

標識再捕時間に基づく死亡係数の推定、最近における理論展開と実際の解析上の問題

メタデータ	言語: Japanese
	出版者:
	公開日: 2025-04-24
	キーワード (Ja):
	キーワード (En):
	作成者: 北田, 修一
	メールアドレス:
	所属:
URL	https://fra.repo.nii.ac.jp/records/2014351

This work is licensed under a Creative Commons Attribution 4.0 International License.



綜 説

標識再捕時間に基づく死亡係数の推定, 最近における 理論展開と実際の解析上の問題

北 田 修 一*

(1989 年 9 月 11 日受理)

は じ め に

1 回放流の標識再捕データから漁獲係数 F や自然死亡係数 M を推定する方法は, 1950 年代の半ばから 1960 年代の半ばにかけて集中的に開発された。その後この問題については理論的な論文は発表されていないが, 近年になってオーストラリアとイギリスの研究者によって再び取り上げられるようになった。この背景は明らかではないが, その端緒は BEVERTON (1954) の log-linear model についての検討を行った論文である (SANDLAND, 1982; FAREBROTHER, 1985)。我が国においても栽培漁業を背景として, 死亡係数推定方法について 2 つの論文が出された (KITAHARA *et al.*, 1986; 北田・平野, 1987)。これまでのこれらの方法では死亡係数は時間の経過に関係なく一定と仮定している (我が国では能勢, 1971; 伊藤・村井, 1977; 北田, 1985, 1987a; 赤嶺, 1988 等の解説がある)。この仮定の制約を初めて克服したのが HEARN, SANDLAND and HAMPTON (1987) の論文 (HSH と略されている) である。ここでは M のみ一定の仮定の下で F を時刻 t の関数 $F(t)$ とし, シミュレーションによって 5 種類の $F(t)$ の関数形に対して M の推定値は robust であるという結論を得た。LEIGH (1988) は HEARN の依頼を受けて HSH 法による M の推定値の分散式を導き, 特殊な場合として $F(t)$ が一定の時, GULLAND (1955) のモデルと M の推定の良さについて比較を行った。また, FAREBROTHER (1988) は更に一步すすめて M についても時刻の関数とした解析方法を提案している。このように最近この問題について国の内外を問わず興味を持たれるようになってきている。これらの新しい 3 つの論文については 平松 (1989) の解説がある。

死亡係数推定の理論は統計学に基いており, log-linear model においては最小二乗法が, その他の手法には最尤法が用いられる。log-linear model は通常の最小二乗法を用いて良く利用されているが, 期間毎の再捕尾数に多項分布モデルを用いるこの場合は, 観測値が互いに相関を持つため一般化最小二乗法を用いるべきである。標識された個体群が一定の漁獲率で再捕されるので, その時の再捕数は海の中の生残標識個体の数を表わしている。このことから, 観測された再捕数が互いに相関を持つことは直感的に理解されよう。このことは最初に PAULIK (1963) が指摘し, 前出の 2 つの論文で更に検討されているが, 我が国のこれまでのレビューや教科書ではこのことにふれていない。また, 死亡係数は死亡時間の確率分布から定義されるが, 日本における水産資源学ではこのような観点から説明されたことはなかったように思われる。ここでは, 死亡係数の定義, log-linear model における加重最小二乗推定及び最尤法による解析手法を統計学の立場から整理して, 現在までの業績をレビューする。また, 実際のデータ解析においてしばしば直面する問題と解析手法の選択についても考察する。

第 1 章 推定理論

1. 死亡係数の定義と漁獲の理論

解析には個々の個体が再捕されるまでの時間が用いられる。この確率論的背景は次のようなものである。放流から死亡までの時間を X とする。 X は failure time と呼ばれる。死亡するまでの時間は 1 尾ずつ異っており, X

* 日本栽培漁業協会調査課 (〒101 東京都千代田区神田小川町 2-12 進興ビル)

は確率変数と考えることができる。 X の確率密度関数を $f(x)$, 分布関数を $\Phi(x)$ とする。放流されてから死亡せずに x 時間経過したという条件の下で, それに続く Δ 時間内に死亡する確率は

$$\Pr \{x < X \leq x + \Delta | X > x\} = \frac{\Phi(x + \Delta) - \Phi(x)}{1 - \Phi(x)}$$

で与えられる。

$$\lim_{\Delta \rightarrow 0} \frac{\Phi(x + \Delta) - \Phi(x)}{\Delta} = \Phi'(x) = f(x)$$

であるから, Δ が十分小さいとき, この確率は

$$\frac{f(x)}{1 - \Phi(x)} \Delta$$

で近似できる。

死亡係数は

$$h(x) = \frac{f(x)}{1 - \Phi(x)} \quad (1.1)$$

と書かれ, 統計学的には hazard function と呼ばれる (定義)。

時刻 t までの生残確率は

$$S(x) = 1 - \Phi(x) \quad (1.2)$$

であり, survivor function と呼ばれる。(1.2) 式を (1.1) 式に用い, (1.1) 式の両辺を 0 から x まで積分して

$$S(x) = \exp \left\{ - \int_0^x h(u) du \right\} \quad (1.3)$$

を得る。

このように死亡係数と死亡時間の確率分布は密接に関係しており, 生残確率は死亡係数から求まる。

この理論は工業製品や疫学等の survival data の解析で用いられているが, 漁業の場合, “漁獲” という行為が入るために多少異なるところがある。一般の life testing においては, サンプル全てについて failure time を計測するが (図1, A), 漁業の場合ではサンプル (ここでは標識放流尾数) の failure time のうち, 漁獲によるものだけが取扱われる (図 1, B)。

このことを F, M 一定の場合で具体的に見てみよう。

$$h(x) = F + M$$

であるから

$$f(x) = (F + M)e^{-(F+M)x},$$

$$\Phi(x) = 1 - e^{-(F+M)x}.$$

時刻 $(x, x + \Delta)$ の間に死亡する確率は $f(x)\Delta$ であるが, 漁獲の理論ではこれを漁獲される確率 $Fe^{-(F+M)x}\Delta$, と自然死亡する確率 $Me^{-(F+M)x}\Delta$ に分けて考える。即ち, $f_F(x) = Fe^{-(F+M)x}$, $f_M(x) = Me^{-(F+M)x}$ とすれば

$$f(x) = f_F(x) + f_M(x)$$

となっている。 $\Phi_F(x) = \int_0^x f_F(t)dt$, $\Phi_M(x) = \int_0^x f_M(t)dt$ と書けば

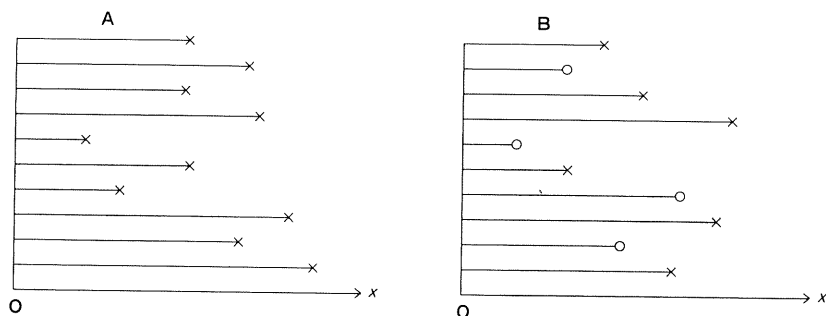


図 1 解析の対象となる failure time

A: 一般の解析 ×; failure

B: 漁業での解析 ×; 自然死亡による failure ○; 漁獲による failure

$$\Phi(x) = \Phi_F(x) + \Phi_M(x)$$

となることがわかる。 $\Phi_F(x)$ は漁獲率と呼ばれている。

2. 最小二乗推定

この節では FAREBROTHER (1985) に沿って解説する。

(1) log-linear model

時刻 0 で N 尾が放流される時、時刻 x_j から x_{j+1} の間に 1 個体が再捕される確率を P_j とし、この間に実際にとられる魚を n_j 、その期待値を W_j とすると、 $W_j = NP_j$ であるから

$$n_j = W_j + E_j \quad j=0, 1, \dots, J-1 \quad (2.1)$$

と表わされる。ここで、 $n_0, n_1, \dots, n_{J-1}$, $n_j = N - n_0 - n_1 - \dots - n_{j-1}$ が多項分布に従うから、 E_j は平均 $E(E_j) = 0$ 、分散 $\text{Var}(E_j) = NP_j(1-P_j)$ 、共分散 $\text{Cov}(E_j, E_k) = -NP_jP_k$ を持つ。

(2.1) 式は

$$\log(n_j/N) = \log P_j + E_j^* \quad j=0, 1, \dots, J-1 \quad (2.2)$$

と変形できる。ここで $E_j^* = \log(1 + E_j/W_j)$ である。 $\log(n_j/N) = \log \hat{P}_j$ を $\log P_j$ のまわりで第 1 項まで TAYLOR 展開することによって $E(E_j^*) = 0$, $\text{Var}(E_j^*) = (1-P_j)/(NP_j)$, $\text{Cov}(E_j^*, E_k^*) = -1/N$ を得る。ここで、 W_j がどの j に対しても 10 以上の時、 E_j^* は近似的に正規分布に従うことが知られている (PAULIK, 1963; FAREBROTHER, 1985)。

