

適合度・適合曲線と有意選出法の新提案

田 口 時 夫

報告者は、ローレンツ曲線とそれを多次元化した曲線を一般的に定義し、[1]でその詳細な解析を示した。そこで曲面への正射影を考察し、比率曲線を定義した。この曲線 $\begin{bmatrix} V(x) \\ W(x) \end{bmatrix}$ は、確率変数 X が連続な密度関数 $f(x)$ をもつ場合は、 $\varphi(x), \psi(x)$ を共に x の連続関数としたとき、条件

$$(1) \quad E\{\varphi(X)\}, E\{\psi(X)\} \neq 0$$

の下で次のように表わせる。

$$(2) \quad \begin{aligned} V(x) &= \int_{-\infty}^x \varphi(\xi) f(\xi) d\xi / E\{\varphi(X)\} \\ W(x) &= \int_{-\infty}^x \psi(\xi) f(\xi) d\xi / E\{\psi(X)\} \end{aligned}$$

(A) 与えられた密度関数 $f(x)$ に対して、適合しようとする密度関数を $\tilde{f}(x)$ とし、

$$(3) \quad \psi(x) = 1, \varphi(x) = \frac{\tilde{f}(x)}{f(x)}$$

とすれば、(2)は

$$(4) \quad \begin{aligned} V(x) &= \tilde{F}(x) \equiv \int_{-\infty}^x \tilde{f}(\xi) d\xi \\ W(x) &= F(x) \equiv \int_{-\infty}^x f(\xi) d\xi \end{aligned}$$

となる。この曲線を特に $f(x)$ の $\tilde{f}(x)$ による分布適合曲線と呼ぶ。

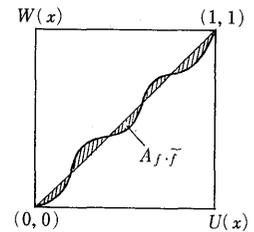


図 1.

もし適合が完成で

$$(5) \quad f(x) = \tilde{f}(x)$$

ならば、これは原点と $(1,1)$ を結ぶ線分となる。これを適合の基準線と呼ぶ。適合曲線と基準線が囲む面積は、不適合性を示す。従って $\tilde{f}(x)$ の $f(x)$ に対する不適合度 $F_{f, \tilde{f}}$ は、

$$(6) \quad F_{f, \tilde{f}} = 2A_{f, \tilde{f}} = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{sgn}(x_2 - x_1) \begin{vmatrix} \tilde{f}(x_1) & \tilde{f}(x_2) \\ f(x_1) & f(x_2) \end{vmatrix} dx_1 dx_2$$

で与えることが出来る。但し面積は基準線の下で正、上で負値をとる。企業統計にパレート分布を予想し

$$(7) \quad F_{f, \tilde{f}} = 0$$

により係数を決定した結果は、最尤法によるよりもよい結果を示した。

(B) 確率ベクトル (X, Y) において、 Y の X に関する回帰関数を $R(X)$ 、近似的回帰関数を $\tilde{R}(X)$ とする。更に

$$(8) \quad \begin{aligned} \psi(x) &= E(Y|X=x) \equiv R(x) \\ \varphi(x) &= \tilde{R}(x), E(Y), E\{\tilde{R}(X)\} \neq 0 \end{aligned}$$

とすれば、(2)式は、

$$(9) \quad \begin{aligned} V(x) &= \int_{-\infty}^x \tilde{R}(\xi) f(\xi) d\xi / E\{\tilde{R}(X)\} \\ W(x) &= \int_{-\infty}^x R(\xi) f(\xi) d\xi / E(Y) \end{aligned}$$

となる。この曲線を $R(x)$ の $\tilde{R}(x)$ による近似曲線と呼ぶ。又、(6) の $F_{f,f}$ に対応して $F_{R,\tilde{R}}$ が得られる。これを非近似度と呼ぶ。 Y を目的変量、 X を補助変量とすると【2】で示した方法でこの曲線をもとに有意選出法を与えることが出来る。

参 考 文 献

- [1] 田口。「経済分析と多次元解析」東洋経済新報社、昭和59年
 [2] 田口。「順位を基礎とする統計学の数学的方法のある種の体系化について」統計研彙報、第11巻第2号、昭和39年

目視にもとづく鯨資源推定——その2

岸 野 洋 久

ライントランセクト法は予め設定されたコースに沿って目視観測するべきだが、実際には鯨種および群れサイズの確認の為発見された鯨は接近される接近方式が採られていた。シミュレーションにより、鯨が不均質に分布しているときは、この方式は推定にバイアスを生ずることがわかった。ここで双眼鏡による発見メカニズムとハザードレイトとの関係を見てみる。

噴気間隔を T_0 とし、双眼鏡の距離別識別関数を $g(r)$ とし、一回の噴気が双眼鏡の視野に入る時間を \tilde{T} とおくと、船との相対位置 (r, θ) における鯨のハザードレイト $h(r, \theta)$ は

$$1 - e^{-h(r, \theta)T_0} = \int (1 - e^{-g(r)t}) P_\theta(\tilde{T} \in (t, t+dt))$$

と表わされる。 \tilde{T} の分布は角度別努力量分布と双眼鏡の視角および角速度から計算される。

ところで、鯨がパッチ状に分布しているときは調査の発見頭数の分散は大きくなるが、このことについて考察を加える。パッチの中心がなす分布を特性測度 $\lambda_0(x) dx$ をもつポアソン場 $N_0(dx)$ 、中心 x のパッチ内における鯨の分布を特性測度 $\lambda_1(z-x) dz$ をもつポアソン場 $N_1^{(x)}(dz)$ とすると、海域 A 内の鯨の頭数は

$$N(A) = \int_A \int_{R^2} N_0(dx) N_1^{(x)}(dz)$$

と表わされる。ここで $N_0, N_1^{(x)}$; $x \in R^2$ は独立である。

$$V(N(A)) = E(N(A)) + \int_{R^2} \lambda_0(x) dx \int_A \int_A \lambda_1(z_1-x) \lambda_1(z_2-x) dz_1 dz_2$$

となる。第2項が不均質性による分散の増加を表わす。 λ_1 として正規分布型 $C(2\pi\sigma^2)^{-1} \exp\{-|z-x|^2/2\sigma^2\}$ をとると、平均濃度 $\lambda_0 C$ を固定したとき、この項はパッチ内の濃度 C に比例し、パッチ半径 σ の減少関数となっている。また、半径が無小になった極限では $CE(N(A))$ となる。