capítulo 2 foi reportado, ao leitor, uma série de particularidades desta técnica. Por razões já definidas no contexto desta tese, foi proposto uma extensão da Rede Neural de Kohonen denominada Rede Neural de Kohonen via simulação de Monte Carlo (MCRNK). Com relação aos resultados, estes foram processados e questionados no Capítulo 6;
2. Análise Envoltória de Dados (DEA): o Capítulo 3 abrange os principais fundamentos desta metodologia. Este procedimento visa uma melhor compreensão dos conceitos DEA permitindo, assim, uma maior versatilidade de análise dos resultados editados no Capítulo 7. Vale destacar o desenvolvimento de um novo modelo

DEA com restrições aos pesos (*adjusted contingent restrictions on weights*);

3. Modelos de fronteira estocástica (SFA) com ênfase freqüencista: assunto que envolve o conhecimento de vários conceitos e premissas. Ciente dessa necessidade, no escopo do Capítulo 5 o leitor encontra um texto claro, objetivo e necessário para que ele possa utilizar de forma produtiva. Continuando, foram produzidos 8 modelos e os resultados descritos no Capítulo 8 juntamente com algumas considerações pertinentes;
4. Modelos de fronteira estocástica (SFA) com enfoque Bayesiano: principal contribuição desta tese. Vale frisar que esta técnica, bastante complexa, permite algumas vantagens:
 - possibilita ao especialista incorporar informações adicionais sobre os parâmetros através das distribuições a priori. Estas, por sua vez, são combinadas com o auxílio do teorema de Bayes com a informação oriunda dos dados (função de verossimilhança). É importante sublinhar que esta flexibilidade (*Bayesian Adaptive Desing*) torna este método altamente confiável;
 - Parâmetros que não constituem o principal interesse no problema (parâmetros de distúrbio) podem ser facilmente descartados;
 - Inferências sobre os parâmetros podem ser feitas apenas com base nas informações disponibilizadas pela distribuição a priori e amostral.Aproveitando-se destas particularidades, foram gerados mais 10 modelos. Novamente, os resultados obtidos foram publicados e comentados no Capítulo 8;
5. Análise Fatorial: ferramenta fundamental para o desfecho desta tese. Com esta técnica, foi possível uma melhor interpretação dos 21 modelos. Além disso, também foram realizadas valiosas inferências.

Em concordância com o texto acima, é interessante mencionar que a escolha do método mais apropriado depende da aplicação. Por exemplo, no caso da transmissão de energia elétrica, a ANEEL adotou o método DEA [8], pois neste segmento há somente oito (8) DMUs, cada uma com três (3) registros anuais referentes ao período 2003/2005. A hipótese de que não houve mudança tecnológica neste intervalo de tempo, o modelo DEA adotado, na transmissão, considera uma transmissora em um determinado ano como sendo uma DMU. Assim, o modelo compara a performance de 24 DMUs (3 anos x 8 transmissoras). Com clareza percebe-se que esta situação não é recomendável para SFA clássica, mas plenamente aplicável em um contexto Bayesiano.

Na distribuição esta escolha não é tão simples, em virtude do maior número de DMUs. Neste caso, independente do método a ser escolhido pelo regulador, é interessante utilizar as três (3) abordagens, pois uma complementa a outra e permite avaliar a robustez dos resultados.

Por fim, na passagem dos índices de eficiência para a eficiência dos custos operacionais, a ANEEL pode considerar um critério semelhante ao adotado para as transmissoras, no qual são reconhecidos pelo menos 80% dos custos operacionais. Isto implica na seguinte normalização os índices de eficiência ($\theta_j \equiv r_j$) obtidos pelos modelos DEA e SFA:

$$\theta_j^* = 0,20 \cdot [\theta_j - \min(\theta_j)] / [\max(\theta_j) - \min(\theta_j)] + 0,80$$

10.2

Perspectivas

Todo estudo ou pesquisa visa atender um objetivo, propor soluções a um problema, ou mesmo aumentar o conhecimento sobre o problema. Porém, no decorrer de uma pesquisa, novas e diferentes questões vão surgindo sem que a maioria delas possam ser solucionadas em conjunto com o tema central do estudo. Desta forma, este tópico visa apresentar algumas questões consideradas relevantes para estudos futuros das metodologias mencionadas nesta tese:

- (1) DEA:
 - Utilizar métodos econométricos mais sofisticados para examinar o fenômeno de economia de escala;
 - Utilizar outras formas de fazer restrições nos pesos e analisar os resultados conjuntamente;
 - Utilizar modelos não-radiais;
 - Fazer análise de sensibilidade do *score* de eficiência de uma DMU. Em particular, na presença de outras DMU's na amostra, as quais são seus grupos de *benchmarks*;

- (2) Análise *Cluster*:
 - Utilizar outras medidas de similaridades tais como Kullback-Liebler, outros algoritmos de *clustering* tais como: dendograma, *fuzzy clustering method* visando análises comparativas entre as técnicas.

