Mathematical and Computational Applications, Vol. 15, No. 1, pp. 148-155, 2010.

RANKING DECISION MAKING UNITS WITH STOCHASTIC DATA BY USING COEFFICIENT OF VARIATION

F. Hosseinzadeh Lotfi[1], N. Nematollahi[2],
M.H. Behzadi[3] and M. Mirbolouki[1]
[1]Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran
[2]Department of Statistics, Faculty of Economics, Allameh Tabataba'i University, Tehran, Iran
[3]Ph.D Student, Department of Statistics, Science and Research Branch, Islamic Azad University, 1477893855, Tehran, Iran, behzadi@srbiau.ac.ir

Abstract- Data Envelopment Analysis (DEA) is a non-parametric technique which is based on mathematical programming for evaluating the efficiency of a set of Decision Making Units (DMUs). Throughout applications, managers encounter with stochastic data and the necessity of having a method that is able to evaluate efficiency and rank efficient units has been under consideration. In this paper considering the concept of coefficient of variation among efficient DMUs, two ranking methods has been proposed. Within these ranking methods, a DMU will have a higher rank if it's coefficient of variation be smaller. These methods are suitable when managers are able to determine weights on coefficient of variations or on inputs and outputs. At the end we applied these methods on a numerical example.

Key Words- Coefficient of variation, Data envelopment analysis, Ranking

1. INTRODUCTION

In DEA, the efficiency is measured through the comparison process with the efficient frontier. The evaluation of efficiency measure, inorder to rank DMUs seems to be problematic since in a circumstances when efficient units are more than one there will be no further discrimination between these units based on their efficiency scores. This problem is more significant when the number of DMUs relative to the sum of inputs and outputs is small. Many researchers have proposed different methods inorder to rank efficient units [1] [2] [3]. Anderson and Peterson (A.P) [4] proposed a method for ranking efficient units on basis of the position of each eliminated efficient DMUs in relation to its corresponding new Production Possibility Set (PPS). Since in A.P model nonstability can occur, Mehrabian et al. [5] introduced another method (MAJ) for ranking efficient DMUs. This method does not suffer from nonstability but it would be infeasible in some cases.

Variety of DEA models have been formulated for performance evaluation and ranking DMUs in various fields with different data such as: deterministic, interval, fuzzy, etc. In different real world applications managers encounter with data that in which the uncertainty is inherent. In such circumstances the necessity of having a model which has the ability to rank and evaluate the efficiency of DMUs with stochastic data has been under consideration. A model in which DMUs have stochastic data have been defined by Cooper et al. [6] [7] and they have defined the stochastic efficient DMUs.

The efficiency by stochastic data has been analyzed by Sengupta [8]. In addition Morita et al. [9] have discussed DEA efficiency reliability and probability as being efficient but a suitable model for ranking such DMUs has not been proposed yet.

In this paper in accordance with the useful specifications of coefficient of variation, some indexes for ranking have been defined. These methods consider a multi criteria problem for minimizing coefficient of variation of input and outputs of efficient DMUs. In these methods we use weighting the coefficient of variation regarding the priority of each input and output coefficient of variation or weighting inputs and outputs by managers. Using numerical examples of car manufacturing industry, we will demonstrate how to use the result.

The remainder of paper is organized as follows: First the stochastic DEA models are introduced, Section 3 provides two methods for ranking DMUs with stochastic data. Section 4 and 5 provide an example and conclusion.

2. STOCHASTIC DEA MODELS

Let us assume that there exists n, DMUs and $\tilde{X}_j = (\tilde{x}_{1j},...,\tilde{x}_{mj})$ and $\tilde{Y}_j = (\tilde{y}_{1j},...,\tilde{y}_{sj})$ are random input and output vectors of DMU_j, $j = 1\ldots,n$. These components have been deemed to be normally distributed that $\tilde{x}_{ij}$: $N(\mu_{ij},\sigma_{ij}^2)$ and $\tilde{y}_{rj}$: $N(\mu'_{rj},\sigma'^2_{rj})$. Chance constrained version of output oriented stochastic CCR model is as follows:

