

確率的効率性に関する一考察

神戸大学工学部 森田 浩 (Hiroshi Morita)

1 はじめに

DEA(Data envelopment analysis)は多入力多出力システムの相対的効率性を評価するための有用なノンパラメトリック手法として広く用いられるようになったが、確率的なデータは扱えるようにはできていない。実際に解析に用いられるデータには測定誤差やノイズなどが含まれていると考えられることから、確率的 DEA 法の開発の重要性が言われている(例えば Seiford(1996))。確率的なデータによる効率性分析の研究として、Olesen and Petersen(1995)は機会制約条件付き DEA モデルを提案し、機会制約条件を与えたときの効率性尺度を示した。このモデルは非凸計画問題に定式化されるため、得られた効率性尺度もその上下限値を与えるに止まっている。Sengupta(1990)も1出力の場合に機会制約モデルを示している。また、統計的な手法では Banker(1993)の最尤推定量や Simar(1992)のブーツストラップ法などがある。

本稿では観測データに測定誤差が含まれているとき、その確率的変動を考慮したときの効率性について考察し、確率的変動に対する新たな効率性尺度を提案する。確率的なデータを扱うときに最も興味あることは、得られた効率値が確率的変動に対して如何にロバストであるかとかどの程度信頼できるものであるかであろう。効率性評価の信頼性やある確率レベルにおける最小効率値を提案し、さらに期待効率値や効率的である確率、効率値の $\alpha\%$ 点などの確率的尺度についても考察する。

n 個のユニット(DMUと呼ぶ)の相対的効率性を考える。各々のDMUは m 個の入力と s 個の出力を持ち、DMU $_j$ は入力 $X_j = \{x_{ij}\}(i = 1, \dots, m)$ から出力 $Y_j = \{y_{rj}\}(r = 1, \dots, s)$ を産出しているものとする。ここでは主としてCCR入力指向モデルを考える。通常のDEAモデルでは確率的変動は無視して、次の確定的な線形計画問題により効率性を決定している。

$$\text{CCR-D} \left\{ \begin{array}{ll} \max_{u,v} & w_o = v'Y_o, \\ \text{subject to} & u'X_o = 1, \\ & v'Y - u'X \leq 0, \\ & u, v \geq \epsilon, \end{array} \right. \quad (1)$$

これはDMU $_o$ の効率性を評価するための計画問題であり、すべてのDMUの効率性を求めるには $(X_o, Y_o) = (X_j, Y_j), j = 1, \dots, n$ としながら n 回解く必要がある。目的関数の値 w_o^* により、 $w_o^* = 1$ となるDMUは効率的、 $w_o^* < 1$ となるDMUは非効率的と呼ばれる。

2 DEAにおける確率的データ

すべてのデータには潜在的に不確実性が存在しており、最も典型的な確率的データは加法的観測誤差を含んだデータであろう。このとき繰り返し測定は誤差の推定精度を上げる

ために取られる方法である。データにはある一定期間にわたって観測することがある。例えば、1週間とか1ヶ月当たりの観測値であるが、これらは複数の観測値の和として表され、確率的変動の影響を少なくするために行われている。

ここで分散成分の推定方法について整理してみよう。DMU_jが N_j 回の繰り返し観測値を持っているとして、それらを $x_j^k, k=1, \dots, N_j$ と表す。 x_j^k は平均 μ_j と分散 σ_j^2 をもつ正規母集団からの標本と仮定し、分散 σ_j^2 の推定法を母集団の仮定により以下のように分類する。

仮定なし： $\sigma_j^2, j=1, \dots, n$

各々の分散はそれぞれ独立に推定されることになり、分散の推定量 $\hat{\sigma}_j^2$ は

$$\hat{\sigma}_j^2 = \frac{1}{N_j - 1} \sum_{k=1}^{N_j} (x_j^k - \bar{x}_j)^2 \quad (2)$$

で与えられる。ただし、 $\bar{x}_j$ は標本平均である。この場合十分な推定精度を得るにはかなりの繰り返しが必要となる。

等分散：すべての j に対して、 $\sigma_j^2 = \sigma^2$

すべてのDMUが同じ分散 σ^2 をもっていると仮定する。つまり、観測誤差はDMUには依らず、入力あるいは出力項目に固有のものであると考える。推定量 $\hat{\sigma}^2$ は(2)をプーリングして

$$\hat{\sigma}^2 = \frac{\sum_{j=1}^n (N_j - 1) \hat{\sigma}_j^2}{\sum_{j=1}^n N_j - n} \quad (3)$$