- (3) SFA Clássica:
 - Utilizar outras funções de produção tais como *Constant Elasticity of Substitution* (CES), *Generalized Production Function* (GPF);

- (4) SFA Bayesiana:
 - Utilizar outras funções de produção tais como *Constant Elasticity of Substitution* (CES), *Generalized Production Function* (GPF);
 - Utilizar outras funções de distribuição para o termo aleatório, u_j , tais como Gama Generalizada, Log-Normal, Weibull;
 - Modelos com dados em painel;
 - Fator de Bayes para determinar qual o modelo que melhor se ajusta aos dados.

Em todas as situações acima, incorporar aos modelos novas variáveis. Como exemplo, incluir a variável IASC que mensura a qualidade dos serviços prestados pelas empresas distribuidoras de energia elétrica.

Referências bibliográficas

- [1] AIGNER, D.; LOVELL, K.; SCHMIDT, P. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, n. 6, p. 21-37, 1977.
- [2] ALLEN, R.; et al. Weight restrictions and value judgements in data envelopment analysis: evolution, development and future directions. *Annals of Operations Research*, v. 73, p. 13-34, 1997.
- [3] ANEEL Nota Técnica no 097/2001 - SRE/ANEEL.
- [4] ANEEL Revisão tarifária periódica da concessionária de distribuição de energia elétrica Companhia de eletricidade do Estado da Bahia Coelba, Nota Técnica no 52/2003 - SRE/ANEEL, 3 de marco de 2003.
- [5] ANEEL Resolução Normativa no 55/2004, 5 de abril de 2004.
- [6] ANEEL Nota Técnica no 166/2006 - SRE/ ANEEL, 19 de maio de 2006a.
- [7] ANEEL Nota Técnica no 262/2006 - SRE/SFF/SRD/SFE/SRC/ANEEL, 19 de outubro de 2006b.
- [8] ANEEL Nota Técnica no 125/2007 - SRE/ANEEL, 11 de maio de 2007.
- [9] ARCOVERDE, F. D.; TANNURI-PIANTO, M. E.; SOUSA, M. C. S. Mensuração das eficiências das distribuidoras do setor energético brasileiro usando fronteiras estocásticas. In: XXXIII ENCONTRO NACIONAL DE ECONOMIA, Natal, RN, 2005, *Anais*.

- [10] BADIN, N. T. *Avaliação da produtividade de supermercados e seu benchmarking*. 1997. Dissertação de Mestrado em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis.
- [11] BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, n. 10, p. 3-66, 1995.
- [12] BATTESE, G. E.; CORRA, G. S. Estimation of a production frontier model with application to the pastoral zone of eastern Australia. *Australian Journal of Agricultural Economics*, n. 21, p. 169-179, 1977.
- [13] BAUER, P. Recent developments in the econometric estimation of frontiers. *Journal of Econometrics*, n. 46, p. 39-56, 1990.
- [14] BESAG, J.; et al. Bayesian computation and stochastic systems. *Statistical Science*, v. 30, n. 9, p. 1078-1092, 1984.
- [15] BOUSSOFIANE, A.; DYSON, R. G.; THANASSOULIS, E. Applied data envelopment analysis. *European Journal of Operational Research*, n. 52, p. 1-15, 1991.
- [16] CASELLA, G.; GEORGE, E. I. Explaining the Gibbs sampler. *The American Statistician*, v. 46, n. 3, p. 167-174, 1992.
- [17] CESPEDES, J. G. *Eficiência de produção: um enfoque bayesiano*, 2003. Dissertação de Mestrado em Agronomia, Escola Superior de Agricultura “Luiz de Queiroz” – Universidade de São Paulo, Piracicaba.
- [18] CHARNES, A.; et al. *Data envelopment analysis: theory, methodology, and application*. Boston, MA: Kluwer Academic, 1994.