$$
\begin{array}{lll}
\max & \phi & \\
s.t. & p\{\sum_{j=1}^{n}\tilde{y}_{rj}\lambda_j \geq \phi\tilde{y}_{ro}\} \geq 1-\alpha, & r = 1,...,s, \\
& p\{\sum_{j=1}^{n}\tilde{x}_{ij}\lambda_j \leq \tilde{x}_{io}\} \geq 1-\alpha, & i = 1,...,m, \\
& \lambda_j \geq 0, & j = 1,...,n.
\end{array} \tag{1}
$$

where in the above models, p means “probability” and $\alpha \in [0,1]$ is a level of error which is a predefined number. In accordance with the definitions and theorems which have been proposed in [7], the above model can be converted into the following deterministic model:

$$
\begin{aligned}
\max \quad & \varphi + \varepsilon\left(\sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^-\right) \\
s.t. \quad & \varphi\mu'_{ro} - \sum_{j=1}^{n} \mu'_{rj}\lambda_j + s_r^+ - \Phi^{-1}(\alpha)\sigma_r^o(\varphi,\lambda) = 0, \quad && r = 1,\ldots,s, \\
& \sum_{j=1}^{n} \mu_{ij}\lambda_j + s_i^- - \Phi^{-1}(\alpha)\sigma_i^I(\lambda) = \mu_{io}, && i = 1,\ldots,m, \\
& s_i^- \geq 0, \quad s_r^+ \geq 0, && i = 1,\ldots,m, \quad r = 1,\ldots,s, \\
& \lambda_j \geq 0, && j = 1,\ldots,n.
\end{aligned}
\tag{2}
$$

where:

$$(\sigma_r^o(\phi,\lambda))^2 = \sum_{j\neq o}\sum_{k\neq o}\lambda_j\lambda_k cov(\tilde{y}_{rj},\tilde{y}_{rk}) + 2(\lambda_o - \phi)\sum_{j\neq o}\lambda_j cov(\tilde{y}_{rj},\tilde{y}_{ro}) + (\lambda_o - \phi)^2 var(\tilde{y}_{ro}),$$

and

$$(\sigma_i^I(\lambda))^2 = \sum_{j\neq o}\sum_{k\neq o}\lambda_j\lambda_k cov(\tilde{x}_{ij},\tilde{x}_{ik}) + 2(\lambda_o - 1)\sum_{j\neq o}\lambda_j cov(\tilde{x}_{ij},\tilde{x}_{io}) + (\lambda_o - 1)^2 var(\tilde{x}_{io}).$$

Here, Φ is the cumulative distribution function of the standard normal distribution and $\Phi^{-1}(\alpha)$, is its inverse in level of α. The above model is nonlinear programming which can be converted into a quadratic programming model. Cooper et al. [7] have proposed a stochastic efficient DMU_o as follows:

DEFINITION- DMU_o is stochastic efficient if and only if the following conditions are both satisfied:

(i) $\phi^* = 1$,

(ii) Slack values are all zeros for all optimal solution.

3. RANKING DMUS BY USING COEFFICIENT OF VARIATION

Variance and standard deviation which have been used inorder to dispersal explanation, greatly depend on measurement unit in population. For comparing dispersion between populations, analysts should use indexes which are not related to measurement unit. One of these indexes is coefficient of variation, which is defined as $c = \sigma/\mu$ where σ and μ are standard deviation and mean of population, respectively.

One of the application of coefficient of variation is when an identical property in different population, is under measurement but the magnitude of observations differ considerably. For instance, inorder to measure dispersal comparison among profit and loss between high and low technique industries, there will be no alternative at hand but coefficient of variation. Note that in a population when we have a lower coefficient of variation, there will be more stability in performance.

Let E be the set of all stochastic efficient DMUs and $C_j = (c_{1j}, c_{2j}, ..., c_{mj})$ and $C_j' = (c'_{1j}, c'_{2j}, ..., c'_{sj})$ be the input and output coefficient of variation respectively for every $j \in E$, in a way that:

$$c_{ij} = \frac{\sigma_{ij}}{\mu_{ij}} \quad , \quad c'_{rj} = \frac{\sigma'_{rj}}{\mu'_{rj}}$$

3.1. Method 1- Weighting coefficient of variations

Considering the concept of coefficient of variation among efficient DMUs, a DMU which its corresponding input and output coefficients of variation are minimum is more desirable and it will have a better rank. Therefore among the efficient DMUs, a DMU will have a better rank if it is the optimal solution of the following problem:

$$\min_{j \in E} \{c_{1j}, c_{2j}, ..., c_{mj}, c'_{1j}, c'_{2j}, ..., c'_{sj}\} \tag{3}$$