と与えられる。これは繰り返しのある一元配置分散分析と同じである。

等変動係数：すべての j に対して、 $\sigma_j = K\mu_j$

すべてのDMUが同じ変動係数 K をもっていると仮定する。つまり、分散の大きさは絶対量に比例するというもので、大きい入出力値ほど大きい誤差を含んでいると考える。対数変換をした $\hat{x}_j^k = \log(x_j^k)$ は近似的に等分散¹となることから、 $\hat{x}_j^k$ を(3)に適用して変動係数の推定値を求めることができる。変数変換法を使うと分散の他の仮定に対しても適用できる²。分散の推定には最尤推定法も有効である。

3 DEAにおける確率的変動

確率的変動があるとき、どの段階で確率的変動を効率性評価に組み入れるかで2つの立場が存在する。一つは確率的データを直接用いて効率値を計算するもので、確率的な効率

¹ $V(x_j) = \sigma_j^2, \sigma_j = K\mu_j$ とする。 $f(x_j)$ を $x_j = \mu_j$ の周りでテイラー展開すると、 $f(x_j) = f(\mu_j) + f'(\mu_j)(x_j - \mu_j) + \dots$ となり、近似的に期待値 $E[f(x_j)] = f(\mu_j)$ と分散 $V[f(x_j)] = \{f'(\mu_j)\}^2 \sigma_j^2$ が求められる。 $V[f(x_j)]$ が j に依らないようにするには、関数 f を $f'(\mu_j)\sigma_j$ が定数となるように選ばねばよいので、微分方程式 $f'(x)x = 1$ を解くと $f(x) = \log(x)$ が得られる。よって、 $V[\log(x_j)] = K^2$ が得られ、 j に依らず一定となる。

² $\sigma_j^2 = g(\mu_j)$ と仮定する。等分散となるように変数変換するには、関数 f を $\{f'(x)\}^2 g(x) = 1$ となるように選ばねばよいので、 $f(x) = \int \frac{1}{\sqrt{g(x)}} dx$ が得られる。例えば、 $\sigma_j^2 = K\mu_j$ ならば、 $f(x) = 2\sqrt{x}$ とすると $V[f(x)] = K$ となる。

性尺度が導かれる。もう一つは確定的な情報から効率値を求めた後で、確率的変動の情報を用いて効率的 DMU のロバスト性を評価するもので、感度分析や信頼性分析につながるものである。

3.1 確率的効率性分析

入出力データ $X(\omega)$ と $Y(\omega)$ が確率変数として与えられるとき、問題 (1) は次の確率的計画問題となる。

$$\text{CCR}(\omega)\text{-D} \quad \left\{ \begin{array}{ll} \max_{u,v} & w_o = v'Y_o(\omega), \\ \text{subject to} & u'X_o(\omega) = 1, \\ & v'Y(\omega) - u'X(\omega) \leq 0, \\ & u, v \geq \epsilon. \end{array} \right. \quad (4)$$

目的関数の値は確率変数となり、効率的か非効率的かの判定も確率的となる。このとき、効率値の期待値と分散、効率的である確率、効率値の $\alpha\%$ 点などの確率的効率性の尺度が考えられる。

$X(\omega)$ と $Y(\omega)$ は加法的な正規雑音をもつデータとしているが、上の確率的計画問題を解析的に解くことは非常に難しく、以下に示すモンテカルロ法によるシミュレーションが確率的効率性の尺度を求めるには有効な手法となる。

1. データから推定された分布に従う乱数を発生させ、人工的に入出力データを作る。
2. 問題 (1) を解いて効率値を計算する。
3. 上の操作を十分な回数繰り返す。

N 回の繰り返しで得られた効率値を $\theta_1, \dots, \theta_N$ とする。このとき、確率的効率性尺度の推定値は以下のように求められる。

効率値の期待値と分散： $\bar{\theta} = \frac{1}{N} \sum_{t=1}^N \theta_t$ と $s_{\theta}^2 = \frac{1}{N-1} \sum_{t=1}^N (\theta_t - \bar{\theta})^2$