漁業が CCU* (PAULIK, 1963) の下では $P_j = e^{-Zx_j} \{F(1-e^{-Z})/Z\}$ だから

$$\log P_j = \log \{F(1-e^{-Z})/Z\} - Zx_j. \quad (2.3)$$

ここで、 $\alpha = \log \{F(1-e^{-Z})/Z\}$, $\beta = -Z$ とすれば

$$\log P_j = \alpha + \beta x_j \quad j=0, 1, \dots, J-1 \quad (2.4)$$

と書ける。ここで $Z = F + M$ である。

(2.4) 式を (2.2) 式に用いて、 $\log(n_j/N) = y_j$ とすると

$$y_j = \alpha + \beta x_j + E_j^* \quad j=0, 1, \dots, J-1 \quad (2.5)$$

となり、通常の線形回帰モデルが得られる。

ここで E_j^* が平均 0、分散 σ^2 の正規分布に従うなら通常の最小二乗法が使えるが、多項分布モデルが適用される再捕データの解析においては E_j^* の分散共分散は上で見たように少し異なるものになっている。これをもう少し具体的に見てみよう。

(2) 加重最小二乗推定量

$$\mathbf{y} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{J-1} \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{J-1} \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \sigma E_0^* \\ \sigma E_1^* \\ \vdots \\ \sigma E_{J-1}^* \end{bmatrix} \quad \mathbf{W} = \begin{bmatrix} W_0 & & & \\ & W_1 & & 0 \\ & \vdots & \ddots & \\ & 0 & & W_{J-1} \end{bmatrix}$$

とすると、(2.5) 式は

$$\mathbf{y} = \mathbf{ac} + \beta \mathbf{x} + \boldsymbol{\varepsilon} \quad (2.6)$$

となる。

E_j^* の分散共分散行列は

$$\text{Var}(\boldsymbol{\varepsilon}) = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2 \begin{pmatrix} \frac{1-P_0}{NP_0} & -\frac{1}{N} & \dots & -\frac{1}{N} \\ -\frac{1}{N} & \frac{1-P_1}{NP_1} & \dots & -\frac{1}{N} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{N} & \vdots & \dots & \frac{1-P_{J-1}}{NP_{J-1}} \end{pmatrix} \quad (2.7)$$

* constant な漁獲圧で continuous に営まれる漁業を言う。ここで漁具が漁場内で uniformly に配置されている。CCU の下では、標識魚が漁場内に均一に分布していなくても、再捕尾数は標識個体群からのランダムサンプルと考えることができる。

である。ここで、 ϵ' は ϵ の転置を意味する。これは PAULIK (1963, p. 39) の (31) 式と同じである。(31) 式においては $\sigma^2=1$ としている。モデル (2.5) でも $\sigma^2=1$ ある。このように E_j^* の分散共分散行列は、非対角要素も 0 ではないので、一般化最小二乗法を用いるべき状況にある。PAULIK (1963, p. 40) は (2.7) の逆行列を導き、それで重みづけられた一般化最小二乗推定量を求めた。

FAREBROTHER (1985) はこの一般化最小二乗推定量が、ウェイトを適切に選べば加重最小二乗法によって得られることを示した。

今、 $y_* = W^{1/2}y$, $c_* = W^{1/2}c$, $x_* = W^{1/2}x$, $\epsilon_* = W^{1/2}\epsilon$, $W^{1/2}$ は対角行列 $\{W_0^{1/2}, \dots, W_{J-1}^{1/2}\}$ とすると (2.6) は

$$y_* = \alpha c_* + \beta x_* + \epsilon_* \quad (2.8)$$

となる。

$X_* = [c_* \ x_*]$, $\gamma = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ とすれば (2.8) 式は

$$y_* = X_* \gamma + \epsilon_* \quad (2.9)$$

となる。このモデルにおいて $E(\epsilon_*) = 0$, $E(\epsilon_* \epsilon_*') = \sigma^2 V$ である。 $\sigma^2 V$ を具体的に書くと

$$E(\epsilon_* \epsilon_*') = \sigma^2 \begin{bmatrix} 1-P_0 & W_0^{1/2} W_1^{1/2} \left(-\frac{1}{N}\right) \cdots W_0^{1/2} W_{J-1}^{1/2} \left(-\frac{1}{N}\right) \\ W_1^{1/2} W_0^{1/2} \left(-\frac{1}{N}\right) & 1-P_1 & \cdots W_1^{1/2} W_{J-1}^{1/2} \left(-\frac{1}{N}\right) \\ \vdots & \vdots & \cdots & \vdots \\ W_{J-1}^{1/2} W_0^{1/2} \left(-\frac{1}{N}\right) & \cdots & \cdots & 1-P_{J-1} \end{bmatrix} \quad (2.10)$$

のようになる。

ここで、 $R'R = RR' = R^2 = V$ になるような行列 R を考え R^{-1} をモデル (2.9) に右がけて

$$R^{-1}y_* = R^{-1}X_*\gamma + R^{-1}\epsilon_*$$

これを

$$T = Q\gamma + f \quad (2.11)$$

と書くと

$$E(f) = E(R^{-1}\epsilon_*) = 0$$

$$\begin{aligned} \text{Var}(f) &= E(ff') = E\{R^{-1}\epsilon_* \epsilon_*' (R^{-1})'\} = R^{-1} \text{Var}(\epsilon_*) R^{-1'} \\ &= R^{-1} \sigma^2 V R^{-1'} = \sigma^2 I \end{aligned}$$

となるから通常の最小二乗法を (2.11) に応用して問題ない。

正規方程式は

$$Q'Q\gamma = Q'T.$$

従って

$$\begin{aligned} (R^{-1}X_*)' R^{-1}X_* \gamma &= (R^{-1}X_*)' R^{-1}y_*, \\ \hat{\gamma} &= (X_*' V^{-1} X_*)^{-1} X_*' V^{-1} y_* \end{aligned} \quad (2.12)$$

$$\begin{aligned} \text{Var}(\hat{\gamma}) &= \sigma^2 (Q'Q)^{-1} \\ &= \sigma^2 (X_*' V^{-1} X_*)^{-1} \end{aligned} \quad (2.13)$$

となり、 V^{-1} で加重されている。これは一般化最小二乗推定量である。

このままでは V^{-1} を求めなければならないが、FAREBROTHER (1985, Appendix) は

$$\begin{aligned} \hat{\gamma} &= (X_*' V^{-1} X_*)^{-1} X_*' V^{-1} y_* \\ &= (X' W X)^{-1} X' W y \end{aligned} \quad (2.14)$$

$$\begin{aligned} \text{Var}(\hat{\gamma}) &= \sigma^2 (X_*' V^{-1} X_*)^{-1} \\ &= \sigma^2 (X' W X)^{-1} + \sigma^2 A \\ &= \sigma^2 B^{-1} + \sigma^2 A \end{aligned} \quad (2.15)$$

となることを示した。ここで $[X=c \ x]$ である (原著では $[x \ c]$ となっているが印刷ミスであろう)。また、

$$A = \begin{bmatrix} -1/N & 0 \\ 0 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} \sum_{j=0}^{J-1} W_j & \sum_{j=0}^{J-1} W_j x_j \\ \sum_{j=0}^{J-1} W_j x_j & \sum_{j=0}^{J-1} W_j x_j^2 \end{bmatrix}$$

である。

このように、モデル (2.6) の一般化最小二乗推定量はモデル (2.8) の一般化最小二乗推定量と同じであるので、(2.14), (2.15) はモデル (2.5) の一般化最小二乗推定量であることが示された。これは同時に W 行列で加重した加重最小二乗推定量になっている。

行列表現を具体的な形で書けば、モデル (2.5) では $\sigma^2=1$ であるので

$$(X'WX)^{-1}B^{-1} = \frac{1}{\det(B)} \begin{bmatrix} \sum_{j=0}^{J-1} W_j x_j^2 & -\sum_{j=0}^{J-1} W_j x_j \\ -\sum_{j=0}^{J-1} W_j x_j & \sum_{j=0}^{J-1} W_j \end{bmatrix} \quad (2.16)$$

$$X'WY = \begin{bmatrix} \sum_{j=0}^{J-1} W_j y_j \\ \sum_{j=0}^{J-1} x_j W_j y_j \end{bmatrix}.$$

従って

$$\hat{\gamma} = \frac{1}{\det(B)} \begin{bmatrix} \sum W_j x_j^2 & \sum W_j y_j - \sum W_j x_j \sum x_j W_j y_j \\ -\sum W_j x_j & \sum W_j y_j + \sum W_j \sum x_j W_j y_j \end{bmatrix}. \quad (2.17)$$