The above problem is multi criteria problem and for solving such a problem we consider the significance of the magnitude of input and output coefficient of variation, the specific weights can be assigned to these coefficients by managers. For instance in an industrial project the magnitude of coefficient of variation of manpower may be of less importance or the coefficient of variation of consuming material and manufacturing apparatuses be of different importance. Regarding this topic, different weights can be considered for coefficient of variation. Problem (4) can be converted into the following problem:

$$\min_{j \in E} \{\bar{c}_j = \sum_{i=1}^{m} v_i c_{ij} + \sum_{r=1}^{s} u_r c'_{rj}\} \tag{4}$$

where $v_i, i = 1, ..., m$ and $u_r, r = 1, ..., s$ are definite weights. We propose $\bar{c}_j$ as the ranking index for efficient DMUs.

3.2. Method 2- Weighting inputs and outputs

Let us assume that D_j be the indicator of inputs and outputs vector when $j \in E$, i.e:

$$D_j = (\tilde{x}_{1j}, \tilde{x}_{2j}, ..., \tilde{x}_{mj}, \tilde{y}_{1j}, \tilde{y}_{2j}, ..., \tilde{y}_{sj})^t.$$

Also $W \in R^{m+s}$ is the vector of weights which indicates the importance of vector of inputs and outputs. Let $Z_j = W^t D_j, j \in E$. Thus Z_j is a stochastic variable which has normal distribution. The coefficient of variation of this variable is a sign of the total coefficient of variation of DMU_j. Therefore we have:

$$\mu''_j = E(Z_j) = W^t (\mu_{1j}, \mu_{2j}, ..., \mu_{mj}, \mu'_{1j}, \mu'_{2j}, ..., \mu'_{sj})^t,$$

$$\sigma''^2_j = Var(Z_j) = W^t \Sigma W.$$

where Σ is the variance-covariance matrix of D_j. Therefore considering above relations, coefficient of variation of Z_j is as follows:

$$CV_j = \frac{\sigma''_j}{\mu''_j}, \quad j \in E. \tag{5}$$

Since CV_j are reflecting the extent of constancy in performance of DMU_j, therefore, the less CV is, more constancy of performance will be and in different settings with more confidence this unit can be considered efficient. Therefore in this method CV_j is the index for ranking efficient DMUs. The less this score is, the higher rank of DMU_j is.

4. AN APPLICATION OF RANKING

A car manufacturing company wants to investigate into the specific car through their factories in different countries and also, to rank the efficient units. For this purpose company consider “production expenses” and “service expenses” as inputs and “acceleration from 0-100” and “the maximum horse power rotation per minute” as outputs. In the opinion of managers each of the inputs and outputs has different significance and priority, where these priorities are predefined from the managers. All the data have normal distribution and they are gathered in tables (1) through (4). These data have been obtained by sampling through ten succeeding temporal period. Also, the results of applying the aforementioned methods in section 3 with two priority vectors are indicated in tables (5) and (6).

Table 1. input 1

X_{1j}	period1	period2	period3	period4	period5	period6	period7	period8	period9	period10
x11	10037.7	9975.7	10013.2	10032.6	9997.5	10010.1	9973.5	10030.3	9993.8	10045.6
x12	9997.4	9964.5	9975.2	10039.5	9977.8	9898.5	9970.8	9987.4	9940	9999.7
x13	10009.4	9991.1	10011.3	10034.7	10044.7	10000.6	10035	10016.5	10005.7	10030.4
x14	10032.8	10006.2	10047.2	9990.3	10026.6	10035.6	10021	10050.8	10058.7	10029.6
x15	9964	10009.1	9965.3	9938.1	9989.8	9975.6	9970.5	9979.2	9970.2	10053.9
x16	9983.7	10008.3	10050.6	10055.8	9994.1	9948.3	10053.3	10075.4	10081.2	9921.8
x17	10016.5	10050.7	10030.9	10003.9	10029.9	9998.4	10042.3	10022.7	10044	10016.3
x18	10016.8	10027.2	10038.6	10004.6	10025.2	10032.6	9984.7	10047.8	10017	10021.5
x19	9986.6	9989.9	10015	9966.2	10059	9954.6	9976.1	9959.8	9960.5	10000.7
x1,10	9996.7	10021.8	10023.6	10026.3	10025.8	10009.8	10034.4	9975	10031	9999.5