効率値の $\alpha\%$ 点： $\theta^{\alpha} = \theta_{(N\alpha)}$

ただし、 $\theta_{(i)}$ は i 番目に大きな効率値である。

効率的である確率： $p = \#(\theta_t = 1)/N$

二項分布を用いた統計的推定理論を用いることになるので、十分な精度で推定するには極めて大きな回数の繰り返しが必要となる。

3.2 信頼性分析

入出力データの確率的変動は効率的な DMU の感度分析や信頼性分析にも用いられる。通常の DEA モデルでは確率的変動の情報を使わないで効率的か非効率的かの評価がなされているが、確率的変動が存在するときには、効率性の評価が確率的変動に対してどの程度信頼できるものであるか、どんな確率で効率性が保証されるのか、どの入出力が確率的

変動に影響を受けやすいかなどが興味の対象となる。このとき確率的変動の情報はロバスト性や信頼性を評価するために用いられ、先に求められた効率性に付加的な評価を与えるものとなる。

効率的な DMU に対する信頼性の尺度としては、その効率性評価を変化させない範囲の確率的変動の取りうる最大の確率レベルを考える。つまり、その確率レベルで起こりうるどんな確率的変動も効率性の評価を変化させることはないという確率レベルの最大値を意味している。次章に信頼性分析の詳細を示すが、そこで得られる結果から各入出力の感度も分析できる。

4 効率的な DMU の信頼性

観測データが平均 0 と有限分散をもつ多変量正規分布に従う確率的変動 $\delta = (\delta_x, \delta_y)$ を持っているとは仮定する。ただし、 $\delta_x = (\delta_{x1}, \dots, \delta_{xm})$ は入力、 $\delta_y = (\delta_{y1}, \dots, \delta_{yr})$ は出力にそれぞれ対応させる。一定の確率レベルで起こりうる確率的変動を考えると、確率的変動が大きくなるにつれて、取りうる効率値の幅も大きくなり、効率的な DMU も非効率と評価される可能性がでてくる。効率的な DMU を非効率と評価させるような確率的変動の大きさは、効率的と評価された DMU の信頼性を反映しているといえよう。すなわち、大きな確率的変動に対しても効率的という評価が不変であれば、その DMU は信頼性が高いものといえる。

α を確率レベルとすると、DMU_o における確率的変動 δ の信頼領域は

$$S_\alpha = \{ \delta | \delta' \Sigma_o^{-1} \delta \leq \chi_{m+s-1}^2(\alpha) \} \quad (5)$$

となる。ここで、 Σ_o は DMU_o の確率的変動のもつ分散共分散行列であり、 $\chi_{m+s}^2(\alpha)$ は自由度 $m+s$ の χ^2 分布の $\alpha\%$ 点である。入出力の確率的変動 δ は確率 α で (5) の中の値を取ることになり、DMU_o の効率値は δ の実現値の関数となる。

DMU_o が確率的変動 $\delta = (\delta_x, \delta_y)$ を取ったとすると、観測データは $(X_o + \delta_x, Y_o - \delta_y)$ となり、効率値は次の計画問題によって与えられる。

$$\left| \begin{array}{ll} \max_{u,v} & \tilde{w}_o(\delta) = v'(Y_o - \delta_y) \\ \text{subject to} & u'(X_o + \delta_x) = 1, \\ & v'Y - u'X \leq 0, \\ & u, v \geq \epsilon. \end{array} \right. \quad (6)$$

$\tilde{w}_o^*(\delta)$ を δ が与えられたときの効率値とする。ここで、確率レベル α での確率的変動を考えると、次の計画問題で与えられる効率値の最小値は確率レベル α における最小の効率値を保証するものである。

$$\left| \begin{array}{ll} \min_{\delta} & w_o(\alpha) = \tilde{w}_o^*(\delta) \\ \text{subject to} & \delta' \Sigma_o^{-1} \delta \leq \chi_{m+s-1}^2(\alpha), \end{array} \right. \quad (7)$$

または

$$\text{MIN}(\alpha)\text{-D} \left| \begin{array}{l} \min_{\delta} \max_{u,v} \quad w_o(\alpha) = v'(Y_o - \delta_y) \\ \text{subject to} \quad u'(X_o + \delta_x) = 1, \\ \quad \quad \quad v'Y - u'X \leq 0, \\ \quad \quad \quad u, v \geq \epsilon \\ \quad \quad \quad \delta' \Sigma_o^{-1} \delta \leq \chi_{m+s-1}^2(\alpha). \end{array} \right. \quad (8)$$

