

COMPUTING CONFIDENCE INTERVALS FOR OUTPUT ORIENTED DEA MODELS: AN APPLICATION TO AGRICULTURAL RESEARCH IN BRAZIL

Geraldo da Silva e Souza
Mirian Oliveira de Souza
Eliane Gonçalves Gomes

Brazilian Agricultural Research Corporation (Embrapa) – SGE
Parque Estação Biológica, Av. W3 Norte final, Asa Norte, 70770-901, Brasília, DF, Brazil
{geraldo.souza, mirian.souza, eliane.gomes}@embrapa.br

ABSTRACT

We define and model the research production at Embrapa, the major Brazilian institution responsible for applied agricultural research. The main theoretical framework is Data Envelopment Analysis – DEA. We explore the economic interpretation and the statistical properties of these models to compute confidence intervals for output oriented efficiency measurements, based on a parametric flexible model, defined by the truncated normal distribution. These results provide a better insight on the efficiency classification and allow comparison among the DMUs involved in the evaluation process taking into account inefficiency random variation.

Key words: Confidence intervals; DEA; Agricultural research.
Main area: DEA

RESUMO

Neste artigo é definido e modelado o sistema de produção de pesquisa da Embrapa, a maior instituição brasileira de pesquisa agropecuária. A ferramenta teórica principal é Análise Envoltória de Dados – DEA. Exploram-se a interpretação econômica e as propriedades estatísticas desses modelos, para calcular intervalos de confiança para as medidas de eficiência orientadas a produto. Aqueles são baseados em um modelo paramétrico flexível definido pela distribuição normal truncada. Esses resultados geram melhores entendimentos sobre a classificação das medidas de eficiência e permitem comparações entre as DMUs envolvidas no processo de avaliação, considerando a variação aleatória da ineficiência.

Palavras chave: Intervalos de confiança; DEA; Pesquisa agropecuária.
Área de classificação principal: DEA

1. Introduction

It is of importance to administrators of research institutions to have at their disposal measures and procedures making feasible an evaluation of the quantum of production, as well as the technical efficiency of the production process of their institutions. In times of competition and budget constraints a research institution needs to know by how much it may increase its production, with quality, without absorbing additional resources. The quantitative monitoring of the production process allows for an effective administration of the resources available and the observation of predefined research patterns and goals. In this context, the Brazilian Agricultural Research Corporation (Embrapa) developed a production model based on the input-output data of its research units. The theoretical framework for this model is the analysis of production frontiers, known as Data Envelopment Analysis (DEA).

Several uses are made of the efficiency measurements by Embrapa's administration. Those include monitoring of production targets, resource allocation and rewarding. Administrative actions regarding a given ranking of units will have more impact if they take into account the stochastic variation imbedded in the measurements of production variables. This leads to the consideration of statistical production models, from which one may infer statistical properties for efficiency estimates. For the stochastic frontier analysis, with proper parametric specifications of the production or cost functions, this is a natural process, as can be seen in Kumbhakar and Lovell (2000) and Coelli et al. (2005).

For the nonparametric frontier approaches induced by classical DEA (Coelli et al., 2005) or the Free Disposal Hull (FDH) of Deprins et al. (1984), some technical issues arise and a proper approach has to be put forward to guarantee the derivation of sound statistical results. This is the line of work carried out by Banker (1993), Banker and Natarajan (2004, 2008), Simar and Wilson (2004, 2007), Daraio and Simar (2007), Souza and Staub (2007) and Souza et al. (2009a).

In this article we combine the results of Banker (1993), Banker and Natarajan (2008), Simar and Wilson (2007) and Souza and Staub (2007) to come up with confidence intervals for DEA efficiency measurements, robust relative to production function choices and efficiency distributions within reason. These intervals are more appealing than those generated by the bootstrap of Simar and Wilson (2004, 2007) that may produce unexpected results, like one unit being regarded as inefficient after being observed as a benchmark or generating confidence limits that do not include observed efficiency measurements.

