

# UMA APLICAÇÃO DE MODELOS DE FRONTEIRA DE PRODUÇÃO ESTOCÁSTICA COM ERRO NORMAL E COM ERRO T-STUDENT

Juliana Garcia CESPEDES<sup>1</sup>  
Roseli Aparecida LEANDRO<sup>1</sup>

- RESUMO: Os Escritórios de Desenvolvimento Rural (EDRs) dividem o estado em 40 sub-regiões onde são coletadas informações diárias sobre preços e produção agrícola. O objetivo deste trabalho é analisar a eficiência técnica de 40 EDRs do Estado de São Paulo, com relação a produção de carne, considerando-se os modelos de fronteira estocástica: com erro normal e com erro t-Student, utilizando a abordagem Bayesiana.
- PALAVRAS-CHAVE: Fronteira de produção estocástica; inferência Bayesiana; produção de carne; regressão com erro t-Student; regressão com erro normal.

## 1 Introdução

Os Escritórios de Desenvolvimento Rural foram criados a partir do Decreto nº 41.559, de 1º de janeiro de 1997, que dispõe sobre a estrutura administrativa da Coordenadoria de Assistência Técnica Integral (CATI), da Secretaria de Agricultura e Abastecimento do Estado de São Paulo, e divide o estado em 40 sub-regiões, que podem ser visualizadas na Figura 1, onde são coletadas informações diárias sobre preços e produção agrícola (Donadelli et al., 1997).

As informações coletadas dos EDRs neste trabalho são referentes a produção de carne de cada sub-região em razão do número de bovinos de corte, número de bovinos mistos, área de pasto natural, área de pasto cultivado e área de cana forrageira.

Considerando que o papel econômico de uma fazenda é importante para o mercado, a eficiência com que desempenha suas atividades torna-se um

---

<sup>1</sup>Departamento de Ciências Exatas, Universidade de São Paulo, Campus de Piracicaba - ESALQ/USP, CEP 13418-900 Piracicaba, SP, Brasil. E-mail: [jgcesped@esalq.usp.br](mailto:jgcesped@esalq.usp.br) / [rleandr@esalq.usp.br](mailto:rleandr@esalq.usp.br).



Figura 1 - Mapa dos EDRs do estado de São Paulo. **Fonte:** Instituto de Economia Agrícola, <<<http://www.iea.sp.gov.br/out/banco/mapas-edr-ra.php>>>

determinante central do bem-estar econômico da sociedade. Uma maneira de se calcular a eficiência de uma fazenda é por meio de um modelo de fronteira estocástica.

Os primeiros a modelar fronteira estocástica foram Aigner, Lovell e Schmidt (1977) e Meeusen e Van Den Broeck (1977) que, independentemente, propuseram a adição de um erro aleatório, com distribuição normal, nos modelos de fronteira para incluir uma variação natural que os dados podem conter. Nessa modelagem, além do erro com distribuição normal, é subtraído uma variável aleatória não-negativa para especificar a ineficiência do uso da técnica, assim, parte do resíduo se deve a ineficiência técnica e parte se deve a variação natural dos dados.

Algumas distribuições são geralmente utilizadas para o termo de ineficiência: Meeusen e Van Den Broeck (1977) propõem a distribuição exponencial, Aigner, Lovell e Schmidt (1977) consideram além da distribuição exponencial, a distribuição meia-normal (*distribution half normal*,  $|x| \sim N^0(a, b)$ ), Stevenson (1980), Greene (1980a) e Greene (1980b) propõem a distribuição gama.

Um dos primeiros artigos sobre modelos de fronteira estocástica utilizando a abordagem Bayesiana foram apresentados por Van Den Broeck et al. (1994), nesse trabalho, a variável responsável pela ineficiência técnica possui distribuição gama com parâmetro de forma conhecido. Migon e Medici (2004), também utilizando a abordagem Bayesiana, propõem utilizar a distribuição log-normal para o termo de ineficiência técnica da firma. Fernandez, Koop e Steel (2000) generalizam o modelo de fronteira de produção estocástica para dados que apresentam múltiplos produtos. Um estudo detalhado desse artigo pode ser encontrado em Cespedes (2003).