問題 (8) の最適値を $w_o^*(\alpha)$ と表す。しかしながら、この問題はミニマックス型で非凸の制約条件を持つ計画問題であるから、このままでは大域的な最適解を求めることは困難である。 δ を固定し、問題 (8) の双対問題を作ると

$$\text{MIN}(\alpha)\text{-P} \left| \begin{array}{l} \min_{\delta, \theta, \lambda, s^+, s^-} \quad z_o(\alpha) = \theta - \epsilon' s^+ - \epsilon' s^- \\ \text{subject to} \quad Y \lambda - s^+ = Y_o - \delta_y, \\ \quad \quad \quad X \lambda + s^- = \theta(X_o + \delta_x), \\ \quad \quad \quad \lambda, s^+, s^- \geq 0, \\ \quad \quad \quad \delta' \Sigma_o^{-1} \delta \leq \chi_{m+s}^2(\alpha). \end{array} \right. \quad (9)$$

が得られる。これはミニマックス型ではないが、まだ非凸計画問題である。

$w_o^*(\alpha)$ は確率レベル α における最小の効率値である。すなわち、確率レベル α での信頼領域内のどんな確率的変動でもその効率値を $w_o^*(\alpha)$ 以下にすることはできない。つまり $w_o^*(\alpha) = 1$ のとき、DMU_o は (5) の中の任意の δ に対して効率的となる。 $w_o^*(\alpha) = 1$ であるような α の最大値を $\alpha_{\max}$ と表す。

$$\alpha_{\max} = \{\alpha | w_o^*(\alpha) = 1\}. \quad (10)$$

DMU_o は確率レベル $\alpha_{\max}$ での信頼領域内の任意の確率的変動に対して効率的となることが保証されるが、確率レベル $\alpha_{\max}$ 以上では非効率となることもある。この確率 $\alpha_{\max}$ を DMU_o の信頼性 (reliability) と呼ぶことにする。

信頼性 $\alpha_{\max}$ を求めるために、Andersen and Petersen (1993) の拡張効率値を用いる。

$$\text{SE} \left| \begin{array}{l} \min_{\theta, \lambda} \quad \hat{z}_o = \theta \\ \text{subject to} \quad Y_{-o} \lambda \geq Y_o, \\ \quad \quad \quad X_{-o} \lambda \leq \theta X_o, \\ \quad \quad \quad \lambda \geq 0, \end{array} \right. \quad (11)$$

ここで、 $X_{-o} = (X_1, \dots, X_{o-1}, X_{o+1}, \dots, X_n)$ と $Y_{-o} = (Y_1, \dots, Y_{o-1}, Y_{o+1}, \dots, Y_n)$ とする。効率のフロンティアは DMU_o 以外で構成されるため効率値 $\hat{z}_o$ は 1 以下とはならず、効率的な DMU は 1 以上の効率値を取るが、非効率的な DMU は CCR モデルと同一の効率値を取る。この拡張効率値モデル (9) における確率的変動を考える。

$$\text{SE}(\alpha) \left| \begin{array}{l} \min_{\theta, \lambda, \delta} \quad \hat{z}_o(\alpha) = \theta \\ \text{subject to} \quad Y_{-o} \lambda \geq Y_o - \delta_y, \\ \quad \quad \quad X_{-o} \lambda \leq \theta(X_o + \delta_x), \\ \quad \quad \quad \lambda \geq 0, \\ \quad \quad \quad \delta' \Sigma_o^{-1} \delta \leq \chi_{m+s}^2(\alpha). \end{array} \right. \quad (12)$$

問題 (12) は確率レベル α における最小の拡張効率値を与える。この問題 (12) も凸計画ではない。 $\hat{z}_o^*(\alpha)$ を問題 (12) の最適値とすると、

$$\alpha < \alpha_{\max} \text{ のとき } z_o^*(\alpha) = 1, \hat{z}_o^*(\alpha) > 1 \quad (13)$$

$$\alpha = \alpha_{\max} \text{ のとき } z_o^*(\alpha) = \hat{z}_o^*(\alpha) = 1 \quad (14)$$

$$\alpha > \alpha_{\max} \text{ のとき } z_o^*(\alpha) = \hat{z}_o^*(\alpha) < 1 \quad (15)$$