- [19] CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision-making units. *European Journal of Operational Research*, n. 2, p. 429-444, 1978.
- [20] CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hastings algorithm. *The American Statistician*, n. 49, p. 327-335, 1995.
- [21] COELLI, T.; RAO, D. S. P.; BATTESE, G. E. *An introduction to efficiency and productivity analysis*. Boston, MA: Kluwer Academic, 1998.
- [22] COWLES, M. K.; CARLIN, B. P. Markov chain Monte Carlo convergence diagnostic: a comparative study. *Journal of the American Statistical Association*, n. 91, p. 883-904, 1996.
- [23] COOPER, W. W.; SEIFORD, L. M.; TONE, K. *Data envelopment analysis: a comprehensive text with models, applications, references and dea-solver software*. Boston, MA: Kluwer Academic, 2000.
- [24] DIALLO, M.; SOUZA, M. V. P.; GUEDES, L. E.; SOUZA, R. C. Comparing artificial DMUs and an adjusted contingent weight restrictions approach in the analysis of Brazilian retail banks efficiency. In: PROCEEDINGS OF OPERATIONS RESEARCH, S. NICKEL ET AL (Eds), SPRINGER, OPERATIONS RESEARCH 2007.
- [25] DOUCET, A.; GORDON, N. J.; KRISHNAMURTHY, V. Sequential simulation-based estimation of jump Markov linear systems. In: PROCEEDINGS OF THE IEEE CONFERENCE ON DECISION AND CONTROL, p. 1166-1171, 2000.
- [26] DUDA, R. O.; HART, P. E. *Pattern classification and scene analysis*, New York: Wiley, 1973.
- [27] DYSON, R. G.; et al. Pitfalls and protocols in DEA. *European Journal of Operational Research*, n. 132, p. 245-259, 2001.

- [28] DYSON, R. G., THANASSOULIS, E. Reducing weight flexibility in data envelopment analysis. *Journal of the Operational Research Society*, v. 39, n. 6, p. 563-576, 1988.
- [29] EHLERS, R. S. Métodos computacionalmente intensivos em estatística, versão n. 2, 2004. Disponível em <http://www.est.ufpr.br/~ehlers/notas/mci.pdf>. Acessado em Janeiro de 2005.
- [30] EHLERS, R. S. Métodos computacionais intensivos no R. Disponível em <http://leg.ufpr.br/~ehlers/ce718/praticas/praticas.html>. Acessado em Janeiro de 2005.
- [31] FARRELL, M. J. The measurement of productive efficiency. *Journal of the Royal Statistic Society*, n. 120, p. 253-281, 1957.
- [32] FERNÁNDEZ, C.; OSIEWALSKI, J.; STEEL, M. F. J. On the use of panel data in stochastic frontier models with improper priors. *Journal of Econometrics*, n. 79, p. 169-193, 1997.
- [33] FORNI, A. L. C. *On the detection of outliers in data envelopment analysis methodology*, 2002. Dissertação de Mestrado em Engenharia Mecânica-Aeronáutica, Instituto Tecnológico de Aeronáutica, São José dos Campos.
- [34] GAMERMAN, D. *Stochastic simulation for Bayesian inference*. London: Chapman and Hall, 1997.
- [35] GAMERMAN, D.; MIGON, H. S. *Inferência estatística: uma abordagem integrada*. Rio de Janeiro: UFRJ. Instituto de Matemática, 1997.
- [36] GELFAND, A. E.; SMITH, F. M. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, n. 85, p. 398-409, 1990.

- [37] GELFAND, A. E.; et al. Illustration of Bayesian inference in normal data models using Gibbs sampling . *Journal of the American Statistical Association*, n. 85, p. 972-985, 1990.
- [38] GELMAN, A.; et al. *Bayesian data analysis*. London: Chapman and Hall, 1995.
- [39] GELMAN, A. Inference and monitoring convergence. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Eds). *Markov chain Monte Carlo in practice*. London: Chapman and Hall, 1996. Cap. 8, p. 131-143.
- [40] GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, n. 6, p. 721-741, 1984.
- [41] GILKS, W.R.; et al. (Ed.). *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall, 1996.
- [42] GILKS, W. R.; ROBERTS, G. O. Strategies for improving MCMC. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Eds). *Markov chain Monte Carlo in practice*. London: Chapman and Hall, 1996. Cap. 6, p. 89-114.
- [43] GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. Introducing Markov chain Monte Carlo. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Eds). *Markov chain Monte Carlo in practice*. London: Chapman and Hall, 1996. Cap. 1, p. 1-19.
- [44] GOLANY, B. A note on including ordinal relations among multiplier in data envelopment analysis. *Management Science*, v.34, n. 8, p. 1029-1033, 1988.
- [45] GREENE, W. H. A gamma-distributed stochastic frontier model. *Journal of Econometrics*, n. 46, p. 141-164, 1990.