Our discussion proceeds as follows. In Section 2 we review the concepts leading to the models for which one may view DEA estimates as nonparametric maximum likelihood, and for which statistical properties may be derived for efficiency estimates. In Section 3 we review Embrapa's production model. Section 4 deals with the statistical results of our application and, finally, in Section 5 we summarize our findings.

2. Data Envelopment Analysis Production Models

Consider a production process composed of n decision making units (DMUs). Each DMU uses varying quantities of m different inputs to produce varying quantities of s different outputs.

Denote by $Y = (y_1, y_2, \dots, y_n)$ the $s \times n$ production matrix of the n DMUs. The r th column of Y is the output vector of DMU r . Denote by $X = (x_1, x_2, \dots, x_n)$ the $m \times n$ input matrix. The r th column of X is the input vector of DMU r . The matrices $Y = (y_{ij})$ and $X = (x_{ij})$ must satisfy: $p_{ij} \geq 0$, $\sum_i p_{ij} > 0$ and $\sum_j p_{ij} > 0$, where p is x or y .

The measure of technical efficiency of production (under constant returns to scale) for DMU $o \in \{1, 2, \dots, n\}$, denoted $E^{CR}(o)$, is the solution of the linear programming problem (1).

$$E^{CR}(o) = \max_{u,v} \frac{y'_o u}{x'_o v}$$

subject to

$$i) x'_o v = 1, \quad ii) y'_j u - x'_j v \leq 0, \quad j = 1, 2, \dots, n \quad \text{and} \quad iii) u \geq 0, v \geq 0$$
(1)

If we look at the coefficients u and v as input and output prices, we see that the measure of technical efficiency of production is very close to the notion of productivity (output income/input expenditure). Technical efficiency, in this context, basically, is looking for the price system (u, v) for which DMU o achieves the best relative productivity ratio.

The dual problem of the linear programming problem (1) has an important economic interpretation, which we will explore. This is $\min_{\theta, \lambda} \theta$, subject to

$$\begin{pmatrix} 0 & Y \\ x_o & -X \\ 0 & I \end{pmatrix} \begin{pmatrix} \theta \\ \lambda \end{pmatrix} \geq \begin{pmatrix} y_o \\ 0 \\ 0 \end{pmatrix}$$

or, equivalently, formulation (2).

$$\min_{\theta, \lambda} \theta$$

subject to

$$i) Y\lambda \geq y_o, \quad ii) X\lambda \leq \theta x_o \quad \text{and} \quad iii) \lambda \geq 0, \theta \text{ free}$$
(2)

The matrix products $Y\lambda$ and $X\lambda$, with $\lambda \geq 0$, represent linear combinations of the columns of Y and X , respectively, i.e., a sort of weighted averages of output and input vectors. In this way, for each λ , we can generate a new production relation, a new “pseudo” producer. Trivially, the set of DMUs $1, 2, \dots, n$ are included among those new producers. Making allowance for these newly defined production relationships, the question that the dual intends to answer is: What proportional reduction of inputs θx_o it is possible to achieve for DMU o and still produce at least output vector y_o ? The solution $\theta^*(x_o, y_o)$ is the smallest θ with this property.

We can define the concept of technical efficiency of production in a context of fixed inputs instead of fixed outputs, i.e., in a program of output augmentation. In this environment the measure of technical efficiency of production of DMU o , under constant returns to scale, is the one defined in (3).

$$\phi^*(x_o, y_o) = \max_{\phi, \lambda} \phi$$

subject to

$$i) Y\lambda \geq \phi y_o, \quad ii) X\lambda \leq x_o \quad \text{and} \quad iii) \lambda \geq 0, \phi \text{ free}$$
(3)

In the output augmentation program the question we ask is: what proportional rate ϕ can be uniformly applied to augment the output vector y_o , without increasing the input vector x_o ? The solution ϕ^* is the largest ϕ with this property. This is the approach we will explore here.

Questions of scale can be dealt properly imposing proper restrictions in the linear programming problem. One obtains the variable returns DEA imposing the additional condition $1'\lambda = 1$ on the weight vector λ .