ここで

$$\det(B) = \sum_{j=0}^{J-1} W_j \sum_{j=0}^{J-1} W_j (x_j - \bar{x})^2.$$

これらから

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} \quad (2.18)$$

$$\hat{\beta} = \frac{\sum_{j=0}^{J-1} W_j (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=0}^{J-1} W_j (x_j - \bar{x})^2} \quad (2.19)$$

を得る。ここで

$$\bar{x} = \frac{\sum_{j=0}^{J-1} W_j x_j}{\sum_{j=0}^{J-1} W_j}, \quad \bar{y} = \frac{\sum_{j=0}^{J-1} W_j y_j}{\sum_{j=0}^{J-1} W_j}. \quad (2.20)$$

また推定値の分散はそれぞれ

$$\text{Var}(\hat{\alpha}) = \frac{\sum_{j=0}^{J-1} W_j x_j^2}{\sum_{j=0}^{J-1} W_j \sum_{j=0}^{J-1} W_j (x_j - \bar{x})^2} - \frac{1}{N} \quad (2.21)$$

$$\text{Var}(\hat{\beta}) = 1 / \sum_{j=0}^{J-1} W_j (x_j - \bar{x})^2 \quad (2.22)$$

となる。

これらは PAULIK (1963, p. 40) の一般化最小二乗推定値と本質的に同じである。ただ、PAULIK のモデルは $j-1$ を j でおきかえたものである。

$$y_j^0 = \alpha^0 + \beta x_j + E_j^0 \quad j=1, 2, \dots, J.$$

ここで、 $\alpha^0 = \alpha - \beta$, $y_j^0 = y_{j-1}$, $E_j^0 = E_{j+1}^*$, $x_j = j$ であるので、 $\hat{\beta}$ は同じであるが、 $\hat{\alpha}^0$ は FAREBROTHER の推定量と異っている。また、 $\text{Var}(\hat{\alpha}^0)$ は (2.21) において、 $\sum_{j=0}^{J-1} W_j x_j^2$ を $\sum_{j=0}^{J-1} W_j x_{j+1}^2$ として計算することになる。さらに、PAULIK は $\hat{\alpha}^0$ の分散に A 行列を考慮しなかったため、 $1/N$ の項は除いて計算している。

N が大きいときには、 $\text{Cov}(E_j^*, E_k^*) = -1/N$ は 0 に近づくので、(2.7) の非対角要素は全て 0 に近づき、対角要素の逆数をウェイトにした加重最小二乗法を使うことができる。ここでウェイトは $W_j^* = [\text{Var}(E_j^*)]^{-1} = NP_j / (1 - P_j)$ である。これは SANDLAND (1982) のウェイトであるが、 N が大きいときは $\text{Cov}(E_j^*, E_k^*) = -1/N$ は 0 に近づき、 $\text{Var}(E_j^*) = (1 - P_j) / NP_j = 1 / NP_j - 1/N = 1/W_j$ に近づく。従って、ウェイトは W_j で良いこと

になる。この時の加重最小二乗推定値は先に見たように、一般化最小二乗推定値と同じである。この W ウェイトを持つ加重最小二乗推定値の正確な分散行列は $A+B^{-1}$ で与えられ、 N が大きいときは B^{-1} に近い。

実際のデータで計算する場合には W_j はわからないのでこの代りに n_j を用いる。

(3) 信頼区間の計算について

PAULIK (1963, p. 14) は α, β の区間推定に t 分布を利用した。即ち、 $t_\alpha = (\hat{\alpha} - \alpha) / \sqrt{\widehat{\text{Var}}(\hat{\alpha})}$, $t_\beta = (\hat{\beta} - \beta) / \sqrt{\widehat{\text{Var}}(\hat{\beta})}$ が自由度 $J-2$ の t 分布に従うことを利用する。一方、FAREBROTHER (1985, p. 168) は t_α, t_β が標準正規分布に従うとして信頼区間を計算しているが、その理由は不明である。ところで $\text{Var}(\hat{\alpha})$, $\text{Var}(\hat{\beta})$ がわかっているなら t_α, t_β は標準正規分布に従うが、分散の推定量を用いる場合はこれらは t 分布に従う。この場合、(2.21), (2.23) 式からわかるように W_j がわからないので、代りに実測値 n_j を用いて $\hat{\alpha}, \hat{\beta}$ の分散を推定している。このことは通常の直線回帰でも行われており、そのために一般に t 分布が用いられるのである。ここで FAREBROTHER が正規分布を用いたのは、用いたデータでは n_j が大きかったために、 $\text{Var}(\hat{\gamma}) \approx \widehat{\text{Var}}(\hat{\gamma})$ としたのであろう。 t 分布を用いる方が信頼幅は大きくなるが、一般的には t 分布を用いる方が良いだろう。

また、PAULIK は推定値の分散を $\sigma^2 B^{-1}$ とした。

即ち

$$\text{Var}(\hat{\gamma}) = \sigma^2 (X'WX)^{-1} \quad (2.23)$$

とし、 σ^2 に $\hat{\sigma}^2$ を用いて計算した。ここで

$$\hat{\sigma}^2 = \sum_{j=0}^{J-1} W_j (y_j - \hat{\alpha} - \hat{\beta} x_j)^2 / (J-2) . \quad (2.24)$$

しかし、モデル (2.5) では $\sigma^2=1$ であるし、PAULIK 自身も (2.7) 式において $\sigma^2=1$ としているので、 $\hat{\sigma}^2$ を用いるのは奇妙である。FAREBROTHER はこれ対にして、 $\hat{\sigma}^2$ を用いることはモデル (2.5) が

$$y_j = \alpha + \beta x_j + \sigma E_j^* \quad (2.25)$$

と書き直されたモデルにおいて、active パラメータである時に限って正当化されるが、これを入れておく強い理由はないとしている。しかし、 $\hat{\sigma}^2$ はモデルの仮定や理論の中で行っている近似の前提がくずれた場合に、1 より大きく外れた値をとるだろうとし、 $\hat{\sigma}^2$ をモデルの misspecification の示標とした。すなわち、 $\sum (\hat{E}_j^{*2}) / \sigma^2$ が自由度 $(J-2)$ のカイ 2 乗分布 $\chi^2(J-2)$ に従うことを利用する。ここで帰無仮説 $H_0: \sigma^2=1$ とすると、 H_0 が正しいという仮説の下で、 $\sum (\hat{E}_j^{*2}) = (J-2)\hat{\sigma}^2$ が $\chi^2(J-2)$ に従うので、 H_0 を検定することができる。ところで、(2.7) の分散共分散行列（あるいはこれを W^{-1} で近似した場合でも）はわからないので、実測値を使って推定しているため、その時の誤差が加わる。このことを考慮して PAULIK は $\hat{\sigma}^2$ をつけたものと考えて良いだろう。再捕尾数が多項分布から外れる程、 $\hat{\sigma}^2$ の値は 1 から外れる。 $\hat{\sigma}^2$ はモデルのあてはまりの良さの情報を与えるので、FAREBROTHER の考え方は自然であるが、分散の推定値の計算においては、PAULIK がそうしたようにつけておいた方が良いだろう。

3. 最尤推定

3-1. 全尤度による推定

(1) 正確な再捕時間が知られている場合 (exact-time data)

$Z(t) = F(t) + M(t)$ とすれば、時刻 0 で放流された個体が x までに再捕される確率は

$$P(x) = \int_0^x F(v) S(v) dv \quad (3.1)$$

で与えられる。

N 尾が時刻 0 で放流され、合計 n 尾が時刻 U までの $x_1, x_2, \dots, x_n (=U)$ でそれぞれ 1 尾ずつ再捕される確率は、尤度関数

$$\begin{aligned} L &= m \prod_{i=1}^n F(x_i) S(x_i) \{1 - P(U)\}^{N-n} \\ &= m \prod_{i=1}^n F(x_i) \exp \left\{ - \int_0^{x_i} Z(u) du \right\} \left[1 - \int_0^U F(v) \exp \left\{ - \int_0^v Z(u) du \right\} dv \right]^{N-n} \end{aligned} \quad (3.2)$$

で表わされる。ここで、 $m=N!/[(N-n)!n!]$ である。これは FAREBROTHER (1988, p. 231) の (18) 式と同等であり、再捕の全尤度である。

F も M も一定の時、尤度関数は

$$L=m \prod_{i=1}^n F e^{-Z x_i} \{1-(F/Z) e^{-Z U}\}^{N-n} \quad (3.3)$$

となる。これは PAULIK (1963, p. 34) の (9) 式と同じである。また、GULLAND (1955) はこの尤度関数を導き $U \rightarrow \infty$ の時の最尤推定量を与えた。KITAHARA *et al.* (1986) のモデルは休漁期を含んでいるため尤度関数の形は独特であるが、尤度についての基本的な考え方は同じである。