- [46] GRIFFIN, J. E.; STEEL, M. F. J. Bayesian stochastic frontier analysis using WinBUGS. *Journal of Productivity Analysis*, 2007. (To appear)
- [47] HAIR JR, J. F.; ANDERSON, R. E.; TATHAM, R. E.; BLACK, W. C. *Análise Multivariada de Dados*, 5.ed., Porto Alegre, Bookman, 2005.
- [48] HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, n. 57, p. 97-109, 1970.
- [49] HAYKIN, S. *Redes Neurais: princípios e prática*, 2.ed., Porto Alegre: Bookman, 2002.
- [50] JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. *ACM Computing Surveys*, v. 31, n. 3, p. 265-323, 1999.
- [51] JASMAB, T.; POLLIT, M. Benchmarking and regulation: international electricity experience. *Utilities Policy*, v. 9, n. 3, p. 107-130, 2000.
- [52] JONDROW, J.; et al. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, n. 19, p. 233-238, 1982.
- [53] KLEIT, A. N.; DEK, T. Measuring potential efficiency gains from deregulation of electricity generation: A Bayesian approach. *Review of Economics and Statistics*, v. 3, n. 83, p. 523-530.
- [54] KOOP, G. Recent progress in applied Bayesian econometrics. *Journal of Economic Surveys*. v. 8, n. 1, p. 1-34, 1994.
- [55] KOOP, G.; OSIEWALSKI, J.; STEEL, M. F. J. Bayesian efficiency analysis with a flexible form: the AIM cost function. *Journal of Business & Economic Statistics*. v. 12, n. 10, p. 339-346, 1994.

- [56] KOOP, G.; OSIEWALSKI, J.; STEEL, M. F. J. Posterior analysis of stochastic frontier models using Gibbs sampling. *Computational Statistics*. n. 10, p. 353-373, 1995.
- [57] KOOP, G.; OSIEWALSKI, J.; STEEL, M. F. J. Bayesian efficiency analysis through individual effects: hospital cost frontiers. *Journal of Econometrics*, n. 76, p. 77-105, 1997.
- [58] KUMBHAKAR, S. C.; LOVELL, C. A. K. *Stochastic frontier analysis*. Cambridge University Press, 2000.
- [59] MARTIN-DEL-BRIO, B.; SERRANO-CINCA, C. Self-organizing neural networks: the financial state of spanish companies. In: REFENES A.-P. (Ed). *Neural Networks in the Capital Markets*. Chichester, NY: John Wiley & Sons, 1995. Cap. 23, p. 341-357.
- [60] MARTINEZ, W. L.; MARTINEZ, A. R. *Computational statistics handbook with MATLAB*. London: Chapman and Hall, 2002.
- [61] MEDRANO, L. A. T. *Análise Bayesiana de modelos de fronteiras de produção estocásticas*, 2003. Dissertação de Mestrado – Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- [62] MEDRANO, L. A. T.; MIGON, H. S. *Crítérios baseados na “Deviance” para a comparação de modelos Bayesiano de fronteira de produção estocástica*. Rio de Janeiro: UFRJ, 2004. (*Technical Report 176*).
- [63] MEEUSEN, W.; VAN DEN BROECK, J. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, n. 18, p. 435-444, 1977.
- [64] MELO, L. G.; EHLERS, R. S. *Seminário WinBUGS*. Curitiba: UFPR, 2006. (*Technical Report 01*).