Table 2. input 2

X_{2j}	period1	period2	period3	period4	period5	period6	period7	period8	period9	period10
x21	9.2637	9.0457	8.1849	8.9956	8.5914	9.4166	10.0153	8.5772	9.0778	8.8398
x22	9.0915	9.8106	10.1168	9.3509	9.395	8.2743	9.4816	9.7766	8.973	9.7956
x23	10.3774	9.9203	10.1771	10.1969	10.2716	9.964	10.3591	9.9601	10.1888	10.0459
x24	10.3393	10.5793	10.8264	10.1375	10.4783	10.9337	10.9033	10.554	10.3782	10.2801
x25	7.49104	7.76119	8.41306	9.49551	9.3203	8.5424	7.83482	8.05882	9.74838	8.40874
x26	9.2632	8.8461	8.7745	9.5382	10.1447	7.6051	7.9171	9.1831	9.6143	9.7307
x27	10.1729	10.5935	10.5292	10.3275	10.6438	10.1919	10.3315	10.4361	10.3828	10.3722
x28	10.5844	10.5381	10.4168	10.23	10.1652	10.1414	9.8771	10.3465	10.4199	10.5026
x29	8.764	10.0621	7.7721	9.3339	7.8466	9.8717	8.9566	10.3985	9.9548	9.5601
x2,10	10.7033	10.5632	10.7276	10.5324	10.7364	10.7317	10.6693	10.8936	10.7664	10.6915

Table 3. output 1

Y_{1j}	period1	period2	period3	period4	period5	period6	period7	period8	period9	period10
y1,1	11.3021	11.3748	11.2201	11.144	11.3548	11.5313	11.6184	12.0096	11.1365	11.5129
y1,2	11.1543	11.6795	11.9251	11.443	10.572	9.9516	11.0554	11.6893	11.11	11.214
y1,3	9.8729	10.5942	9.1735	10.3173	9.9395	10.3133	9.2206	9.6419	9.7899	10.2297
y1,4	9.6855	9.6424	9.2758	9.6589	9.8589	10.6956	9.1564	9.7087	8.8202	9.0603
y1,5	11.2969	11.8646	11.6855	11.7956	11.5162	12.2677	11.7501	12.2583	12.4484	12.4793
y1,6	11.4075	11.2272	12.1297	11.7059	11.9491	10.7108	11.4051	10.6782	11.8115	11.2247
y1,7	9.8929	9.9607	9.8242	10.2438	10.1394	9.9547	9.6843	9.6536	10.0772	10.2146
y1,8	9.4471	10.2223	9.5851	10.1988	8.8077	10.0335	9.5656	10.1492	9.6211	10.081
y1,9	10.2396	12.146	11.1536	11.2986	13.06	9.6666	11.5222	12.7845	11.2119	9.7935
y1,10	9.4981	10.0259	10.2198	10.1127	10.3899	10.0516	10.0384	9.8584	10.1073	10.1583

Table 4. output 2

Y_{2j}	period1	period2	period3	period4	period5	period6	period7	period8	period9	period10
y2,1	101.275	111.112	90.201	91.966	97.173	102.852	92.352	97.614	108.875	75.223
y2,2	109.991	76.782	98.868	106.237	113.742	100.53	110.048	109.818	99.717	111.298
y2,3	96.543	99.523	93.344	92.3	98.721	103.952	97.838	100.284	106.861	97.419
y2,4	96.817	103.021	100.573	98.677	101.011	95.643	100.901	93.696	94.981	96.382
y2,5	104.824	100.503	105.669	108.393	105.082	96.15	113.406	98.767	115.951	101.475
y2,6	99.248	108.436	103.164	107.006	101.37	105.586	99.725	109.767	111.135	110.741
y2,7	101.609	99.111	95.57	100.26	103.218	93.401	100.692	93.687	101.054	106.32
y2,8	97.019	102.656	101.899	103.57	96.333	102.167	99.911	101.06	102.101	97.597
y2,9	99.945	88.252	100.557	114.821	100.185	92.762	99.854	94.439	103.795	103.499
y2,10	96.947	99.849	96.071	99.328	104.453	96.836	94.87	103.566	98.573	91.013

