1996年度日本オペレーションズ・リサーチ学会
秋季研究発表会

Data Variations in DEA

01302170 埼玉大学 刀根 薫 TONE Kaoru

Introduction

Data in DEA are not always deterministic but suffer from several factors of disturbance, e.g. errors in data gathering, round-off errors, and statistical noise. Thus, data are a chance realization of the hidden true values and may fluctuate. Since DEA is a data oriented methodology, study of this area of uncertainty is a necessity. Specifically, the following two questions arise:

1. If a DMU is judged *efficient* by a certain DEA model, how *robust* is it with respect to a change in an input/output value?

2. If input/output values suffer from disturbance, the efficiency score may vary accordingly. Then, what is the probability that a DMU remains *efficient*, or what is the distribution of efficiency score for efficient or nearly efficient DMUs?

The first question concerns the *sensitivity* analysis of DEA scores and the latter the *stochastic* analysis.

A string of DEA papers dealt with these subjects, e.g. Charnes and Neralic (1990), Charnes, Cooper and Zlobec (1990), Charnes et al. (1992), Thompson et al. (1994) for the first subject and Sengupta (1987), Morita et al. (1994), Wilson (1994) for the second, among others.

1. Sensitivity Analysis in DEA

In this section, the following two subjects will be dealt with:

1. Suppose a DMU, say DMU_o, is efficient by the DEA model employed. What is the upper bound of increase in a certain input value of DMU_o for preserving efficiency?

2. What is the lower bound of decrease in a certain output value of DMU_o for preserving efficiency?

We will solve these problems using a bi-section method. Although we describe the method for the CCR model, we can also develop similar procedures for other DEA models.

1.1 The Upper Bound of Increase in an Input Value

Let the objective input item of the sensitivity analysis be x_{io}. If we increase x_{io} and measure the efficiency $\theta(x_{io})$ of DMU_o as a function of x_{io}, then either there is an upper bound $\bar{x}_{io}$ beyond which DMU_o is no longer efficient, or there is no bound, i.e. $\theta(x_{io}) = 1$ for every $x_{io} \geq 0$. We can solve these problems using a bisection method as described below. In the algorithm, we use a very small number ε, e.g. $\varepsilon = 1/1000$ and a large number M, e.g. $M = 10^5$.

Step 0. Set

$$\begin{aligned} switch &= 0 \\ x_{LB} &= x_{io} \text{ (the original data)} \\ x_{UB} &= 2x_{io} \\ x_{MID} &= x_{UB} \end{aligned}$$

Step 1.

If $|x_{UB} - x_{LB}| < \varepsilon$, then we have $\bar{x}_{io} = x_{MID}$. Stop.
If $|x_{UB} - x_{LB}| > M$, then there is no bound of increase in x_{io}. Stop.
Otherwise evaluate $\theta(x_{MID})$ by the CCR model.

Step 2

(i) If $switch = 0$ and $\theta(x_{MID}) = 1$, then set

$$x_D = x_{UB} - x_{LB}$$

$$
\begin{aligned}
x_{UB} &= x_{UB} + 2x_D \\
x_{LB} &= x_{UB} \\
x_{MID} &= x_{UB}.
\end{aligned}
$$

Return to Step 1.

(ii) If $\theta(x_{MID}) < 1$, then set

$$
\begin{aligned}
switch &= 1 \\
x_{UB} &= x_{MID} \\
x_{MID} &= (x_{UB} + x_{LB})/2.
\end{aligned}
$$

Return to Step 1.

(iii) If $switch = 1$ and $\theta(x_{MID}) = 1$, then set

$$
\begin{aligned}
x_{LB} &= x_{MID} \\
x_{MID} &= (x_{UB} + x_{LB})/2.
\end{aligned}
$$

Return to Step 1.

Once the *switch* is turned to 1, the interval $x_{UB} - x_{LB}$ that contains $\bar{x}_{io}$ will be reduced by $1/2^k$ after k iterations.

1.2 The Lower Bound of Decrease in an Output Value

In a similar setting to the above case, we can deal with the lower bound of decrease in an output value of an efficient DMU, using the bisection method.

1.3 An Example of Sensitivity Analysis

We will show an example at the presentation.

2. Stochastic DEA

In the previous section, we looked into the sensitivity analysis with respect to a specified input/output item of an efficient DMU and found the upper/lower bound of its variation for keeping the DMU efficient. In this section, we assume that all the data are subject to change according to some probability distributions; uniform, triangle, normal, lognormal, among others. Subsequently, the efficiency score is no more *deterministic* but *stochastic*. So, the subject in this section is called the *stochastic* DEA. We will introduce a simulation study for this purpose.

2.1 Stochastic DEA by Simulation

This simulation is done in the following way. The input/output data are sampled from the fitted probability distribution, usually expressed as

[original data] + [noise].

After sampling the noisy input/output data, we evaluate the efficiency of each DMU by an appropriate DEA model. We will repeat the above process for sufficiently many times, e.g. 1000, 5000 or 10000.

2.2 An Example of Stochastic DEA

We will show an example at the presentation.

References

[1] Charnes, Cooper and Zlobec (1990) in *Parametric Optimization and Related Topics II*, Berlin: Akademic-Verlag.

[2] Charnes, Haag, Jaska and Semple (1992), *International Journal of System Sciences*, 23(5):789-798.

[3] Charnes and Neralic (1990), *EJOR*, 48:332-341.

[4] Morita, Fujii and Kajihara (1994), in *Proceedings of APORS' 94*, World Scientific.

[5] Sengupta (1987), *Computers and Operations Research*, 14:117-169.

[6] Thompson, Dharmapala and Thrall (1994), in *Data Envelopment Analysis*, Kluwer Academic Publishers.

[7] Wilson (1994), working paper.