- [65] METROPOLIS, N. A.; et al. Equations of state calculations by fast computing machine. *Journal of Chemistry*, n. 21, p. 1087-1091, 1953.
- [66] MINGOTI, S. A. *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte: Editora UFMG, 2005.
- [67] NASCIMENTO JR, C. L.; YONEYAMA, T. *Inteligência artificial em controle e automação*. São Paulo: Editora Edgard Blücher, 2000.
- [68] ORME, C.; SMITH, P. The potential for endogeneity bias in data envelopment analysis. *Journal of the Operational Research Society*, v. 47, n. 1, p. 73-83, 1996.
- [69] OSIEWALSKI, J.; STEEL, M. F. J. Numerical tools for the Bayesian analysis of frontier models. *Journal of Productivity Analysis*, n. 10, p. 103-117, 1998.
- [70] PEDRAJA-CHAPARRO, F.; SALINAS-JIMENEZ, J.; SMITH, P. On the role of weight restrictions in data envelopment analysis. *Journal of Productivity Analysis*, n. 8, p. 215-230, 1997.
- [71] PEREIRA, J. C. R. *Análise de dados qualitativos – Estratégias metodológicas para as ciências da saúde, humanas e sociais*. São Paulo: Editora da Universidade de São Paulo, 2004.
- [72] PESSANHA, J. F. M.; SOUZA, R. C.; LAURENCEL, L. C. Usando DEA na avaliação da eficiência operacional das distribuidoras do setor elétrico brasileiro. In: XII CONGRESSO LATINO-IBEROAMERICANO DE INVESTIGACION DE OPERACIONES Y SISTEMAS, Ciudad de La Havana, CUBA, 2004, *Anais*.
- [73] RAFTERY, A. E.; LEWIS, S. M. How many iterations in the Gibbs sampler? In: BERNARDO, J. B. et al (Eds). *Bayesian Statistics 4*. Oxford: University Press, 1992. p. 763-773.

[74] RAFTERY, A. E.; LEWIS, S. M. Implementing MCMC. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Eds). *Markov chain Monte Carlo in practice*. London: Chapman and Hall, 1996. Cap. 7, p. 115-130.

[75] RESENDE, M. Relative efficiency measurement and prospects for yardstick competition in brazilian electricity distribution. *Energy Economics*, 2001.

[76] ROBERT C. P.; CASELLA, G. *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.

[77] ROBERTS, G. O. Markov chain concepts related to sampling algorithms Monte Carlo. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Eds). *Markov chain Monte Carlo in practice*. London: Chapman and Hall, 1996. Cap. 3, p. 45-57.

[78] ROBERTS, G. O. Computer intensive methods, 2000. Disponível em <http://www.maths.lancs.ac.uk/~robertgo/notes.ps>. Acessado em Setembro de 2007.

[79] ROLL, Y.; GOLANY, B. Alternate methods of treating factor weights in DEA. *OMEGA*, v. 21, n. 1, p. 99-109, 1993.

[80] SPALL, J. C. Estimation via Markov chain Monte Carlo. In: PROCEEDINGS OF THE IEEE CONFERENCE ON DECISION AND CONTROL, Anchorage, AK, 2002, p. 2559-2564, 2002.

[81] SOLLERO, M. V. K.; LINS, M. P. E. Avaliação da eficiência de distribuidoras de energia elétrica através da análise envoltória de dados com restrições aos pesos. In: XXXVI SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, São João Del Rei, MG: SOBRAPO, 2004, *Anais*.

[82] SOUZA, M. V. P. *Identificação da eficiência de empresas de telecomunicações empregando análise de envoltória de dados e redes neurais de*

Kohonen. 2002. Dissertação de Mestrado em Engenharia Mecânica-Aeronáutica, Instituto Tecnológico de Aeronáutica, São José dos Campos.

[83] SOUZA, M. V. P.; SOUZA, R. C.; PESSANHA, J. F. M. Custos operacionais eficientes das distribuidoras de energia elétrica: comparando modelos DEA e SFA. In: XXXIX SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, Fortaleza, RN: SOBRAPO, 2007, *Anais*.

[84] STEVENSON, R. E. Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics*, n. 13, p. 57-66, 1980.

[85] SUEYOSHI, T. Data envelopment analysis: formulations and economic implications. *Mathematic Japonica*, v. 42, n. 1, p. 187-200, 1995.

[86] TANNER, M.; WONG, W. The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association*, n. 82, p. 528-550, 1987.

[87] THANASSOULIS, A.; ALLEN, R. Simulating weight restrictions in data envelopment analysis by means of unobserved DMUs. *Management Science*, n. 44, p. 586-594, 1998.

[88] THANASSOULIS, A.; BOUSSOFIANE, A.; DYSON, R. G. Exploring output quality targets in the provision of perinatal care in England using data envelopment analysis. *European Journal of Operational Research*., n. 80, p. 588-607, 1995.