HSR (1987) の推定式は

$$\sum_{i=1}^n e^{M x_i} = N \quad (3.4)$$

であり、これを数値計算で解いて $\hat{M}$ を求める。推定式の中に $F(t)$ が入っていないが、 x_i で再捕された 1 尾が $x_0=0$ では $e^{M x_i}$ 尾であったということを示したものである。つまり、自然死亡した魚以外は全てとられ、推定時点では海の中に 1 尾も残っていないという仮定を置いている。

LEIGH (1988) はこの時の尤度関数を与えている。この尤度は基本的には (3.2) 式と同じであるが、その構造を理解するためには良い表現をとっている。即ち

$$L=m \prod_{i=1}^n [\{\Phi(x_i+D)-\Phi(x_i)\} e^{-M x_i}] \left\{1 - \int_0^\infty e^{-M x} d\Phi(x)\right\}^{N-n}. \quad (3.5)$$

ここで $\Phi(x)$ は再捕時間 X の分布関数である。 x_i まで生きた魚は漁獲によってとれるから、漁獲係数一定の条件で考えると

$$\Phi(x)=1-e^{-F x} \quad (3.6)$$

である。従って

$$\begin{aligned} \{\Phi(x_i+D)-\Phi(x_i)\} e^{-M x_i} &= f(x_i) e^{-M x_i} D \\ &= F e^{-(F+M) x_i} D, \\ \int_0^\infty e^{-M x} d\Phi(x) &= \int_0^\infty e^{-M x} \frac{d\Phi(x)}{dx} dx = \int_0^\infty e^{-M x} f(x) dx \\ &= \int_0^\infty F e^{-(F+M) x} dx \end{aligned}$$

となって、 $U \rightarrow \infty$ の時の GULLAND の尤度関数となる。密度関数を用いて書くと

$$L=m \prod_{i=1}^n \{f(x_i) e^{-M x_i}\} \left\{1 - \int_0^\infty e^{-M x} f(x) dx\right\}^{N-n} \quad (3.7)$$

となる。この時、 $h(x)=F e^{-F x}/e^{-F x}=F$ である。即ち、 $e^{-M x_i}$ の確率で x_i まで生きた魚が $f(x_i)$ という確率（密度）でとられるという考えである。(3.1) 及び (3.2) 式では、密度関数を使わない表現になっているので、 $S(x)=\exp\left\{-\int_0^x Z(u) du\right\}$ である。即ち、 x_i での漁獲確率は、漁獲係数と全減少に対応した生残率の積の形となっている。これはまた、漁獲の確率密度関数と自然死亡のみに対応した生残率の積で表わされる。

(2) 再捕時間が期間毎にまとめられている場合 (grouped data)

(1) ではそれぞれの個体の再捕までの時間が正確に知られていることを仮定したが、ここでは、それがわからないが、時刻 x_j から x_{j+1} の間にとられたことがわかっている場合について考える。(3.1) からこのことが起る確率は

$$\begin{aligned} P_j &= P(x_{j+1}) - P(x_j) \\ &= \int_{x_j}^{x_{j+1}} F(v) S(v) dv \\ &= \int_{x_j}^{x_{j+1}} F(v) \exp\left\{-\int_0^{x_j} Z(u) du\right\} dv \end{aligned}$$

であるから

$$\begin{cases} P_j \equiv (F_j/Z_j)(1-e^{-Z_j}) & (j=0) \\ P_j \equiv \exp\left(-\sum_{i=1}^{j-1} Z_i\right) (F_j/Z_j)(1-e^{-Z_j}) & (j=1, \dots, J-1) \end{cases} \quad (3.8)$$

表わされる。

時刻 0 から U までの間の x_j から x_{j+1} の間に n_j が再捕され、 $N-n_0-n_1-\cdots-n_{J-1}$ が再捕されない場合の尤度は

$$L = m^* \prod_{j=0}^{J-1} P_j^{n_j} (1 - \sum_{j=0}^{J-1} P_j)^{N-n} \quad (3.9)$$

で与えられる。ここで $m^* = N! / \{n_0! \cdots n_{J-1}! (N-n)!\}$ である。これは FAREBROTHER (1988) の (22) 式と同じである。FAREBROTHER は、この尤度が、PAULIK (1963) の (20) 式と同じであることを導いている。これによって、PAULIK の推定方法が見通し良く理解できるので、以下に記しておこう。

(3) PAULIK の考え方 (PAULIK's reparametarization)

PAULIK は F, M が一定の場合で考え

$$\mu = (F/Z)(1 - e^{-ZU}) \quad (3.10)$$

$$s = e^{-Z} \quad (3.11)$$

とした。これから

$$F = -\mu \log s / (1 - s^U) \quad (3.12)$$

$$Z = -\log s. \quad (3.13)$$

従って、(3.3) 式は次のように書ける。

$$\begin{aligned} L &= m F^n e^{-Z \sum x_i} \{1 - P(U)\}^{N-n} \\ &= m \mu^n \left\{ \frac{-\log s}{1 - s^U} \right\}^n s^{\sum x_i} (1 - \mu)^{N-n}. \end{aligned} \quad (3.14)$$

これは PAULIK (1963) の (10) 式であるが、印刷ミスがある。

$$\log L = \log m + n \log \mu - n \log(\log s) - n \log(1 - s^U) + \sum x_i \log s + (N - n) \log(1 - \mu) \quad (3.15)$$

だから

$$\frac{\partial \log L}{\partial \mu} = \frac{n}{\mu} - \frac{(N - n)}{1 - \mu} \quad (3.16)$$

$$\frac{\partial \log L}{\partial s} = \frac{n}{s \log s} + \frac{n U s^{U-1}}{1 - s^U} + \frac{\sum x_i}{s} \quad (3.17)$$

これから

$$\hat{\mu} = n / N \quad (3.18)$$

$$\frac{-1}{\log \hat{s}} - \frac{U \hat{s}^U}{1 - \hat{s}^U} = \frac{\sum x_i}{n} \quad (3.19)$$

を得る。ここで (3.19) 式は PAULIK (1963) の (12) 式、FAREBROTHER (1988) の (B.8) 式と同じであるが、第 1 項の形が異なる。これは計算ミスではないかと思われる。

$\hat{s}$ の漸近分散は、 $\hat{s}$ と $\hat{\mu}$ は無相関なので

$$\begin{aligned} \text{Var}(\hat{s}) &= \left[-E \left\{ \frac{\partial^2 \log L}{\partial s^2} \right\} \right]^{-1} \\ &= \frac{1}{n} \left\{ \frac{1}{(\log s)^2} - \frac{U^2 s^U}{(1 - s^U)^2} \right\}^{-1} \end{aligned} \quad (3.20)$$

として求まる。

次に grouped data の場合を考える。

$$P(x) = (F/Z)(1 - e^{-Zx}) \quad (3.21)$$

だから

$$\begin{aligned} P(x) &= (F/Z)(1 - e^{-ZU})(1 - e^{-Zx}) / (1 - e^{-ZU}) \\ &= \mu(1 - s^x) / (1 - s^U) \end{aligned} \quad (3.22)$$

$$\begin{aligned} P(j+1) - P(j) &= \mu s^j (1 - s) / (1 - s^U) \\ &= \mu s^j / \sum_{k=0}^{J-1} s^k \end{aligned} \quad (3.23)$$

と書ける。(3.9) 式は

$$L = m^* \prod_{j=0}^{J-1} (\mu s^j / \sum_{k=0}^{J-1} s^k)^{n_j} (1-\mu)^{N-n} \quad (3.24)$$

となる。ここで、 $T = \sum_{j=0}^{J-1} j n_j$ とすれば

$$L = m^* (1-\mu)^{N-n} \mu^n s^T (1 / \sum_{k=0}^{J-1} s^k)^n \quad (3.25)$$

となり、両辺の対数をとれば

$$\log L = \log m^* + (N-n) \log(1-\mu) + n \log \mu + T \log s + n \log (1/\sum s^k). \quad (3.26)$$

μ, s で偏微分とすると

$$\frac{\partial \log L}{\partial \mu} = \frac{n-N}{1-\mu} + \frac{n}{\mu} \quad (3.27)$$

$$\frac{\partial \log L}{\partial s} = \frac{T}{s} - \frac{n \sum_{j=0}^{J-1} j s^{j-1}}{\sum_{k=0}^{J-1} s^k}. \quad (3.28)$$

これから

$$\hat{\mu} = n/N \quad (3.29)$$

$$\frac{\sum_{j=0}^{J-1} j \hat{s}^j}{\sum_{k=0}^{J-1} \hat{s}^k} = \frac{T}{n} \quad (3.30)$$

が得られる。 $\hat{\mu}$ の漸近分散は

$$\text{Var}(\hat{\mu}) = \left[-E \left\{ \frac{\partial^2 \log L}{\partial \mu^2} \right\} \right]^{-1} = \mu(1-\mu)/N, \quad (3.31)$$