が得られる。もし $\theta = 1$ が問題 (12) の実行可能解であれば、 $\alpha \geq \alpha_{\max}$ となるから、 $\alpha_{\max}$ は $\theta = 1$ のときの α の最小値として求めることができる。さらに、 $\chi_{m+s}^2(\alpha) = \delta' \Sigma_o^{-1} \delta$ が成立し、 $\delta' \Sigma_o^{-1} \delta$ が最小となるのは α が最小のときであるから、次の 2 次計画問題によって $\alpha_{\max}$ を求められる。

$$\begin{array}{ll} \min_{\lambda, \delta} & \Delta = \delta' \Sigma_o^{-1} \delta, \\ \text{subject to} & Y_{-o} \lambda \geq Y_o - \delta_y, \\ & X_{-o} \lambda \leq X_o + \delta_x, \\ & \lambda \geq 0. \end{array} \quad (16)$$

問題 (16) の最適値を Δ^* と表すと、信頼性 $\alpha_{\max}$ は

$$\chi_{m+s}^2(\alpha_{\max}) = \Delta^*. \quad (17)$$

から得られる。 $\sqrt{\Delta^*}$ はデータ点 $(X_o + \delta_x, Y_o - \delta_y)$ から包絡面への確率的距離を示している。

Δ^* を $\Delta^* = d'd$ によって各入出力に分解すると、確率距離ベクトル d は各入出力の確率的変動に対する感度もわかる。

$$d = \Sigma_o^{-\frac{1}{2}} \delta^*, \quad (18)$$

点 (X_o, Y_o) から確率的に最も近い包絡面上の点はベクトル d の方向にある。つまり、 $d_{x1} > d_{x2}$ ならば、入力-1 は入力-2 より確率的変動に対してより敏感となっていることがわかる。図 1 にベクトル δ^* と d を例示している。左の図の楕円は元のデータ空間に書かれたもので、 δ^* は点 (X_o, Y_o) から確率的に最も近い包絡面上の点へのベクトルを表している。右の図の円は分散行列によって変換された空間に書かれたもので、 d は点 (X_o, Y_o) から包絡面へのベクトルを表している。

ある確率レベルにおける最小効率値は問題 (12) を解けば求められるが、非凸であるため容易に求められない。問題 (16) を導いたように効率値 θ を固定すると、効率値 θ を保証する最大の確率レベル $\alpha(\theta)$ は次の問題を解くことによって求めることができる。

$$\begin{array}{ll} \min_{\lambda, \delta} & \Delta(\theta) = \delta' \Sigma_o^{-1} \delta, \\ \text{subject to} & Y_{-o} \lambda \geq Y_o - \delta_y, \\ & X_{-o} \lambda \leq \theta(X_o + \delta_x), \\ & \lambda \geq 0. \end{array} \quad (19)$$

確率 $\alpha(\theta)$ は $\chi_{m+s}^2(\alpha(\theta)) = \Delta^*(\theta)$ から θ の関数として求められる。いくつかの θ に対して $\alpha(\theta)$ を計算すれば、 (θ, α) -空間上に曲線を描くことができ、確率レベル α_0 における最小効率値は $\alpha(\theta) = \alpha_0$ となる θ によって求めることができる。