We now turn our attention to production statistical models. We follow along the lines of Banker (1993). Suppose $m = 1$ (a single output) and assume the existence of a continuous frontier production function $g: K \rightarrow R$ defined on the convex and compact subset K of the positive orthant of R^s . For each DMU o , the input observations x_o are points in K . Let (4).

$$K^* = \left\{ x \in K; x \geq \sum_{i=1}^n \lambda_i x_i, \lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1 \right\} \quad (4)$$

The DEA frontier production function is defined for $x \in K^*$ by (5), and it can be shown that for DMU o , $g^*(x_o) = \phi_o^* y_o$, where we are assuming variable returns to scale.

$$g^*(x) = \sup_{\lambda} \left\{ \sum_{j=1}^n \lambda_j y_j; x \geq \sum_{i=1}^n \lambda_i x_i, \lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1 \right\} \quad (5)$$

Suppose that observations (x_o, y_o) are interior points to K and that they are generated in accordance with the statistical model (6),

$$y_o = g(x_o) - \varepsilon_o \quad (6)$$

where:

- The inefficiencies ε_o are *iid* with a common density $f(\varepsilon)$.
- The common distribution function $F(x)$ of the inefficiencies is strictly positive in $(0, +\infty)$.
- The inputs x_o represent a random sample from a density $h(x)$ strictly positive in the interior of K .
- The inputs x_o and the inefficiencies are independent.

Then:

- $g^*(x_o)$ is the nonparametric maximum likelihood estimate of $g(x_o)$ if $f(\varepsilon)$ is monotonically decreasing in $(0, +\infty)$.
- $g^*(x_o)$ is weakly consistent for $g(x_o)$.
- Let M be any fixed subset of DMUs. If n is large, the joint distribution of the estimated inefficiencies $\hat{\varepsilon}_j = y_j - g^*(x_j)$, $j \in M$, is, approximately, the joint distribution of the true inefficiencies ε_j , $j \in M$.

These results can be used to test hypothesis about the nature of the production process. An example is the verification of whether the technology shows constant returns to scale or variable returns to scale. We may perform this test comparing the empirical distribution functions of the $\hat{\varepsilon}_o$ (estimated inefficiency errors) under the assumptions of constant and variable returns to scale computing Kolmogorov-Smirnov test statistic (Conover, 1998).

The statistical properties of univariate DEA estimates were extended by Banker and Natarajan (2004, 2008) to encompass stochastic formulations of the production model, and by Souza and Staub (2007) to allow for non *iid* inefficient components.

Checking whether or not the production model fits the data is a matter of verifying if the postulated inefficiency distribution fits the efficiency observations. As in stochastic frontier analysis, the three commonly used family of distributions used to model inefficiency errors are the exponential, the half-normal and the truncated normal, the latter having flexibility properties (Coelli et al., 2005). Again, Kolmogorov-Smirnov test can be used to assess proper fits of these families. For the exponential and half-normal cases, Banker (1993) and Souza and Staub (2007) have derived simple statistics, which allow the study of the quality of the fit.

Under the exponential model, the statistic (7) will have qui-square distribution with $2n$ degrees of freedom, where d is the standard error of the inefficiency errors.

$$w_1 = \frac{2 \sum_{i=1}^n \hat{\varepsilon}_i}{d} \quad (7)$$

Under the assumption of a half-normal model, both w_2 and w_3 , in (8), are approximately chi-square with n degrees of freedom.

$$w_2 = \left(\frac{2}{\pi} \right) \frac{\sum_{i=1}^n (\hat{\varepsilon}_i)^2}{(\bar{\varepsilon})^2} \quad w_3 = \left(1 - \frac{2}{\pi} \right) \frac{\sum_{i=1}^n (\hat{\varepsilon}_i)^2}{d^2} \quad (8)$$