$\hat{s}$ のそれは (3.30) 式を利用して

$$\begin{aligned} \text{Var}(\hat{s}) &= \left[-E \left\{ \frac{\partial^2 \log L}{\partial s^2} \right\} \right]^{-1} \\ &= \frac{s^2 (\sum_{j=0}^{J-1} s^j)^2}{n \{ \sum_{j=0}^{J-1} s^j \sum_{j=0}^{J-1} j^2 s^j - (\sum_{j=0}^{J-1} j s^j)^2 \}} \end{aligned} \quad (3.32)$$

で与えられる。

3-2 部分尤度による推定

前節では再捕の全尤度による推定について述べた。外国で開発されてきた推定方法は全てこのタイプであるし、KITAHARA *et al.* (1986) の方法もこの尤度を用いている。しかし、 F/Z が一定の条件の下では別の方法が考えられる。

北田・平野 (1987) は死亡係数が一定の時は死亡時間の分布は指数分布であり、その平均値の逆数が Z である点に着目した。推定の手順はまず指数分布の平均値を推定して Z の推定値を求め、その後放流数に対する再捕尾数の割合からこれを F と M に分離するというものである。そして期間毎のデータに基づいて、実験期間が有限、無限の両方の場合について推定方法を示した。この方法ではデータをまとめる幅が必ずしも同じでなくても良い。

TANAKA (1967) は GULLAND (1955) と同じ推定量を導いたが、尤度関数のつくり方において両者は異なっている。TANAKA の方法もまず指数分布のパラメータ Z を推定し、この後 F と M に分離する。TANAKA は実験期間が無限の場合に、データが exact-time でとられたときの推定方法を提案した。これら 2 つの方法は本質的に同じ考え方にたっており、3-1 で述べた方法と少し視点を異にしている。この推定の考え方は (3) で説明する。

F と M のそれぞれの推定は exact-time data, grouped data いずれの場合でも次のようにすれば良い。(3.1) 式より時刻 U までに再捕される確率は

$$P(U) = F(1 - e^{-ZU})/Z$$

だから、 U までに再捕される期待尾数 n は

$$n = NF(1 - e^{-ZU})/Z$$

である。 $\hat{Z}$ が与えられれば、 F の推定量 $\hat{F}$ は $n=NF(1-e^{-\hat{Z}U})/\hat{Z}$ を F について解いた

$$\hat{F}=n\hat{Z}/\{N(1-e^{-\hat{Z}U})\} \quad (3.33)$$

で与えられる。また M の推定量 $\hat{M}$ は

$$\hat{M}=\hat{Z}-\hat{F}$$

で与えられる。なお、非打ち切りの場合 UZ が十分大きいので $1-e^{-ZU}=1$ として扱われ、(3.33) 式は $\hat{F}=n\hat{Z}/N$ となり (3.34) 式はこの $\hat{F}$ に対して $\hat{M}$ を与える。

このように、死亡係数が一定の時は、指数分布のスケールパラメータの推定問題に帰結する。データの状況に応じた推定方法は統計学において知られており、これを利用する。以下にデータの状態別の推定式を記す。

(1) exact-time data

N 尾が時刻 0 で放流され、時刻 U までに合計 n 尾が $x_1, x_2, \dots, x_n (=U)$ でそれぞれ 1 尾ずつ再捕される。 $h(x)=Z$ であるから、確率密度関数は

$$g(x)=Ze^{-Zx}/(1-e^{-ZU}) \quad (3.35)$$

である。尤度関数は

$$\begin{aligned} L &= g(x_1) \cdots g(x_n) \\ &= Z^n \exp(-Z \sum_{i=1}^n x_i) (1-e^{-ZU})^{-n}. \end{aligned} \quad (3.36)$$

$$\frac{\partial \log L}{\partial Z} = \frac{n}{Z} - \frac{nUe^{-ZU}}{1-e^{-ZU}} - \sum x_i \quad (3.37)$$

より、 Z の最尤推定量は方程式

$$\frac{1}{Z} - \frac{Ue^{-ZU}}{1-e^{-ZU}} = \frac{\sum x_i}{n} \quad (3.38)$$

の一意的な正の根として与えられる (DEEMER and VOTAW, 1955)。(3.38) 式は (3.19) 式と全く同じであることがわかる。

U が大きい時は、 e^{-ZU} は 0 に近づくので (3.38) 式から

$$\hat{Z} = \sum x_i / n \quad (3.39)$$

を得る。これは TANAKA (1967) の推定量であり、GULLAND (1955) のそれと同じである。

(2) grouped data

打ち切り時刻 U の指数分布の分布関数は

$$\Phi(x) = (1-e^{-Zx})/(1-e^{-ZU}) \quad (3.40)$$

であるから、時刻 x_j から x_{j+1} に死亡する確率は

$$Q_j = \Phi(x_{j+1}) - \Phi(x_j) \quad j=0, 1, \dots, J-1 \quad (3.41)$$

である。

x_j で n_j 尾が再捕される時、尤度関数は

$$L = m^{**} \prod_{j=0}^{J-1} Q_j^{n_j} \quad (3.42)$$

である。ここで $m^{**} = n!/(n_0!n_1!\cdots n_{J-1}!)$ である。

Z の最尤推定量 Z は $n_1 < n$ かつ $\sum_{j=0}^{U-1} n_j(x_j + x_{j+1}) < nU$ のとき、方程式

$$\sum_{j=0}^{J-1} \frac{n_j(x_{j+1} - x_j)}{e^{Z(x_{j+1} - x_j)} - 1} - \frac{nU}{e^{ZU} - 1} - \sum_{j=1}^{J-1} n_j x_j = 0 \quad (3.43)$$

の一意的な正の根として与えられる。

期間の幅 $(x_{j+1} - x_j = x_1)$ が全て等しい場合では $\hat{Z}$ は、 $\sum_{j=0}^{J-1} (j+1)n_j < (J+1)n/2$ のとき

$$\frac{1}{e^{Zx_1} - 1} - \frac{J}{e^{ZU} - 1} - \frac{1}{n} \sum_{j=1}^{J-1} j n_j \quad (3.44)$$

の一意的な正の根として与えられる。これらは KULLDORFF (1961) によって導かれている。

実験を途中で打ち切らない場合は、期間幅を全て同じとして解析できる。

分布関数は

$$\Phi(x)=1-e^{-Zx}$$

であるから

$$\begin{aligned} Q_j &= \Phi\{(j+1)x_1\} - \Phi(jx_1) \quad j=0, 1, \dots, J-2, \\ Q_{J-1} &= 1 - \Phi\{(J-1)x_1\} \end{aligned}$$

である。尤度関数は

$$L = m^{**} \prod_{j=1}^{J-2} Q_j^{n_j} Q_{J-1}^{n_{J-1}} \quad (3.45)$$

となる。この時の最尤推定量は知られており（例えば高橋・脇本・平野（1975）），

$$\hat{Z} = \log \{(T^* - n_{J-1}) / (T^* - n)\} / x_1 \quad (3.46)$$

で与えられる。ここで、 $T^* = \sum_{j=0}^{J-1} (j+1)n_j$ である。

（３） 推定の考え方

ここでは、 F に起因する再捕尾数だけから尤度をつくって Z を推定する。 F が一定なら、再捕尾数のヒストグラムは海の中の標識魚の生残数を相対的に表わしている。これから Z を推定し、その後標識尾数に対する再捕数をみることによって F を求めるという考え方である。このことについては、岸野*が次のような解釈を与えている。

「再捕尾数 $n_0, \dots, n_{J-1}$ の尤度（全尤度）は 3-1 で示されたが、この尤度 $L = (Z, F | n_0, \dots, n_{J-1})$ を

$$\begin{aligned} L &= (Z, F | n_0, \dots, n_{J-1}) \\ &\equiv \Pr(n_0, \dots, n_{J-1} | Z, F) \\ &= \Pr(n_0, \dots, n_{J-1} | n, Z) \Pr(n | Z, F) \\ &\equiv L_1(Z | n_0, \dots, n_{J-1}, n) L_2(Z, F | n) \end{aligned} \quad (3.47)$$

と分解して、全尤度から直接 Z, F を推定する代りに、まず第 1 項の条件付き尤度（部分尤度）から Z を推定し、得られた Z を用いて第 2 項から F を求めるということを行っている。」

第 1 項の条件付き尤度は (3.36), (3.42), (3.45) 式で示したが、第 2 項の尤度は

$$L_2 = \frac{N!}{(N-n)!n!} \left\{ \frac{F}{Z} (1 - e^{-ZV}) \right\}^n \left\{ 1 - \frac{F}{Z} (1 - e^{-ZV}) \right\}^{N-n} \quad (3.48)$$

である。

$$\frac{\partial \log L_2}{\partial F} = \frac{n}{F} - \frac{(N-n)(1 - e^{-ZV})/Z}{1 - F(1 - e^{-ZV})/Z} = 0$$