Fonseca et al. (2005) propõem o modelo de fronteira estocástica com erro t-Student, fazendo um estudo da influência de diferentes distribuições *a priori* assumidas para o parâmetro “grau de liberdade” para modelar distribuições com caudas pesadas e melhor avaliar a ineficiência técnica quando existem “outliers” nos dados.

Neste trabalho será analisada a eficiência técnica de 40 Escritórios de Desenvolvimento Rural (EDRs) do estado de São Paulo, com relação a produção de carne, considerando-se os modelos de fronteira estocástica: com erro normal e com erro t-Student utilizando a abordagem Bayesiana.

## 2 Material e métodos

### 2.1 Material

Os dados consistem de um produto e cinco insumos referentes a 40 EDRs do estado de São Paulo.

O produto  $y$  refere-se a produção de carne. E os insumos  $x$  são:

- i) número de bovinos de corte;
- ii) número de bovinos mistos;
- iii) área de pasto natural;
- iv) área de pasto cultivado; e
- v) área de cana forrageira.

### 2.2 Métodos

O modelo de fronteira de produção estocástica é definido (Pinilla, 2001):

$$y_i = f(x_i, \beta) - u_i + v_i, \quad (1)$$

sendo  $i = 1, 2, \dots, N$  o número de firmas,  $f(x_i, \beta)$  uma função de produção, em que,  $x_i$  são os insumos associados à produção do produto  $y_i$  e  $\beta = (\beta_0, \beta_1, \dots, \beta_k)$  é um vetor dos coeficientes de regressão;  $v_i$  uma variável aleatória com distribuição simétrica, independente e identicamente distribuída, com média zero e variância constante ( $\sigma_v^2$ ), representando a variação natural dos dados; e  $u_i$  uma variável aleatória com distribuição assimétrica, independente e identicamente distribuída, maior ou igual a zero e estatisticamente independente de  $v_i$ , que quantifica a ineficiência do uso da técnica  $i$ .

Uma função de produção relaciona insumos  $x_i$  e produtos  $y_i$  de um processo de produção, definindo o produto máximo conseguido por meio de um dado vetor de insumos (COELLI; RAO; BATTESE, 1999). As relações de insumo-produto

dependem em parte das quantidades de recursos empregados e, em parte, da forma pela qual esses insumos são combinados (a tecnologia de produção empregada).

As funções de produção são, geralmente, representadas por funções matemáticas, chamadas formas funcionais e a mais utilizada é a Cobb-Douglas:

$$y = Ax_1^{\beta_1} x_2^{\beta_2} \dots x_k^{\beta_k},$$

sendo  $\beta_1, \beta_2, \dots, \beta_k$  parâmetros positivos, os quais medem a elasticidade do produto, por exemplo, a elasticidade parcial  $\beta_1$  do produto com respeito ao insumo  $x_1$ , mede a variação percentual em  $y$  resultante do aumento de 1% em  $x_1$ , mantendo os outros insumos constante (Gujarati, 1995).

De forma equivalente, tem-se:

$$\ln(y) = \beta_0 + \beta_1 \ln(x_1) + \beta_2 \ln(x_2) + \dots + \beta_k \ln(x_k), \quad (2)$$

sendo o intercepto,  $\beta_0 = \ln A$ , estritamente maior que zero e os outros coeficientes maiores ou iguais a zero, para assegurar que os produtos marginais dos insumos sejam positivos, ou seja, que a produção tenha rendimentos crescentes e garantir  $\beta_1 + \beta_2 + \dots + \beta_k = 1$ , que são propriedades impostas pela estrutura de produção que exige retorno constante à escala.