[89] THOMPSON, R. G.; SINGLETON, F.; THRALL, R.; SMITH, B. Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces*, v. 16, n. 6, p. 35-49, 1986.

[90] THOMPSON, R. G.; LANGEMEIER, L. N.; LEE, C. T.; THRALL, R. The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics*, v. 46, n.(1/2), p. 93-108, 1990.

- [91] TIERNEY, L. Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics*, n. 22, p. 1701-1762, 1994.
- [92] TRIOLA, M. F. *Introdução à estatística*. Rio de Janeiro: Editora LTC, 2008.
- [93] TSIONAS, E. G. Full likelihood inference in normal-gamma stochastic models. *Journal of Productivity Analysis*, n. 13, p. 183-205, 2000.
- [94] VAN DEN BROECK, J.; et al. Stochastic frontier models: a Bayesian perspective. *Journal of Econometrics*, n. 61, p. 273-303, 1994.
- [95] VIDAL, D. N. A.; TÁVORA JÚNIOR, J. L. Avaliação da eficiência técnica das empresas de distribuição de energia elétrica brasileiras utilizando a metodologia DEA. In: XXXV SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, Natal, RN: SOBRAPO, 2003, *Anais*.
- [96] WONG, Y.; BEASLEY, J. E. Restricting weight flexibility in data envelopment analysis. *Journal of the Operational Research Society*, v. 41, n. 9, p. 829-835, 1990.
- [97] ZANINI, A. *Regulação econômica no setor elétrico brasileiro: uma metodologia para definição de fronteiras de eficiência e cálculo do fator X para empresas distribuidoras de energia elétrica*, 2004. Tese de Doutorado em Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.
- [98] ZHAO, X. *Essays on the Bayesian estimation of stochastic cost frontier*, 2006. Disponível em http://etd.lsu.edu/docs/available/etd-11142006-124649/unrestricted/Zhao_dissertation.pdf. Acessado em Dezembro de 2007.
- [99] ZHU, J. *Quantitative models for performance evaluation and benchmarking*. Massachusetts: Springer, 2003.

Anexo 1

Conjunto de Dados

Concessionária	OPEX (1000 X RS)	CONSUMO (MWh)	NCONS	REDE (Km)
AES-SUL (1)	71078	7210207	945758	30618
BANDEIRANTES (2)	224515	18190382	1256051	11442
BOA VISTA (3)	19245	272988	52147	700
BRAGANTINA (4)	15058	568207	89990	2638
CAIUÁ (5)	26004	693894	172411	4227
CAT-LEO (6)	32405	909697	283214	15577
CEA (7)	22414	450311	103728	1582
CEAL (8)	63825	1716664	589634	10701
CEB (9)	145797	3220449	594749	8071
CEEE (10)	314731	6083952	1239523	25300
CELB (11)	11440	455613	120787	1669
CELESC (12)	298866	12424970	1772221	57911
CELG (13)	259062	5957214	1650166	121725
CELPA (14)	191828	3660284	1048235	15407
CELPE (15)	146498	6768699	2244483	38984
CELTINS (16)	43563	665081	245377	11470
CEMAR (17)	105760	2294546	1089195	23978
CEMAT (18)	122108	3015825	597359	56137
CEMIG (19)	664659	34279987	5395922	294022
CENF (20)	9877	263154	75228	891
CEPISA (21)	59854	1244252	598927	14863
CERJ (22)	171248	6632670	1924552	21395
CERON (23)	51300	1098511	327205	14476
CFLO (24)	5099	194883	38067	423
CHESP (25)	2441	58741	23063	2030
COCEL (26)	3960	161612	28045	979
COELBA (27)	253846	8533268	2858271	32839
COELCE (28)	155384	5356489	1881919	45648
COPEL (29)	198014	8600422	2952251	150194
COSERN (30)	56943	2562271	710348	15796
CPEE (31)	5847	230183	42285	1708
CPFL (32)	262031	18386196	2936710	38597
CSPE (33)	6402	296201	56000	2228
DEMEI (34)	2421	81780	22585	110
ELEKTRO (35)	151626	9996081	1757598	70861
ELETROACRE (36)	14540	361584	127044	2071
ELETROCAR (37)	4112	129222	27568	1293
ELETROPAULO (38)	647659	32563414	5584566	17749
ENERGIPE (39)	39048	1661057	403232	7294
ENERSUL (40)	71250	2647522	572272	95652
ESCELSA (41)	97563	5790063	910676	32572
JAGUARI (42)	5636	339742	23691	505
JOÃO CESA (43)	393	19647	2025	14
LIGHT (44)	395162	20916104	3477668	17035
MANAUS (45)	68067	2622032	345955	1670
MOCOCA (46)	4569	156390	33229	747
MUXFELDT (47)	478	20271	5437	54
NACIONAL (48)	9185	344838	79251	609
NOVA PALMA (49)	1171	46654	12212	593
PANAMBI (50)	2076	55158	11218	164
PIRATININGA (51)	99920	9258348	1138768	8399
POÇOS DE CALDAS (52)	8257	253696	51356	6368
RGE (53)	72134	5766016	1001638	37323
SAELPA (54)	64546	1946799	869309	16909
SANTA CRUZ (55)	21123	674281	146410	5826
SANTA MARIA (56)	8171	241305	59014	4091
SULGIPE (57)	6322	157908	77067	3220
URUSSANGA (58)	1737	47182	3507	93
V. PARANAPANEMA (59)	22547	551396	131361	4027
XANXERÊ (60)	5135	131661	22286	887