から

$$\hat{F} = n \hat{Z} / \{N(1 - e^{-\hat{Z}V})\}$$

を得る。これは (3.33) 式であり、この $\hat{F}$ は最尤推定量になっていることがわかる。

このように岸野の与えた解釈が実際に確かめられる。

第 2 章 実際の解析における問題

1. 死亡係数を一定と仮定する場合

（１） 実験期間の打ち切りと仮定の検証

我々が実際に扱わなければならないデータで、きれいな指数分布にそのままあてはまるものはむしろ少ない。これは、実験の条件設定に問題がある場合も多いが、注意深く実験を行った場合においてもみられる現象である。この原因は、死亡係数、特に漁獲係数が季節とともに変化していることにある。つまり、死亡係数は本来、時刻の関数であり、これを短い期間では一定と仮定して解析するのである。従って死亡係数一定の指数分布モデルを実際に使う場合には、その仮定を満足するためにデータを途中で打ち切ることが殆んどの場合で必要になる。もちろん、実験を続けても 1 尾も再捕されない状況もあり得る。従来は、どの解析手法をどのような状況のデータに使えば良いのかという点において、必ずしも整理が十分とは言えなかったため、誤った解析事例もよく見られる。データの状態、即ち、打ち切りか非打ち切りか、あるいは、1 日毎のデータがあるのか期間毎のデータなのか、このような

* 岸野洋久（私信）

表 1 指数分布に基づく死亡係数推定モデルと適応データ

モ デ ル	1 日毎のデータ		期間毎のデータ	
	打ち切り	非打ち切り	打ち切り	非打ち切り
BEVERTON (1954)			○	
GULLAND (1955)		○		
PAULIK (1963)	○		○	
TANAKA (1967)		○		
KITAHARA et al. (1986)	○*			
北田・平野(1987)	○	○	○**	○

* 休漁期を含む ** 期間幅が不等でも可

表 2 GRAHAM のデータに指数分布をあてはめた場合の実測値と理論値

期 間	再捕尾数	期待度数
1	139	138.42
2	91	89.16
3	52	57.43
4	40	36.99
Total	322	322.00

放流尾数; 1000 尾

状況に応じて適切な解析手法を選択する必要がある (表 1)。このことについては、この節の (3) で更に検討する。

また、打ち切り、非打ち切りにかかわらず、データが指数分布に従うと判断できるかどうかを検証することが必要である。前提となる理論の検証を経て初めて解析結果が利用できる。北田・平野 (1987) は、この検証の重要性を主張し、期間毎にまとめた再捕数の実測値と理論値を比較することで、モデルへのあてはまりの良さを示した。

ここで、解析方法の計算例でしばしば使われる GRAHAM (1938) のデータ (PAULIK, 1963, p. 38) を、北田・平野 (1987) に従って解析すると、実測値と理論値がほぼ良く一致しており (表 2)、このデータが指数分布で良く近似されることがわかる。FAREBROTHER (1988) は、このデータを用いて、 F 一定の仮定のもとに M を 3 種類の関数形 (一つは M 一定) で与え、彼が提案した方法で解析した。その結果、組合せの定数項を除いた対数尤度の最大値が 3 つの場合で殆んど変らない (-1038.778 , -1038.771 , -1038.683) ことから、 F も M も一定のモデルを採用することを勧めている。

なお、FAREBROTHER (1988) は 3 つのモデルで最大対数尤度が殆んど変らないことを興味深いと述べている。この考え方は、最大対数尤度が変わらないならパラメータの少ないモデルを採用しようというものであり、AIC につながるものである。

(2) データのまとめの影響

1 日毎のデータを使って解析する場合に比べて、期間毎にデータをまとめた場合には情報がそれだけ失なわれる。つまり推定の効率が悪くなる。この効率がどう変わるかを $\hat{Z}$ の漸近分散をみることで考察してみたい。

データをまとめる期間幅が等しい場合の尤度関数は具体的には、(3.42) 式から

$$L = m^{**}(1 - e^{-x_1 Z})^{n_0}(e^{-x_1 Z} - e^{-2x_1 Z})^{n_1} \dots \{e^{-(j-1)x_1 Z} - e^{-jx_1 Z}\}^{n_{j-1}}(1 - e^{-ZU})^{-n} \quad (2.1.1)$$

$\partial \log L / \partial Z$ の第 j 項の分母及び分子を $e^{-jx_1 Z}$ で除して

$$\frac{\partial \log L}{\partial Z} = -\frac{x_1 n}{e^{Zx_1} - 1} - \frac{nx_1 j}{e^{ZU} - 1} - x_1 \sum_{j=1}^{J-1} j n_j \quad (2.1.2)$$

これより $\hat{Z}$ の漸近分散

$$\begin{aligned} \text{Var}(\hat{Z}) &= \left[-E \left\{ \frac{\partial^2 \log L}{\partial Z^2} \right\} \right]^{-1} \\ &= \frac{1}{n} \left[\frac{x_1^2 e^{Zx_1}}{(e^{Zx_1} - 1)^2} - \frac{U^2 e^{ZU}}{(e^{ZU} - 1)^2} \right]^{-1} \end{aligned} \quad (2.1.3)$$

を得る。この漸近分散は KULLDORFF (1961) によって導かれている。

適当な Z の値に対して、期間幅 x_1 を 5~30 日の範囲で与えて $\text{Var}(\hat{Z})$ を計算すると、 Z の値にかかわらず、 x_1 が小さくなる程分散は小さくなる (図 2)。このことから、1 日毎のデータで推定するのが良いことがわかる。

しかし、計算に当っては計算機への入力が煩雑なものとなる。また、1 日毎のデータが得られない場合もある。従ってどの程度までのまとめ方なら許されるのかを検討しておくことは実用上重要である。そこで、ありそうな Z の範囲に対して、 x_1 を変えて $\text{Var}(\hat{Z})$ を計算し、 $x_1=1$ の $\text{Var}(\hat{Z})$ に対する比を計算した (表 3)。この値は相対的な推定効率を示すと考えられる。ここでは打ち切り時刻 $U=120$ とした。死亡係数は季節によって変化すると考えられるが、120 日は死亡係数を一定と仮定するには十分な期間であろう。また、前出の GRAHAM (1938) のデー

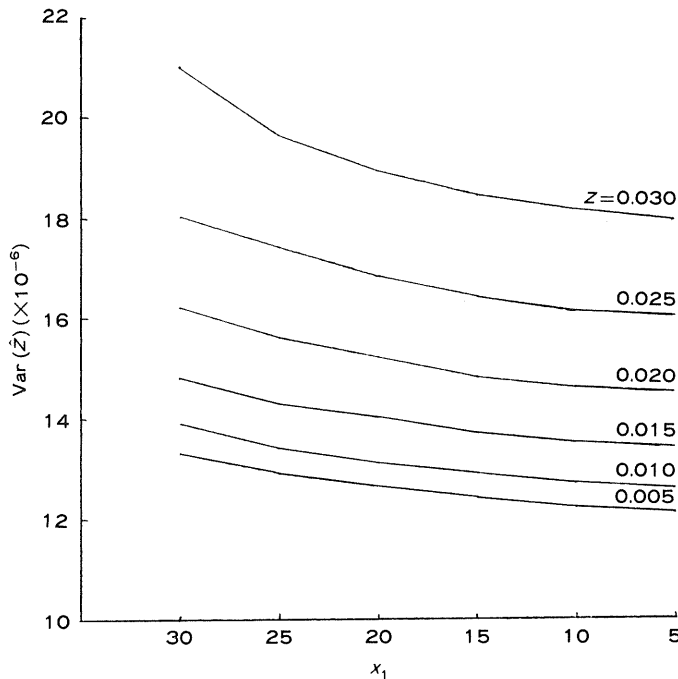


図 2 Z 及び期間幅 x_1 に対する $\hat{Z}$ の漸近分散

表 3 $U=120$ の場合の期間数 J に対する Z の相対的な推定効率*

Z	x_1						
	5	10	12	15	20	30	40
	J						
	24	12	10	8	6	4	3
.002	.9988	.9941	.9905	.9858	.9721	.9383	.8894
.004	.9964	.9906	.9882	.9825	.9700	.9355	.8862
.006	.9988	.9930	.9907	.9850	.9716	.9365	.8869
.008	.9989	.9932	.9899	.9842	.9710	.9356	.8844
.010	.9989	.9933	.9900	.9835	.9707	.9332	.8817

* $\{x_1=1$ の時の $\text{Var}(\hat{Z})\}/\{\text{各 } x_1 \text{ に対する } \text{Var}(\hat{Z})\}$

とも $U=120$ であり、比較の上でも便利であると考えた。表 3 から見て、期間数が 10 以上あれば、1 日毎のデータを使って推定した場合と殆んど同じ推定効率が得られると考えて良いだろう。

ただ、ここで得られた結果は、再捕数がランダムサンプルになっている時のものである。実際には魚のグループ行動のためにまとまって獲られたり、逆に獲られなかったりして 1 日毎のデータはかなりバラツキが大きいことが多い。このような時にはデータを何日か毎にまとめた方が推定精度が良くなる場合もあろう。この点については未検討であるが、前述の推定効率のことも考え併せるとデータを何日か毎にまとめて解析することは实际的である。