A partir do termo de ineficiência  $u_i$  e considerando o modelo 1, tem-se que a eficiência técnica é medida como:

$$ET_i = \frac{f(x_i, \beta) \exp(\varepsilon_i - u_i)}{f(x_i, \beta) \exp(\varepsilon_i)} = \exp(-u_i). \quad (3)$$

A eficiência técnica (eq. 3) assume valores entre zero e um e indica a magnitude do produto da  $i$ -ésima firma, relativa ao produto que poderia ser produzido por uma firma completamente eficiente usando a mesma quantidade de insumos (COELLI; RAO; BATTESE, 1999). Assim sendo, uma firma será considerada eficaz se sua eficiência técnica estiver próxima de 1.

### 2.2.1 Modelo de fronteira de produção estocástica normal-gama

No modelo de fronteira de produção estocástica normal-gama, denominado Normal/G, considera-se:

$$y_i \sim N(f(x_i, \beta) - u_i, \sigma_v^2),$$

e que a distribuição do termo de ineficiência técnica adotada é:

$$u_i | P, \theta \sim G(P, \theta).$$

Tomando como base a abordagem utilizada em Tsionas (2000), as seguintes distribuições *a priori* para os parâmetros são especificadas:

$$\beta \sim N_k(b, H),$$

$$\begin{aligned}
P &\sim G(d, r), \\
\theta &\sim G(v, w), \\
\sigma_v^{-2} &\sim G(n, a),
\end{aligned}$$

sendo  $k = 1, \dots, K$  o número de coeficientes de regressão,  $N_k(b, H)$  a distribuição normal k-variada e  $G(a, b)$  a distribuição gama com esperança  $\frac{a}{b}$  e variância  $\frac{a}{b^2}$ .

Tsionas (2000) afirma que a distribuição *a priori* de  $\sigma_v^2$  precisa ser informativa para assegurar a existência da distribuição *a posteriori* em modelos de fronteira estocástica com dados “cross section” (dados de insumos e produtos de todas as  $N$  firmas referentes a um único tempo), e que a distribuição *a priori* para  $P$  não pode ser muito difusa, pois não seria possível identificar os desvios da fronteira relativos a ineficiência técnica.

Algumas funções dos parâmetros serão analisadas no trabalho: a média da ineficiência técnica, dada por  $m_u = \frac{P}{\theta}$ , a variância da ineficiência técnica, dada por  $\sigma_u^2 = \frac{P}{\theta^2}$ , a parte da variância da fronteira relativa a variância total explicada pela ineficiência  $VF = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$  e a eficiência técnica (eq. 3).

### 2.2.2 Modelo de fronteira de produção estocástica com erro t-Student

No modelo de fronteira de produção estocástica t-Student-lognormal, denominado de t-Student/LN, considera-se:

$$Y_i \sim t(f(x_i, \beta) - u_i, \sigma_i^2, \eta),$$

e que a distribuição do termo de ineficiência técnica é dada por:

$$u_i | \sigma^2 \sim LN(0, \sigma^2),$$

sendo  $LN(a, b^2)$  a distribuição lognormal.

Seguindo a abordagem de Fonseca et al. (2005), especifica-se as seguintes distribuições *a priori* para os parâmetros:

$$\begin{aligned}
\beta &\sim N_k(b, H), \\
\sigma^{-2} &\sim G(z, w), \\
\sigma_t^{-2} &\sim G(n, a), \\
\eta &\sim G(\alpha, \lambda).
\end{aligned}$$

Outras especificações *a priori* para o parâmetro que representa os graus de liberdade, ( $\eta$ ), da distribuição t-Student podem ser encontradas em Fonseca et al. (2005). Geweke (1993) mostra que a distribuição *a priori* para os graus de liberdade também não pode ser muito difusa, como para o parâmetro  $P$ , pois não seria possível identificar os desvios da fronteira relativos a ineficiência técnica.