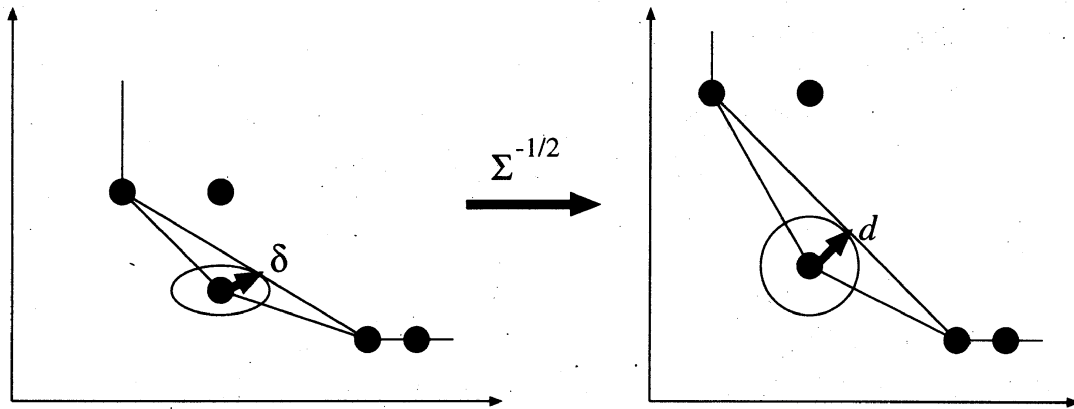


図 1: Stochastic distances

5 簡単な例題

2つの入力と1つの単位出力をもつ5つのDMUを考える。これらの入力値はそれぞれA(3,5), B(5,3), C(8,2), D(9,2), E(5,5)とする。DMU A, B, Cは効率的であり、DMU DとEは非効率的である。DMU Dは効率値1を取るが、入力1に正のスラックを持っている。今、各入力に確率的変動があるものとしてそれぞれの分散を $\sigma_{x_1}^2 = 0.8^2$, $\sigma_{x_2}^2 = 0.5^2$ とおき、DMUの信頼性や確率的効率性尺度を求めてみる。分散は互いに独立としている。

まず、DMU Bの評価をする。拡張効率値は $\hat{z}_B = 1.133$ である。DMU Bの信頼性は次の問題を解いて求められる。

$$P_B \left\{ \begin{array}{ll} \min_{\lambda, \delta} & \left(\frac{\delta_1}{0.8} \right)^2 + \left(\frac{\delta_2}{0.5} \right)^2 \\ \text{subject to} & 3\lambda_A + 8\lambda_C + 9\lambda_D + 5\lambda_E \leq 5 + \delta_1 \\ & 5\lambda_A + 2\lambda_C + 2\lambda_D + 5\lambda_E \leq 3 + \delta_2 \\ & \lambda_A + \lambda_C + \lambda_D + \lambda_E \geq 1 \\ & \lambda_A, \lambda_C, \lambda_D, \lambda_E \geq 0 \end{array} \right. \quad (20)$$

最適解は $\Delta^* = 1.3322$, $\delta_1^* = 0.639$, $\delta_2^* = 0.416$, $\lambda_A^* = 0.472$, $\lambda_C^* = 0.528$, $\lambda_D^* = \lambda_E^* = 0$ である。このとき、 $\chi_2^2(.4863) = 1.3322$ から、信頼性は $\alpha_{\max} = 48.6\%$ となる。また、各入力における確率的距離は $d_1 = 0.80$ と $d_2 = 0.83$ となる。 δ_1^* は δ_2^* より大きい、 δ_1^* の分散は δ_2^* の分散より大きいため、各入力に対する感度はほぼ同じとなっている。

次に、DMU Cの評価をする。拡張効率値は $\hat{z}_C = 1.0625$ である。DMU Cの信頼性は次の問題を解いて求められる。

$$P_C \left\{ \begin{array}{ll} \min_{\lambda, \delta} & \left(\frac{\delta_1}{0.8} \right)^2 + \left(\frac{\delta_2}{0.5} \right)^2 \\ \text{subject to} & 3\lambda_A + 5\lambda_B + 9\lambda_D + 5\lambda_E \leq 8 + \delta_1 \\ & 5\lambda_A + 3\lambda_B + 2\lambda_D + 5\lambda_E \leq 2 + \delta_2 \\ & \lambda_A + \lambda_B + \lambda_D + \lambda_E \geq 1 \\ & \lambda_A, \lambda_B, \lambda_D, \lambda_E \geq 0 \end{array} \right. \quad (21)$$

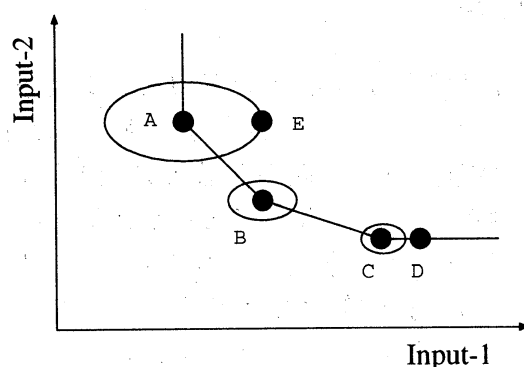


図 2: Ellipsoids of Reliability

表 1: Reliability of efficient DMUs

	DMU				
	A	B	C	D	E
拡張効率値: $\hat{z}$	1.667	1.133	1.063	1.000	0.800
確率的距離: $\sqrt{\Delta^*}$	2.500	1.154	0.464	0.000	—
δ_1^*	2.000	0.639	0.138	0.000	—
δ_2^*	0.000	0.416	0.216	0.000	—
d_1	2.500	0.799	0.173	0.000	—
d_2	0.000	0.832	0.432	0.000	—
信頼性: $\alpha_{\max}$	95.6%	48.6%	10.2%	0.0%	—
90%効率性	1	.908	.824	.751	.665
95%効率性	1	.884	.799	.731	.650