Table 5. results of method 1

α=0.05				α=0.1					
Efficient DMU	**DMU1**	**DMU2**	**DMU6**	**Efficient DMU**	**DMU1**	**DMU2**	**DMU5**	**DMU6**	**DMU10**
$\bar{C}(1)$	0.0449	0.0555	0.0378	$\bar{C}(1)$	0.0449	0.0555	0.0391	0.0378	0.0218
rank	2	3	1	**rank**	4	5	3	2	1
$\bar{C}(2)$	0.179	0.2055	0.1701	$\bar{C}(2)$	0.179	0.2055	0.1784	0.1701	0.0716
rank	2	3	1	**rank**	4	5	3	1	2

Table 6. results of method 2

α=0.05				α=0.1					
Efficient DMU	**DMU1**	**DMU2**	**DMU6**	**Efficient DMU**	**DMU1**	**DMU2**	**DMU5**	**DMU6**	**DMU10**
$CV(1)$	0.002	0.004	0.0052	$CV(1)$	0.002	0.004	0.0028	0.0052	0.0018
rank	1	2	3	rank	2	4	3	5	1
$CV(2)$	0.162	0.3223	0.4147	$CV(2)$	0.162	0.3223	0.2215	0.4147	0.1423
rank	1	2	3	rank	2	4	3	5	1

Note that in Table (5), $\bar{c}_j(1)$ and $\bar{c}_j(2)$ are the sum of the aggregated coefficient of variation which are gained from expression (4) with weight vectors $(u,v)_1 = (0.2,0.1,0.4,0.3)$ and $(u,v)_2 = (1,1,1,1)$, respectively. In Table (6), $CV_j(1)$ and $CV_j(2)$ are gained from expression (5) with weight vectors $W_1 = (0.25,0.05,0.45,0.25)$ and $W_2 = (1,1,1,1)$, respectively.

5. CONCLUSION

Through applications, number of efficient units are more than one. Therefore managers want to find a suitable and analytical method inorder to rank efficient units. In many setting we encounter with stochastic data. Therefore in this paper two different methods for ranking efficient units with stochastic data are proposed. In these methods according to the useful characteristic of coefficient of variation, some indexes for ranking have been defined. These methods are applicable for situations in which some of the input and output coefficient of variation are significant for managers. For instance, when lifetime of a part, product quality, consumer satisfaction and etc. are important. For purpose of illustrations the aforesaid methods in car manufacturing industry are considered. The mentioned methods are based on weighting the coefficient of variation as the criteria of the problem. This weighting method is one of the methods for solving multi criteria problem. Other methods can be suggested for further investigations.

6. REFERENCES

1. F. Liu, and H. H. Peng, Ranking of units on DEA frontier with common weights, *Computer and Operation Research* **35**, 1524-1537, 2008.
2. T. Obata, and H. Ishii, A method of discriminating efficient candidates with ranked voting data, *European Journal of Operational Research* **131**, 233-237, 2008.
3. A. Charnes, W. W. Cooper, and E. Rhodes,Measuring the efficiency of decision making units, *European Journal of Operational Research* **2**, 429-444, 1978.
4. P. Andersen, and N. C. Petersen, A procedure for ranking efficient units in data envelopment analysis, *Management Science* **39**, 1261-1264, 1993.
5. S. Mehrabian, M. R. Alirazaei, and GR. Jahanshahloo, A complete efficiency ranking of decision making units: an application to the teacher training university, *Computational Optimization and Application* **14**, 1998.
6. W.W. Cooper, Z. Huang, and S. Li, Satisfying DEA models under chance constraints, *The Annals of Operations Research* **66**, 279-295, 1996.
7. W.W. Cooper, H. Deng, Z. Huang, and S. X. Li , Chance constrained programming approaches to congestion in stochastic data envelopment analysis, *European Journal of Operational Research* **155**, 487-501 155, 2004.
8. JK. Sengupta, Efficiency analysis by stochastic data envelopment analysis, *Applied Economics Letters* **7**, 379-383, 2000.
9. H. Morita, and L. M. Seiford, Characteristics on stochastic DEA efficiency-reliability and probability being efficient, *Journal of Operations Research Society of Japan* **42**, 389-404, 1999.