Algumas funções dos parâmetros serão analisadas: a média da ineficiência técnica, dada por  $m_u = \exp(\frac{\sigma^2}{2})$ , a variância da ineficiência técnica, dada por  $\sigma_u^2 = \exp(\sigma^2)(\exp(\sigma^2) - 1)$ , a parte da variância da fronteira relativa a variância total explicada pela ineficiência  $VF = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$ , a variância da distribuição t-Student  $\sigma_v^2 = \frac{\eta}{\eta - 2}\sigma_i^2$  e a eficiência técnica (eq. 3).

### 3 Resultados e discussão

A função de produção considerada na análise foi a Cobb-Douglas (eq. 2) com 5 variáveis regressoras. Essa função possui algumas restrições, sendo que a soma dos coeficientes de regressão, menos o intercepto, precisa ser 1, o que garante retorno de escala constante.

Utilizando-se o software Winbugs gerou-se uma cadeia de tamanho 30.000 para cada modelo, descartou-se as 5.000 iterações iniciais, para o período de aquecimento da cadeia (“burn-in”) e para assegurar a independência entre os valores da amostra, foi considerado um espaçamento entre os pontos amostrados de tamanho 3 (“thin”). A convergência foi monitorada pela estatística (R) de Gelman e Rubin (1992).

Os hiperparâmetros das distribuições *a priori* para o modelo Normal/G considerados são:  $b = 0_k$ ,  $H = 10^4 I_k$ ,  $d = 1$ ,  $r = 0, 1$ ,  $v = 0, 1$ ,  $w = 0, 01$ ,  $n = 10^{-3}$ ,  $a = 10^{-3}$  e para o modelo t-Student/LN são  $b = 0_k$ ,  $H = 10^4 I_k$ ,  $n = 10^{-3}$ ,  $a = 10^{-3}$ ,  $\alpha = 0, 1$ ,  $\lambda = 0, 1$ ,  $z = 0, 01$ ,  $w = 0, 01$ .

Numa primeira análise verificou-se que os insumos área de pasto natural e área de cana forrageira não são significativos. Assim, foram considerados na análise apenas os insumos gado de corte, gado misto, área de pasto cultivado, reduzindo o número de variáveis regressoras para três.

A Tabela 1 apresenta um resumo *a posteriori* para os parâmetros  $\beta$ ,  $\sigma_v^2$ ,  $\sigma_u^2$ ,  $m_u$ ,  $VF$  dos modelos Normal/G e t-Student/LN e  $\eta$  para o modelo t-Student/LN.

Observou-se (Tabela 1) que as estimativas dos parâmetros da regressão considerando-se os modelos Normal/G e t-Student/LN são semelhantes, ou seja, a soma dos coeficientes de regressão, sem o intercepto, para o modelo Normal/G é 1,004 e para o modelo t-Student/LN é 1,005, isto é, aproximadamente 1, o que indica retorno de escala constante, satisfazendo as pressuposições da função de produção Cobb-Douglas (eq. 2).

Recomenda-se que o valor esperado da parte explicada pela ineficiência técnica na variação da fronteira relativa a variação total esteja próximo de 0,95 (Tsionas, 2000), ou seja, a ineficiência técnica explica cerca de 95% da variação dos dados. No modelo t-Student/LN encontrou-se o valor 0,837 e no modelo Normal/G encontrou-se o valor 0,240.

A Tabela 2 apresenta o resumo *a posteriori* da classificação do parâmetro eficiência técnica dos modelos Normal/G e t-Student/LN.

Tabela 1 - Resumo *a posteriori* dos parâmetros

Par.	Normal/G		t-Student/LN	
	Média	DP	Média	DP
$\beta_0$	2,463	0,096	3,421	0,103
$\beta_1$	0,551	0,024	0,553	0,029
$\beta_2$	0,374	0,021	0,375	0,026
$\beta_3$	0,079	0,035	0,077	0,043
$m_u$	0,044	0,047	1,001	0,0003
$\sigma_u^2$	0,0004	0,0005	0,0020	0,0005
$\sigma_v^2$	0,0012	0,0005	0,0005	0,0260
$VF$	0,240	0,229	0,837	7,270
$\eta$			7,945	7,000