If one is concerned with $100(1-\alpha)\%$ confidence intervals for the efficiency measurement ϕ_i those can be computed as (9), where \hat{q} is the corresponding quantile of the estimated inefficiency error distribution. This follows from a similar result reported in Souza and Staub (2007).

$$\left[\hat{\phi}_i; \hat{\phi}_i + \frac{\hat{q}}{y_i} \right] \quad (9)$$

The quantile can be computed by parametric and non parametric methods. If the production models are subject also to random errors, assuming that the random error distribution has support in a bounded closed interval, the interval becomes (10), where the constant a should be estimated as in Banker and Natarajan (2008) or in Souza et al. (2009b). These conjectures in the context of interval estimation are new to the best of our knowledge. The extension of these results to the multivariate output is not clear in the literature.

$$\left[\hat{\phi}_i - \frac{a}{y_i}; \hat{\phi}_i - \frac{a}{y_i} + \frac{\hat{q}}{y_i} \right] \quad (10)$$

3. Embrapa's Production System

Embrapa's research system comprises 37 units (DMUs) of research centers. Input and output variables are defined from a set of performance indicators known to the company since 1991. The set comprises 28 outputs and 3 inputs.

We begin our discussion with the output. The output variables are classified into four categories: Scientific production; Production of technical publications; Development of technologies, products, and processes; Diffusion of technologies and image.

By scientific production we mean the publication of articles and book chapters aimed mainly to the academic world. We require each item to be specified with complete bibliographical reference.

The category of technical publications groups publications produced by research centers aiming, primarily, agricultural businesses and agricultural production.

The category of development of technologies, products, and processes groups indicators related to the effort made by a research unit to make its production available to society in the form of a final product. Only new technologies, products and processes are considered. Those must be already tested at the client's level in the form of prototypes, or through demonstration units, or be already patented.

Finally, the category of diffusion of technologies and image encompasses production variables related to Embrapa's effort to make its products known to the public and to market its image.

The input side of Embrapa's production process is composed of three variables. Personnel expenditure, Operational Costs (consumption materials, travel and services less income from production projects), and Capital (measured by depreciation).

All output variables are measured as counts and normalized by the company's mean of the corresponding variable for the year under analysis. Likewise, the inputs are normalized by the mean. As a final output we take a weighted average of all variables in all categories of production. The weights are user defined and reflect the administration perception of the relative importance of each variable to each research center or DMU.

Defining weights is a hard and questionable task. In our application in Embrapa we followed an approach based on the Law of Categorical Judgments of Thurstone (Torgerson, 1958). More details on this issue may be seen in Gomes and Souza (2008). The model is competitive with the AHP method of Saaty (1994) and is well suited when several judges are involved in the evaluation process. Basically, we sent out about 500 questionnaires to researchers and administrators and asked them to rank in importance – scale from 1 to 5, each production category and each production variable within the corresponding production category. A set of weights was determined under the assumption that the psychological continuum of the responses projects onto a normal distribution.

DEA models implicitly assume that the DMUs are comparable. This is not strictly the case in Embrapa. To make them comparable it is necessary an effort to define an output measure adjusted for differences in operation and perceptions. At the level of the partial production categories we induced this measure allowing a distinct set of weights for each DMU. In principle one could go ahead and use DEA with multiple outputs. This would minimize the effort of defining weights leaving to DEA the task of finding these coefficients. The problem with such approach is that there is a kind of dimensionality curse in DEA models. As the number of factors (inputs and outputs) increases, the ability to discriminate between DMUs decreases. Thus we found convenient to extend the weight system to produce a single measure of output y_o .

A personnel score was created for each unit dividing its number of employees by the company's mean. Outputs and inputs were further normalized by this variable. This further established a common basis to compare research units and avoided the incidence of spurious efficient units and zero output (shadow) prices, another common occurrence in multiple output models, and also a disturbing fact for management interpretation. A single output also allows the use of the statistical tests described in the previous section.

DEA models are known to be sensitive to outliers. In our application control of outliers is particularly important for output variables. In this context we use box plot fences to adjust the values of outlying observations. Values above $Q3 + 1.5(Q3 - Q1)$ are reduced to this value for any variable. Here $Q1$ and $Q3$ denote the first and third quartiles, respectively.