(3) 解析手法の選択

前述したように実際のデータ解析においては、殆んどの場合で打ち切りデータを扱うことになる。また、前節で見たように、期間数を 10 以上にするならば、それ程推定効率は落ちないし、データのバラツキを丸める効果も期待できる場合があるので、实际的であるということであった。従って、ここでは、打ち切りで期間毎のデータの場の解析方法に絞って考え方を述べる。この状況で適用可能な方法は、PAULIK (1963)、加重最小二乗法 (PAULIK 1963, FAREBROTHER 1985)、北田・平野 (1987)、FAREBROTHER (1988) の 4 つである。

PAULIK の方法では (3.30) 式を数値計算で解かなければならないが、この推定式が CHAPMAN AND ROBSON

(1960)と同じであるため、 $\hat{s}$ は ROBSON AND CHAPMAN (1961, Table 3, p. 187) の数表を利用して求めることができる。CHAPMAN AND ROBSON の推定方法は田中 (1985, p. 181) で解説されている。この場合、期間の数は $3 \leq J \leq 9$ である。北田 (1985) はこの数表を拡張し、 $10 \leq J \leq 19$ までを付け加えた。ただ、PAULIK の推定式 (3.30) は $\hat{s}$ が含まれているため J が大きくなると、 $\hat{s}$ が小さい場合には特に、 T/n の値が期間の数が増えても変わらないため、期間の数に制限を持っている。 $\hat{s}$ が 0.40 以下の場合では、 J が 10 以上に増えても T/n の値は殆んど変わらないので注意が必要である。ROBSON AND CHAPMAN がこの数表で J の上限を 9 とした理由はこの点にあるのであろう。従って、期間数を 10 以上にする場合においては、他の手法を使う方が無難である。PAULIK 自身も、期間数に注意しなくても良いため一般的には最小二乗法の使用を勧めている。

ただし、この log-linear model においては、当然 $n_j=0$ の場合は用いることはできない。また、PAULIK はこのことに加えて、いくつかの n_j が小さい場合も最小二乗法は使えないと述べており、最尤法による解析手法を勧めている。 n_j が小さい場合というのは、このモデル (2.5) の前提となっている $E(n_j) \geq 10$ と関係しているものと思われる。

従って、データをまとめた時に、 $n_j=0$ や小さい n_j を含む場合には、最尤法で解析するのが望ましい ($n_j=0$ がある場合は最尤法でしか解析できない)。

この状況においては、北田・平野 (1987) と FAREBROTHER (1988) によって解析できる。北田・平野の方法を用いる場合は、BASIC プログラム (北田 1987b, 北田・石岡 1988) が利用できる。FAREBROTHER では尤度関数を直接最大にするようにパラメータを定める強力な数値計算プログラムが必要となる。

最小二乗法が使えるデータの場合に、最尤法とどちらを使うべきかは不明である。これについて若干の検討を加えるために、GRAHAM のデータ (表 2) をそれぞれの方法で解析してみた (表 4)。通常の最小二乗法 (BEVERTON, 1954) による結果も参考のため併記したが、これ以外の方法による推定値はほぼ同じである。

最尤法ではどの方法による推定値も殆んど変わらない。PAULIK の方法と F, M 一定の場合の FAREBROTHER の方法は全尤度に基づく同じ推定方法であるが、 $\hat{Z}$ が小数点第 4 位で少し異っている。これは、PAULIK の方法では数表によって解を近似的に求めるためである。ここでは、全尤度 (FAREBROTHER) と部分尤度 (北田・平野) による推定値が完全に一致している点が注目される。

一方、加重最小二乗法による推定値は $\hat{Z}, \hat{F}$ と少し異っている。これは、最小二乗法では多項分布を正規近似して扱っていることによるものと思われる。一般に log-linear model と呼ばれる最小二乗法では、再捕数の経時データから先ず Z を推定し、 $\hat{Z}$ を用いて F を推定する。このときの $\hat{F}$ は、(2.3), (2.4) 式から

$$\hat{F} = \hat{Z} e^{\hat{\alpha}} / (1 - e^{-\hat{\alpha}}) \quad (2.1.4)$$

であるが、 $\alpha = \log\{F(1 - e^{-\alpha})/Z\}$ であるので、結果的に (2.1.4) 式は (3.33) 式に一致する。これらのことから、log-linear model (BEVERTON 1954, PAULIK 1963) と部分尤度に基づく推定方法 (北田・平野, 1987) は同一の procedure をとっており、本質的に同じものであることがわかる。ただ最小二乗法では多項分布を正規近似する点が異っている。モデル間の推定精度の比較は今後の問題であるが、最近、HIRAMATSU AND KISHINO (1989) によって部分尤度を用いる場合の $\hat{F}$ の漸近分散が導かれており、 $\hat{F}$ の分散が計算できるという観点に立てば北田・平野 (1987) を用いれば良い。HIRAMATSU AND KISHINO の $\hat{F}$ の漸近分散は、この場合

表 4 GRAHAM のデータの手法別の解析結果

パラメータ	最尤法 (全尤度)		最尤法 (部分尤度)	加重最小二乗法	最小二乗法
	PAULIK	FAREBROTHER*1	北田・平野	PAULIK	BEVERTON
Z	0.4397 [0.3377, 0.5534]	0.4399 —	0.4399 [0.3342, 0.5456]	0.4384 [0.2935, 0.5833]*2 (0.3338, 0.5430)*3	0.4296 [0.2856, 0.5736]*2
F	0.1711	0.1711	0.1711 [0.1554, 0.1866]	0.1709	0.1678

*1 F, M 一定の場合を論文中から引用した。この場合は PAULIK の方法と同じであるが、PAULIK の推定値が数表を使って求めているため値が少し異っている。

*2 t 分布による信頼区間 ($\hat{\sigma}^2=0.3983$)

*3 正規分布による信頼区間 (FAREBROTHER による、 $\sigma^2=1$ とした場合)

$$\text{Var}(\hat{F}) = \left\{ \left(-\frac{\partial^2 \log L_2}{\partial F^2} \right)^2 \right\}^{-1} \left\{ -\frac{\partial^2 \log L_2}{\partial F^2} + \left(-\frac{\partial^2 \log L_2}{\partial Z \partial F} \right)^2 \text{Var}(\hat{Z}) \right\} \quad (2.1.5)$$

として求まる。

ここで

$$\frac{\partial^2 \log L_2}{\partial F^2} = -\frac{Nn}{F^2(N-n)}, \quad (2.1.6)$$

$$\frac{\partial^2 \log L_2}{\partial Z \partial F} = -\frac{N^2(ZUe^{-ZU} + e^{-ZU} - 1)}{Z^2(N-n)} \quad (2.1.7)$$

であり、 F , Z に $\hat{F}$, $\hat{Z}$ を用いて計算する。これによって、期間毎の打ち切りデータでは初めて $\hat{F}$ の分散が推定できるようになった。このデータの場合、 $\widehat{\text{Var}}(\hat{F}) = 7.0530 \times 10^{-8}$, (2.1.3) 式より $\widehat{\text{Var}}(\hat{Z}) = 3.2332 \times 10^{-6}$ である。

なお、このデータの場合、 $J=4$, $\hat{\sigma}^2=0.3983$ であるから、 $(J-2)\hat{\sigma}^2=0.797$ である。有意水準 5% の両側仮説検定の $Z^2(J-2)$ の採択域は (0.05, 7.38) である。0.797 はこの中に入るので仮説 $\sigma^2=1$ は棄却されない。

2. 死亡係数の変化とその推定

漁業には周期がある。魚の側からみても漁場内に多く加入してくる時期は決まっており、また、人間の方も魚が多い時には多くの漁船が出る。沿岸漁業においては多様な魚種を対象に、その時どの魚を獲るのが最も利益が得られるかを考えて魚種を選択している。図3は広島県豊島地区のマダイの月別漁獲量、出漁隻数、CPUE（高場ら、1986）のコレログラムである。生データは季節的な変動をしているが、図3を見るときれいな1年周期があることがわかる。一般的に、魚の側からも、それを利用する人間側からも、漁業は1年という周期を持っていると考えることは妥当であろう。このような漁業の中で死亡係数もまた、周期的な変化をしていると考えられる。

もし、毎年の漁獲量が季節に応じて変動していても、その強さが年々変わらないのであれば、この1年という周期でデータをプールすれば良い。何年間かのデータがあれば、それは指数分布に従うはずである。このデータから死亡係数を推定すれば、周期的変化をしている死亡係数の平均値が推定される。しかし、年齢によって漁獲量が変わる場合は、データを1年でプールしても、数年分のデータのヒストグラムが指数分布から外れることになる。この場合は、死亡係数は年齢によって変るステップ関数となる。この状況は本質的にコホート解析のそれと同じである。死亡係数を時刻の関数として推定する意味があるのは、このように数年分のデータがある場合であろう。あるいは、季節毎にステップする死亡係数を仮定して推定することも実際的で意味のあることであろう。