Anexo 2

Modelo Bayesiano

```
#####  
Análise Bayesiana de 60 empresas Distribuidoras de Energia Elétrica.  
# Termo de ineficiência distribuído como uma Normal-Truncada:  
#  $u \sim N^+(neta, lambda)$   
# Forma Funcional: Translog  
# Erro distribuído como uma Normal  
# 2 Variáveis Dummies (Cluster A = 0 1 / Cluster B = 1 0 / Cluster C = 1 1 )  
# Betas distribuídos como uma Normal Truncada  
# Thin inicial de 20.000  
# Simulação de 400.000  
# Thin final de 40 resultando em 10.000 amostras válidas  
# Arquivos para analisar no BOA: "nttlcomd095.ind" e "nttlcomd095.out"  
# Resultados no arquivo:  
# "resultados_eficiencias_normal_truncada_translog_COM_dummies_XX"  
# XX -> 0.50; 0.55; 0.60; 0.65; 0.75; 0.80; 0.85; 0.90; 0.95
```

```
#####
```

```
model  
{  
  fi0 <- psi / sqrt(invlambda)  
  
  for (i in 1:N){  
  
    neta[i] <- fi0 + fi[1] * dummy[i,1] + fi[2] * dummy[i,2] + teta * VZp[i]  
  
    u[i] ~ djl.dnorm.trunc(neta[i], invlambda, 0, 1000) # normal-truncada  
  
    eff[i] <- exp(-u[i])  
  
    mu[i] <- u[i] + alpha + beta[1] * z[i,1] + beta[2] * z[i,2] + beta[3] * z[i,3] + beta[4]  
* z[i,4]  
+ beta[5] * z[i,5] + beta[6] * z[i,6] + beta[7] * z[i,7] + beta[8] * z[i,8] + beta[9]  
* z[i,9]  
  
    Y[i] ~ dnorm(mu[i], prec)  
  }  
  
  alpha <- alpha0 - beta[1]*(0.0059) - beta[2]*(0.1050) - beta[3]*20.4382  
- beta[4]*0.0774 - beta[5]*1.0057 - beta[6]*428.4159
```

Priors

```
for (j in 1:p) {  
  
  beta[j] ~ djl.dnorm.trunc(0, tau, 0, 1000) # coeficientes independentes  
}  
  
alpha0 ~ dnorm(0.0, 0.000001)  
tau ~ dgamma(0.01, 0.01)  
prec ~ dgamma(0.01, 0.01)
```