4. Statistical Results

In Table 1 we show Embrapa's production data for 2007.

We begin our analysis of the data in Table 1 checking the scale of operation of Embrapa's research units. Efficiency estimates under constant and variable returns and the corresponding inefficiency errors are shown in Table 2. This test is particularly important since we further normalize the data by the index of quantum of personnel. The Kolmogorov-Smirnov statistics for this hypothesis is $D=0.324$ with a p-value of 0.041, significant. The data used in this analysis is shown in Table 2 and are inefficiency errors $(\phi - 1)y$ computed from the efficiency measurements. On the basis of this test statistic we assume a variable returns to scale technology.

Table 1. Embrapa's mean normalized production data. Y is the output, $X1$ is personnel expenditures, $X2$ is other expenses and $X3$ is capital.

Research Center	Y	$X1$	$X2$	$X3$
1	1.991	1.411	2.266	1.922
2	1.415	1.010	0.935	0.765
3	1.289	1.009	1.444	1.344
4	0.647	0.847	0.872	0.905
5	1.395	1.150	1.311	1.222
6	0.553	0.891	0.923	0.835
7	1.991	1.182	1.848	3.305
8	1.086	1.252	0.845	0.702
9	0.444	0.967	1.259	1.592
10	1.118	0.973	0.537	1.114
11	1.024	0.832	0.822	0.829
12	0.745	1.152	1.038	3.212
13	0.886	1.147	0.940	1.202
14	0.773	1.079	0.949	0.816
15	0.760	0.993	1.029	1.026
16	0.989	1.188	0.722	1.003
17	1.164	0.862	0.918	0.981
18	1.010	1.003	1.298	0.983
19	1.098	0.958	0.819	0.904
20	1.108	1.360	1.484	1.448
21	1.185	0.943	0.998	1.525
22	0.419	0.931	0.639	0.658
23	0.555	0.930	1.113	0.800
24	0.978	0.913	0.919	0.941
25	0.569	1.057	1.255	0.978
26	0.802	1.134	0.985	0.714
27	1.262	0.908	0.696	0.505
28	1.421	0.963	1.223	1.038
29	0.589	0.758	0.628	0.480
30	0.651	0.940	1.147	0.830
31	0.574	1.116	0.911	0.850
32	0.977	1.061	1.015	0.718
33	1.670	0.814	0.686	0.606
34	0.502	1.056	0.788	0.715
35	1.131	1.028	0.965	1.101
36	0.633	0.937	0.821	1.112
37	1.841	1.089	1.400	1.754

Table 2. Product oriented measures of technical efficiency and inefficiency errors computed under constant returns to scale (CRS) and variable returns to scale (VRS).

Research Center	Efficiency		Error	
	CRS	VRS	CRS	VRS
1	1.453	1.000	0.903	0.000
2	1.464	1.208	0.657	0.294
3	1.605	1.393	0.780	0.506
4	2.688	2.625	1.092	1.051
5	1.692	1.292	0.965	0.407
6	3.300	3.106	1.272	1.164
7	1.218	1.000	0.434	0.000
8	1.783	1.560	0.850	0.608
9	4.464	4.049	1.538	1.354
10	1.170	1.000	0.190	0.000
11	1.667	1.647	0.683	0.663
12	3.175	2.370	1.620	1.020
13	2.584	1.965	1.403	0.855
14	2.865	2.227	1.442	0.949
15	2.681	2.299	1.278	0.987
16	1.776	1.698	0.768	0.690
17	1.520	1.471	0.605	0.548
18	2.037	1.745	1.047	0.753
19	1.789	1.553	0.866	0.607
20	2.519	1.661	1.683	0.733
21	1.631	1.481	0.748	0.571
22	3.717	2.882	1.139	0.788
23	3.436	3.096	1.352	1.163
24	1.916	1.764	0.896	0.747
25	3.817	3.096	1.603	1.193
26	2.451	2.114	1.164	0.894
27	1.103	1.000	0.129	0.000
28	1.391	1.236	0.555	0.335
29	2.247	1.000	0.735	0.000
30	2.959	2.653	1.275	1.076
31	3.861	2.994	1.642	1.145
32	2.024	1.736	1.001	0.719
33	1.000	1.000	0.000	0.000
34	3.817	3.367	1.414	1.188
35	1.866	1.538	0.979	0.609
36	3.040	2.695	1.291	1.073
37	1.214	1.001	0.393	0.002