Tabela 2 - Classificação em ordem crescente da eficiência técnica dos EDRs do estado de São Paulo

Classificação	EDR	EDR	Classificação	EDR	EDR
	normal	t-Student		normal	t-Student
1	33	33	21	4	4
2	34	34	22	10	30
3	17	17	23	30	10
4	40	40	24	36	36
5	8	3	25	19	19
6	3	8	26	37	27
7	5	5	27	27	37
8	12	12	28	28	26
9	15	15	29	20	28
10	31	31	30	26	20
11	16	16	31	22	25
12	29	29	32	25	22
13	39	39	33	35	35
14	14	14	34	9	9
15	21	21	35	32	32
16	2	2	36	7	7
17	24	24	37	1	1
18	11	11	38	23	23
19	18	18	39	13	13
20	6	6	40	38	38

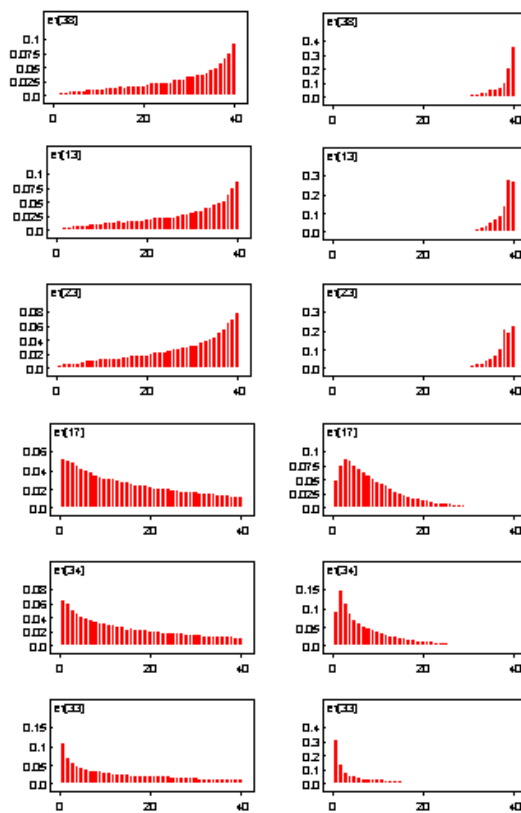


Figura 2 - Gráfico da distribuição *a posteriori* da classificação da medida de eficiência técnica de EDRs eficientes e de EDRs não eficientes para o modelo Normal/G (esquerda) e o modelo t-Student/LN (direita).

A Figura 2 mostra o gráfico da distribuição *a posteriori* da classificação da medida de eficiência técnica de EDRs eficientes e de EDRs não eficientes para os dois modelos.

Observou-se (Tabela 2) que os modelos mantiveram a mesma classificação para a eficiência técnica, os EDRs mais eficientes (EDR38 e EDR13) são os mesmos no modelo Normal/G ou no modelo t-Student/LN. Algumas posições foram trocadas, mas a diferença no valor das estimativas desses EDRs está na ordem da terceira casa decimal.

Observou-se (Figura 2) que a distribuição *a posteriori* da classificação da medida de eficiência no modelo t-Student/LN diferencia mais os EDRs eficientes dos EDRs não eficientes.



## Conclusões

Observando os resultados, pode-se concluir que o EDR mais eficiente é o número 38, e o menos eficiente o número 33.

A pressuposição de retorno de escala constante da função de produção Cobb-Douglas foi satisfeita nos dois modelos.

A classificação dos EDRs quanto a eficiência técnica manteve-se a mesma nos dois modelos.

A parte explicada pela variação da ineficiência técnica relativa a variação total da fronteira no modelo t-Student/LN é de 83,7% e no modelo Normal/G é de 24%.

## Agradecimentos

Ao pesquisador Oscar Tupy da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), São Carlos, SP, que gentilmente cedeu os dados para o artigo.