ただ、標識放流の場合は、再捕の報告もれや標識の脱落が主な原因で、長い期間にわたって正確なデータを得ることが極めて困難である。従って、再捕データの解析は短期間にならざるを得ない。

しかし、理論的にはこれらの状況については、FAREBROTHER (1988) の用いた尤度で対応可能である。

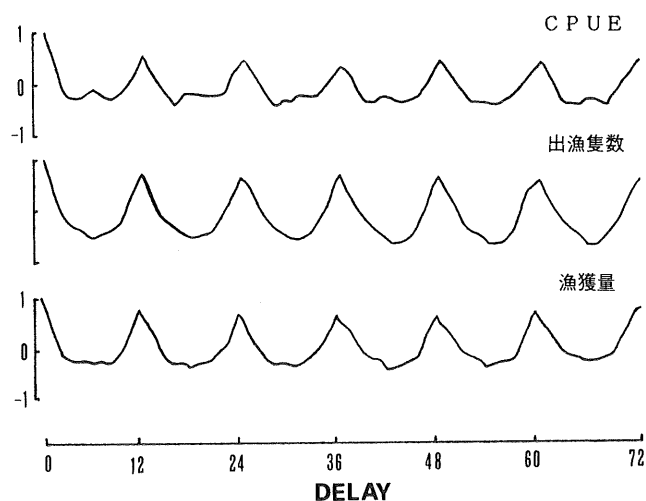


図3 大崎下島、豊浜町漁協における月別マダイ漁獲量、出漁隻数、CPUE のコレログラム
(高場ら 1986 のデータから計算)

なお、HSH 法 (1987) は、色々な $F(t)$ に対してその形が解らなくても robust な M の推定値が求められるが、 $F(t)$ の関数形を推定する訳ではない。また、このモデルの「推定時点ですべての標識魚は海の中に生存していない」という特徴的な仮定は多くの場合で満たされないであろう。このモデルは非打ち切りのモデルであるが、LEIGH は HSH 法の打ち切りの場合を考察した。この条件は、打ち切り時点での海や池の中での残存尾数がわかっていることであり、特に海では困難である。

お わ り に

ここで扱った理論では、合計 n 尾の再捕尾数が同時に得られる確率を多項分布で与える。この問題は漁獲尾数の解析 (コホート解析) と本質的に同じであり、特に死亡係数が時刻の関数である時には FAREBROTHER によって一般化された再捕の尤度が、今後我が国においても注目されよう。この場合、尤度関数を直接最大にするようパラメータを定めるための最適化による強力な数値計算プログラムが必要となる。平松*は、この尤度によって、コホート解析、標識再捕法その他 DE LULY 法も統一的に扱えるとしている。

時刻 x_j から x_{j+1} の間で Z_j, F_j が一定なら、部分尤度を用いる方法を拡張できる。この場合、死亡係数の変化点の推定も可能であるが、このことについては別途詳述することとしたい。

全尤度、部分尤度、最小二乗法による推定精度の比較も興味深い課題である。またモデル選択については AIC (Akaike Information Criterion; AKAIKE 1973) を用いて判断すれば良い。AIC については、坂本・石黒・北川 (1983) などが参考となろう。

なお、実際の標識放流実験においては、標識装着から放流までのハンドリングによる影響、標識の脱落、再捕の報告もれ等、推定値に影響を与える原因が多い。これらについては従来からも比較的意識されてきたが、特に標識個体群のグループ行動に今後注意を払うべきである。実際には、調査対象漁場内に分散放流することが望ましい。また、報告もれについては、今まで考えられている以上に大きく、通常のポスター等による協力依頼では解析に耐え得るデータをとることは難しい。ここで述べたような解析を行おうとするならば、初期のハンドリングによる死亡や標識の脱落に十分配慮することに加えて、漁場内への分散放流の実施やデータ収集体制の見直しを行う必要がある。しかし、報告もれを完全になくすことは難しいだろう。ただ、報告もれは F の推定値に影響を与えるが、いつも報告率が一定と仮定できるなら Z の推定には影響しない。従って、 Z は推定可能であるが、報告もれがあると、 F は過小にそして M は過大に推定される。

謝 辞

本稿をとりまとめるに当たり、貴重な助言を賜った統計数理研究所岸野洋久氏、遠洋水産研究所平松一彦氏、日本海海区水産研究所赤嶺達郎氏、東京大学海洋研究所松宮義晴氏に深謝する。また、統計数理研究所平野勝臣氏及び日本栽培漁業協会須田 明氏には懇切な指導や励ましをいただいた。記して謝意を表したい。

文 献

- 赤嶺達郎 (1988) 1 回放流・多回再捕による生残率の推定方法 (レビュー). 日本海ブロック試験研究集録, 13: 17-38.
- BEVERTON, R. J. H. (1954) Notes on the use of theoretical models in the study of the dynamics of exploited fish populations, Misc. Contr. Biol. Lab. Beaufort, No. 2, 159 pp.
- DEEMER, W. L. Jr. and VOTAW, D. F. Jr. (1955) Estimation of parameters of truncated or censored exponential distributions. *The Annals of Mathematical Statistics*, 26: 498-504.
- FAREBROTHER, R. W. (1985) Weighted least-squares estimates of mortality rates from single-release tagging studies. *J. Cons. int. Explor. Mer.*, 42: 166-170.
- FAREBROTHER, R. W. (1988) Maximum likelihood estimates of mortality rates from single-release tagging studies. *J. Cons. int. Explor. Mer.*, 44: 229-234.
- GULLAND, J. A. (1955) On the estimation of population parameters from marked members. *Biometrika*, 42: 269-270.
- HEARN, W. S., SANDLAND, R. L. and HAMPTON, J. (1987) Robust estimation of the natural mortality rate

* 平松一彦 (私信)

- in a completed tagging experiment with variable fishing intensity. *J. Cons. int. Explor. Mer.*, **43**: 107-117.
- 平松一彦 (1989) 標識放流による死亡率の推定——新しい方法の紹介. 遠洋水研ニュース, No. 74; 5-6.
- HIRAMATSU, K. and KISHINO, H. (1989) Stock identification by maximum likelihood method: Derivation of variance estimator. (Document submitted to the Annual Meeting of the International North Pacific Fisheries Commission, Seattle, U.S.A., November 1989.): 14 pp.
- 伊藤嘉昭・村井 実 (1977) 動物生態学研究法. 古今書院, 東京, 160-164.
- 北田修一 (1985) 標識放流再捕データからパラメータを推定する方法, その理論と応用の留意点, 1 回放流連続再捕の場合. 日裁協研資, **28**: 20 pp.
- 北田修一 (1987a) 標識再捕による放流効果の評価. 資源評価のための数値解析 (嶋津靖彦編), 恒星社厚生閣: pp. 102-117.
- 北田修一 (1987b) 打ち切り型の標識再捕データから死亡係数を推定する BASIC プログラム. 栽培技研, **16**(2): 133-140.
- 北田修一・平野勝臣 (1987) 期間毎の標識再捕に基づく死亡係数の推定について. 日水誌, **53**: 1765-1770.
- 北田修一・石岡清英 (1988) 打ち切り型の標識再捕データからの F , M の推定. パソコンによる資源解析プログラム集, 東海区水産研究所数理統計部編: 73-80.
- KITAHARA, T. *et al.* (1986) Estimation of fishing mortality coefficient from tag recoveries in a finite period with a few closed seasons. *Nippon Suisan Gakkaisi*, **52**: 579-584.
- KULLDORFF, G. (1961) Estimation from grouped and partially grouped samples. John Wiley and Sons, Inc., New York, pp. 55-63.
- LEIGH, G. M. (1988) A comparison of estimates of natural mortality from fish tagging experiments. *Biometrika*, **75**: 347-353.
- 能勢幸雄 (1971) 標識放流解析の理論. 標識放流による資源解析に関するシンポジウム議事録, 西海区水研: 1-12.
- PAULIK, G. J. (1963) Estimates of mortality rates from tag recoveries. *Biometrics*, **49**: 28-57.
- ROBSON, D. S. and CHAPMAN, D. G. (1961) Catch curves and mortality rates. *Trans. Am. Fish. Soc.*, **90**: 181-189.
- 坂元慶行・石黒真木夫・北川源四郎 (1983) 情報量統計学. 共立出版, 東京, pp. 236.
- SANDLAND, R. L. (1982) Estimation, inference, and data analysis for log-linear regression models in tagging studies. *J. Cons. int. Explor. Mer.*, **40**: 291-303.
- 高場稔ら (1986) 豊島地区における放流マダイの漁獲量. 広島水試研報, **16**: 1-18.
- 高橋宏一・脇本和昌・平野勝臣 (1975) 統計学要論. 共立出版, 東京, pp. 100-101.
- TANAKA, S. (1967) Estimation of fishing coefficient of Mojako by tagging experiments on drifting seaweeds—I. Method and example. *Nippon Suisan Gakkaisi*, **33**: 1108-1115.
- 田中昌一 (1985) 水産資源学総論. 恒星社厚生閣, 東京, pp. 181-183.