Table 3 shows goodness of fit statistics related to the fit of the exponential and half-normal distributions. We see that statistic w_1 rejects the exponential model. The statistics w_2 and w_3 do not show evidence against the half-normal distribution.

Table 3. Goodness of fit tests for the exponential and half-normal distributions.

Distribution	df	Qui-square	p-value
Exponential (w_1)	74	119.57	< 0.001
Half-normal (w_2)	37	32.33	0.687
Half-normal (w_3)	37	48.18	0.103

The probability distribution function for the exponential distribution with parameter $\gamma > 0$ is given by (11) and for the truncated normal $N^+(\mu, \sigma)$ is (12), where $\Phi(\cdot)$ is the distribution function of the standard normal. The half-normal obtains making $\mu = 0$.

$$F(\varepsilon) = 1 - \exp\{-\gamma\varepsilon\} \quad (11)$$

$$F(\varepsilon) = \frac{\Phi\left(\frac{\varepsilon - \mu}{\sigma}\right) - \Phi\left(\frac{-\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \quad (12)$$

Estimating the parameters involved by maximum likelihood methods one seeks to maximize (13) for the exponential distribution and (14) for the truncated normal. Here $\phi(\cdot)$ denotes the probability density function of the standard normal distribution.

$$L(\gamma) = n \ln(\lambda) - \gamma \sum_{k=1}^n error_k > 0 \quad (13)$$

$$L(\mu, \sigma) = \sum_{k=1}^n \ln \frac{\phi\left(\frac{error_k - \mu}{\sigma}\right)}{\sigma \Phi(\mu/\sigma)} \quad (14)$$

Table 4 shows maximum likelihood estimates of γ , μ and σ . To compute maximum likelihood estimates we follow Simar and Wilson (2007) and eliminate errors associated with efficient units. This practice does not destroy the asymptotic approximations.

Table 4. Maximum likelihood estimates under the assumptions of exponential, half-normal and truncated normal distribution.

Parameter	Estimate	Standard error	p-value
σ (Half-normal)	0.854	0.108	<0.0001
μ (Truncated normal)	0.791	0.058	<0.0001
σ (Truncated normal)	0.315	0.044	<0.0001
γ (Exponential)	1.256	0.225	<0.0001

The Kolmogorov-Smirnov test statistics for the hypotheses truncated-normal, half-normal and exponential are 0.078, 0.291 and 0.316, respectively. The only non significant result is that related to the truncated normal distribution, since the 95% asymptotic quantile for the Kolmogorov-Smirnov test statistic under the null hypothesis is 0.238. These results are not in agreement with the statistical tests reported in Table 3, which do not reject the half-normal hypothesis. In this context, since location is significantly different from zero in the truncated normal specification, we interpret Table 3 results as lack of enough power to reject the half-normal null hypotheses and follow the truncated normal error assumption.

The quantile plot shown in Figure 1 provides visual support for our conclusions regarding goodness of fit. This is further evidence in favor of the truncated normal.