最適解は $\Delta^* = 0.2155$, $\delta_1^* = 0.138$, $\delta_2^* = 0.216$, $\lambda_B^* = 0.216$, $\lambda_D^* = 0.784$, $\lambda_A^* = \lambda_E^* = 0$ である。このとき、 $\chi_2^2(.1022) = 0.2155$ から、信頼性は $\alpha_{\max} = 10.2\%$ となる。また、各入力における確率的距離は $d_1 = 0.17$ と $d_2 = 0.43$ となる。入力 2 は入力 1 より確率的変動に対して影響を受けやすいことがわかる。図 2 は DMU A, B, C に対する信頼性を表す楕円を例示している。表 1 には分析結果をまとめて示している。

これらの結果から、DMU A は極めて信頼性が高い効率的な DMU であり、入力 2 は確率的変動に全く影響されないことがわかる。DMU D は信頼性の無い DMU であるが、これは入力 1 にスラックがあることから入力 2 のわずかな変動によって非効率とされてしまうからである。DMU C は信頼性は低く、特に入力 2 は確率的変動の影響を受けやすい。DMU B は程々の信頼性を持ち、各入力の感度もほぼ同じである。

6 効率的となる確率

本章では確率的変動のある場合に DMU が効率的となる確率を考え、信頼性との関係を明らかにする。第3章で述べたように、効率的となる確率には2種類の確率が考えられる。一つはすべての DMU の入出力データを確率変数と見たときに効率的となる確率である。これが本来の効率的となる確率といえるものであるが、解析的に求めるのは非常に困難であり、3.1 節に示したシミュレーション手法によって求めるのが有効であろう。しかし、効率的となる確率の推定値を十分な推定精度で求めるにはかなりの回数のシミュレーションが必要となる。

もう一つは評価対象の DMU のみの入出力データを確率変数と見、その他の DMU の入出力データは確定的と見たときに効率的となる確率である。この確率は評価対象の DMU が拡張フロンティアの外側の領域に存在する確率と等しい。拡張フロンティアの外側の領域を求める手法は Seiford and Zhu (1996) によって示されているが、領域が有界でないことから、この領域の大きさを比較して DMU のロバスト性を比較することは困難である。そこで、確率的変動の大きさを考えて、DMU の確率的変動に対する感度を数量的に比較することを可能にするために、この効率的となる確率が有効となる。

何れの確率も正確に計算するのは困難である。これらの確率の下限値の導出などに関する議論の詳細と数値例による検証は現在投稿中の Morita and Seiford を参照されたい。

参考文献

- Andersen, P. and N. C. Petersen, A Procedure for Ranking Efficient Units in Data Envelopment Analysis, *Management Science*, Vol. 39, pp. 1261–1264 (1993).
- Banker, R. D., Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation, *Management Science*, Vol. 39, pp. 1265–1273 (1993).
- Morita, H. and L. M. Seiford, Characteristics on Stochastic Efficiency – Reliability and Probability Being Efficient–, (to be submitted).
- Olesen, O. B. and N. C. Petersen, Chance Constrained Efficiency Evaluation, *Management Science*, Vol. 41, pp. 442–457 (1995).
- Seiford, L. M., Data Envelopment Analysis: The Evolution of the State of the Art (1978–1995), *Journal of Productivity Analysis*, Vol. 7, pp. 99–138 (1996).
- Seiford, L. M. and J. Zhu, Stability Regions for Maintaining Efficiency in Data Envelopment Analysis, *Discussion paper, Department of Mechanical and Industrial Engineering, University of Massachusetts at Amherst*, (1996).
- Sengupta, J. K., Transformations in Stochastic DEA Models, *Journal of Econometrics*, Vol. 46, pp. 109–124 (1990).