CESPEDES, J. G.; LEANDRO, R. A. An application of stochastic frontier models with normal error and with t-Student error. *Rev. Mat. Estat.*, São Paulo, v.24, n.3, p.25-34, 2006.

- **ABSTRACT:** *The Rural Development Offices (RDOs) divide the state in 40 sub-areas where daily information is collected about prices and agricultural production. The objective of this paper is to analyze the efficiency of 40 RDOs in São Paulo state, as regards meat production by taking account the stochastic production frontiers models with normal error and with t-Student error using the Bayesian approach.*
- **KEYWORDS:** *Stochastic production frontier; Bayesian inference; meat production; regression with t-Student error; regression with normal error.*

## Referências

- AIGNER, D. J.; LOVELL, C. A. K.; SCHMIDT, P. Formulation and estimation of stochastic frontier production function models. *J. Econ.*, Amsterdam, v.6, p.21-37, 1977.
- CESPEDES, J. G. *Eficiência de Produção: um enfoque Bayesiano*, 2003. 63f. Dissertação (Mestrado em Estatística e Experimentação Agronômica) – Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo – USP, Piracicaba, 2003.
- COELLI, T.; RAO, D. S. P.; BATTESE, G. E. *An introduction to efficiency and productivity analysis*. Boston: Klumber Academic, 1999. 275p.
- DONADELLI, A. et al. *Atividade agropecuária no estado de São Paulo: distribuição e valor da produção por escritório de desenvolvimento rural, 1995/96 e 1996/97*. *Inf. Econ.*, São Paulo, v.27, n.12, p.21-31, 1997.

- FERNANDEZ, C.; KOOP, G.; STEEL, M. A Bayesian analysis of multiple output production frontiers. *J. Econ.*, Amsterdam, v.98, p.47-79, 2000.
- FONSECA, T. C. O. et al. *Regressão multivariada com erro t-Student: uma análise Bayesiana*. Disponível em: [www.sbe.org.br/ebe26/031.pdf](http://www.sbe.org.br/ebe26/031.pdf). Acesso em: 9 mar. 2005.
- GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences (with discussion). *Stat. Sci.*, Hayward, v.7, p.457-511, 1992.
- GREENE, W. Maximum likelihood estimation of econometric frontier functions. *J. Econ.*, Amsterdam, v.13, p.27-56, 1980a.
- GREENE, W. On the estimation of a flexible frontier production model. *J. Econ.*, Amsterdam, v.13, p.101-115, 1980b.
- GEWEKE, J. F. Bayesian treatment of the independent Student-t linear model. *J. Appl. Econ.*, Chichester, v.8, p.519-540, 1993.
- GUJARATI, D. N. *Basic econometrics*. Singapore: McGraw-Hill International, 1995. 838p.
- MEEUSEN, W.; VAN DEN BROECK, J. Efficiency estimation from Cobb-Douglas production functions with composed error. *Int. Econ. Rev.*, Philadelphia, v.18, n.2, p.435-444, 1977.
- MIGON, H. S.; MEDICI, E. Bayesian hierarchical models for stochastic production frontier. *Estadística*, Santiago do Chile, v.56, p.1-19, 2004.
- PINILLA, A. A. *La medición de la eficiencia y la productividad*. Madrid: Ediciones Pirámide, 2001. 363p.
- STEVENSON, R. Likelihood functions for generalized stochastic frontier estimation. *J. Econ.*, Amsterdam, v.13, p.58-66, 1980.
- TSIONAS, E. Full likelihood inference in normal-gamma stochastic frontier models. *J. Prod. Anal.*, Dordrecht, v.13, p.183-206, 2000.
- VAN DEN BROECK, J.; KOOP, G.; OSIEWALSKI, J.; STEEL, M. F. J. Stochastic frontier models: a Bayesian perspective. *J. Econ.*, Amsterdam, v.61, p.273-303, 1994.

Recebido em 15.11.2005.

Aprovado após revisão em 13.08.2006.