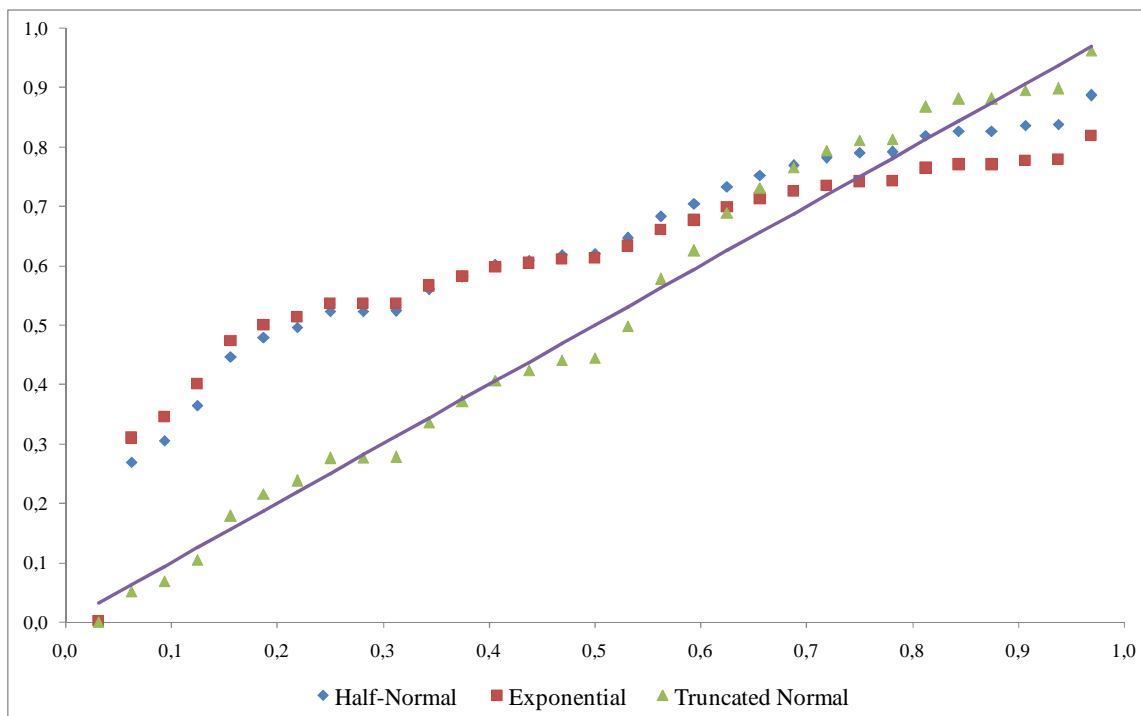


Figure 1. Quantile plots for Half-Normal, Exponential and Truncated Normal distributions.

Finally, Table 5 shows the efficiency estimates under the assumption of variable returns to scale and 95% individual confidence intervals, computed under a nonparametric assumption for DEA residuals. The 95% estimated quantile for the error distribution is 1.193. For the estimated truncated normal distribution the equivalent quantity is 1.310. The former leads to smaller confidence limits in the present case.

The main advantage of the technique used here relative to other confidence intervals suggested in the literature, as the proposal of Simar and Wilson (2004, 2007), is that the actual efficiency estimates are lower bounds and they define real possibilities for the corresponding population values. However, the imposition of stochastic errors may destroy this property. The parametric fitting is particularly convenient when covariates are present, since residuals will not be *iid* and quantiles will change with the DMU level.

5. Summary and Conclusions

Under the assumption of a statistical model containing only inefficiency errors we compute confidence intervals for the efficiency measurements computed using Data Envelopment Analysis. The production variables are normalized, which potentially leads one to consider variable returns to scale technologies. Nonparametric goodness of fit test indicates significant differences, relative to the hypothesis of constant returns to scale.

Residuals from the estimated production model follow a truncated normal distribution, which is used to compute the 95% upper limits of the individual efficient measurements. Only non efficient units are used for parameter estimation via maximum likelihood. The exponential and half-normal distributions fail to provide a good fit.

The intervals provided include actual efficiencies as possible values. This property however may be destroyed if one includes bounded random errors in the model specification.

Table 5. 95% individual confidence intervals for product oriented efficiency measures using empirical quantile. Intervals scaled to be in [0,1].