0.1961, -3.0380, -39.7809, 0.0445, -0.1615, 4.3230, -0.0749, -1.0025, 184.7033, -
0.0028, 1.2461, -1.4004, 0.2182, -1.0327, 3.8594, -0.0272, -0.1452, 161.9573,
-0.2078, 5.4438, -22.5404, -0.2872, 0.8562, 4.7674, 0.0018, -0.0819, 206.9066, -
0.2705, -7.0926, 24.2273, 0.3941, -2.0273, 2.4123, 0.0826, 2.6894, 93.7311,
-0.7689, 9.1404, -43.9252, -0.1780, 0.6383, 4.1652, -0.0477, -0.4533, 176.9142, -
0.1280, -4.2364, 18.2877,
-0.0155, 1.0154, -2.6528, -0.0773, 0.2495, -112.0955, -0.0108, -0.1720, 19.9261,
0.2441, 0.2381, -1.2659, -0.0149, -0.8881, -60.8379, 0.0858, 4.7931, 6.5771, -
0.0591, -0.0153, -0.5387, -0.0745, -0.9977, -32.4258, -0.0048, -1.0598, 1.7843, -
0.0330, -0.9890, 0.4967, -0.0766, -0.2243, 9.8570, 0.0240, -0.5682,
-18.5061, -0.1053, 0.5994, -1.5541, -0.0675, -0.5095, -71.8056, -0.0700,
-1.8770, 13.3024, -0.2080, 0.5744, -1.4316, -0.0365, -0.5443, -67.1652, -0.1373, -
3.8420, 12.9116, -0.2779, 0.8261, -2.3542, -0.0034, -0.1388, -101.3840,
-0.2533, -4.9201, 16.8383, -0.0632, 1.2719, -3.2840, -0.0741, 0.8899,
-134.1466, -0.0790, -0.9839, 23.6185, -0.0488, -0.7154, -3.7311, -0.0755,
-0.6331, -149.2883, 0.0262, -0.7173, -10.1982, 0.2375, 0.4116, -3.2389, -0.0181, -
0.7389, -132.5971, 0.1257, 4.1864, 8.8841, 0.2297, 0.1266, -2.5162, -0.0219 ,
-0.9521, -107.2173, 0.0546, 4.2219, 4.1509, 0.2073, 0.0379, -2.0637, -0.0320,
-0.9853, -90.7928, 0.0304, 3.9159, 2.6252, -0.4240, 2.0079, -4.7566, 0.0975,
3.4587, -182.5027, -0.8835, -6.5571, 33.1341, -0.3387, 0.2840, -1.4658, 0.0334, -
0.8544, -68.4613, -0.1295, -6.3150, 7.3797, 0.1771, 0.0620, -3.2823, -0.0439,
-0.9779, -134.0894, 0.0306, 3.1388, 2.8652, 0.6691, 1.3161, -3.1196 , 0.3781,
1.0138, -128.4794, 0.9591, 11.6887, 24.6119, 0.2412, 2.9118, -8.0491, -0.0163,
8.0953, -274.9240, 0.7454, 3.0612, 37.3756, 0.0304, 0.6543, -3.2734, -0.0761, -
0.4293, -133.7845 , 0.0275, 0.6224, 13.0320, -0.2505, 1.8431, -6.8596, -0.0176,
2.7892, -244.0360, -0.4765, -3.3212, 26.4522, -0.2385, 1.2912, -2.3016,
-0.0233, 0.9435, -99.4774, -0.3248, -4.2195, 25.3217, 0.0734, 0.1961, -4.7922, -
0.0711, -0.9151, -183.6173 , 0.0239 , 1.2408, 4.7113, -0.0197, 1.3983, -5.2754, -
0.0772, 1.2540, -198.5047, -0.0208, -0.2102, 22.7936, 0.3960, -0.9698,
-1.7752, 0.0841, -0.2578, -80.1068, -0.3476, 7.5007, -16.1406, 0.1176, -0.3298 , -
0.5207, -0.0621, -0.9552, -31.7061, -0.0278, 2.4588, -4.4787, 0.1687, -0.5864, -
1.9128, -0.0469, -0.7739, -85.2247, -0.0840, 3.2330, -8.9192, -0.3359, -0.5387, -
2.1397, 0.0316, -0.8176 , -93.5803, 0.1432, -6.0397, -7.9363, 0.6945, 1.6994,
-6.3230, 0.4131, 2.2501, -229.1755, 1.2637, 9.8856, 25.4696, 0.0127, -0.1073,
-0.8786, -0.0770, -1.0057, -45.8366, -0.0000, 0.3625, -0.0464, 0.2805, 0.3799,
-3.5578, 0.0046, -0.7707, -143.4672, 0.1388, 4.8336, 8.1845), .
Dim = c(60,9)), p = 9, N =60,rstar=0.95)

Initial:

list(u=c(1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1,

1,1,1),

1,1,1),

fi=c(1,1), alpha0=0, psi=1, beta=c(0,0,0,0,0,0,0,0,0), tau=1, invlambda=1, prec=1)

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](#)