Research Center	Lower bound	Upper Bound
1	0.625	1.000
2	0.488	0.828
3	0.431	0.718
4	0.224	0.381
5	0.466	0.775
6	0.190	0.322
7	0.625	1.000
8	0.376	0.641
9	0.148	0.247
10	0.484	1.000
11	0.356	0.608
12	0.252	0.422
13	0.302	0.509
14	0.265	0.449
15	0.258	0.435
16	0.344	0.589
17	0.401	0.680
18	0.342	0.574
19	0.379	0.644
20	0.365	0.602
21	0.402	0.675
22	0.174	0.346
23	0.191	0.323
24	0.335	0.567
25	0.193	0.323
26	0.278	0.473
27	0.514	1.000
28	0.482	0.809
29	0.330	1.000
30	0.223	0.378
31	0.197	0.334
32	0.338	0.576
33	0.583	1.000
34	0.174	0.297
35	0.386	0.650
36	0.218	0.371
37	0.606	0.999

6. Acknowledgment

To the National Council for Scientific and Technological Development (CNPq), for the financial support.

7. References

- Banker, R.D.** (1993), Maximum likelihood, consistency and DEA: a statistical foundation, *Management Science*, 39(10), 1265-1273.
- Banker, R.D. and Natarajan, R.**, Statistical tests based on DEA efficiency scores, in Cooper, W.W., Seiford, L.M. and Zhu, J. (Eds.), *Handbook on Data Envelopment Analysis*, Kluwer International Series, Boston, 299-321, 2004.

- Banker, R.D. and Natarajan, R.** (2008), Evaluating contextual variables affecting productivity using data envelopment analysis. *Operations Research*, 56, 48-58.
- Coelli, T.J., Prasada Rao, D.S., O'Donnell, C.J. and Battese, G.E.**, *An Introduction to Efficiency and Productivity Analysis*, 2nd Edition, Springer, New York, 2005.
- Conover, W.J.**, *Practical Nonparametric Statistics*, Wiley, New York, 1998.
- Daraio, C. and Simar, L.** *Advanced Robust and Nonparametric Methods in Efficiency Analysis*, Springer, New York, 2007.
- Deprins, D., Simar, L. and Tulkens, H.** Measuring labor inefficiency in post offices, in Marchand, M., Pestieau, P. and Tulkens, H. (Eds.), *The Performance of Public Enterprises: Concepts and Measurements*, North-Holland, Amsterdam, 243-267, 1984.
- Gomes, E.G. e Souza, G.S.** (2008), Percepções psicométricas sobre a importância de atividades de pesquisa, *Anais do XL Simpósio Brasileiro de Pesquisa Operacional*.
- Kumbhakar, S. and Lovell, C.A.K.** *Stochastic Frontier Analysis*, Cambridge University Press, New York, 2000.
- Saaty, T.L.** *The Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*, RWS Publication, Pittsburgh, 1994.
- Simar, L. and Wilson, P.W.** (2007), Estimation and inference in two-stage, semi-parametric models of production processes, *Journal of Econometrics*, 136 (1), 31-64.
- Simar, L. and Wilson, P.W.** Performance of the bootstrap for DEA estimators and iterating the principle, in Cooper, W.W., Seiford, L.M. and Zhu, J. (Eds.), *Handbook on Data Envelopment Analysis*, Kluwer International Series, Boston, 265-298, 2004.
- Souza, G.S. and Staub, R.B.** (2007), Two-stage inference using data envelopment analysis efficiency measurements in univariate production models, *International Transactions in Operational Research*, 14, 245-258.
- Souza, G.S., Gomes, E.G. and Staub, R.B.** (2009a), Influence of contextual variables: an application to agricultural research evaluation in Brazil, *Proceedings International Data Envelopment Analysis Symposium*.
- Souza, G.S., Moreira, T.B. and Gomes, E.G.** (2009b), An efficiency approach for analyzing the major agricultural economies, *Proceedings International Data Envelopment Analysis Symposium*.
- Torgerson, W.S.** *Theory and Methods of Scaling*. Wiley, New York, 1958.