

1 回放流・多回再捕による生残率の推定方法 (レビュー)

赤 嶺 達 郎

(日本海区水産研究所)

放流実験の目的は以下の 3 つに大別される (石岡 1987)。

- ① 移動・回遊経路の把握
- ② 放流時点の天然群の量の推定
- ③ $F \cdot M$ の推定

栽培漁業における種苗放流の効果判定では③が特に重要である。②については久野 (1986), 加藤 (1987) を参考にするとよいが, 抽出法および Petersen 法についてベイズ統計からのアプローチを検討中 (連絡ニュースに連載) である。著名な教科書としては SEBER (1982) がある。

この小論では③を中心に解説する。②との違いは指数関数をモデルとして仮定している点である。現在の水産資源解析学は BARANOV (1918) の漁獲方程式に基礎をおいており, TANAKA (1967) の方法がよく使われてきた。最近, 北田 (1985, 1987 a, 1987 b), 北田・平野 (1987) により PAULIK (1963) 及び KULLDORFF (1961) の方法 [KULLDORFF は $\theta = 1/Z$ を推定したのみで, $F \cdot M$ の分離は北田・平野 1987 による] が使用可能となった。この 2 手法は TANAKA の方法を区間データに応用したものである。一方, KITAHARA *et al* (1986) は休漁期を含むモデルに拡張した。これらの手法について解説を加える。

まず準備として基礎的事項について述べる。既に理解されている方は読みとばして, 2 の各論に進んでいただきたい。

1. 基礎的事項

(1) 離散モデルと連続モデル

同一年級群の資源尾数は指数関数的に減少することが経験的に知られている。指数関数モデルは BARANOV (1918) によって導入され, BEVERTON and HOLT (1957) において体系としてまとめられた。1960 年前後には様々な手法が開発され, まさに黄金期であった。

モデルの表現として

① 等比数列モデル

$$N_n = N_0 s^n \quad (n = 0, 1, 2, \dots) \quad (1.1)$$

② 指数関数モデル

$$N_t = N_0 e^{-Zt} \quad (1.2)$$

の 2 通りがある。①は離散モデル, ②は連続モデルである。両式より $n = t$ として

$$s = e^{-Z} \quad (1.3)$$

である (図1)。

一般に s を減少率, Z を全死亡係数と呼ぶ。しかし, 種苗放流実験では死亡よりも逸散の占める割合が大きく, このような名称は誤解を招きやすい, むしろ s を縮小率 [通常は $1-s$ を減少率と呼ぶ], Z を全減少係数とも呼んだ方がよい。

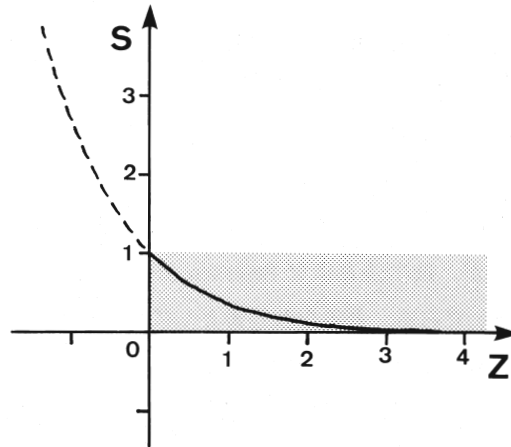


図1 生存率 (s) と全死亡係数 (Z) の関係 ($s = e^{-Z}$)

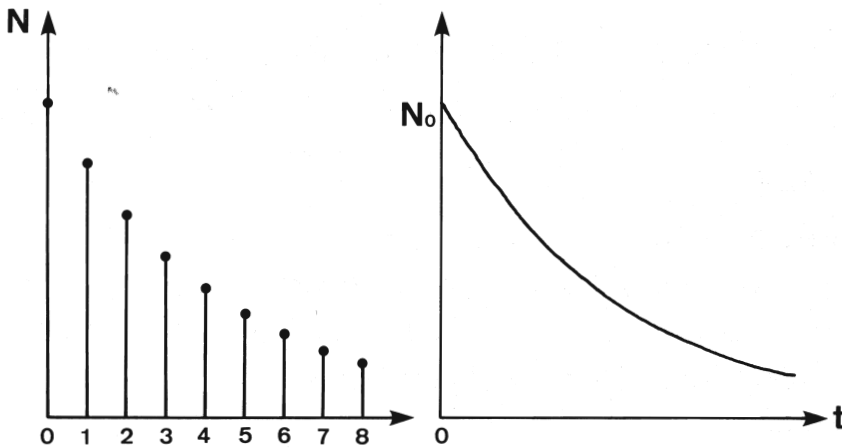


図2 等比数列モデル (左) と指数関数モデル (右)

もともと指数関数は等比数列の拡張なので、②は①の拡張と考えてよい(図2)。しかし、数学的に①と同等な連続モデルは以下の式で与えられる。

$$\frac{d(\ln N)}{dt} = \frac{1}{N} \frac{dN}{dt} = -Zf(t),$$

$$f(t+1) = f(t) \geq 0 \quad (1.4)$$

これを解くと

$$N_t = N_0 e^{-Z \{F(t) - F(0)\}},$$

$$F(t) = \int f(t) dt \quad (1.5)$$

となる。 $f(t)$ は周期1の周期関数であるが、次の基準化

$$\int_0^1 f(t) dt = 1 \quad (1.6)$$

を行った方がよい。この拡張はAKAMINE(1986)の成長式の拡張のアナロジーである。代表例を図3に示す。 t を年と考えた場合、(1.5)の方が(1.2)よりも現実的だろう。

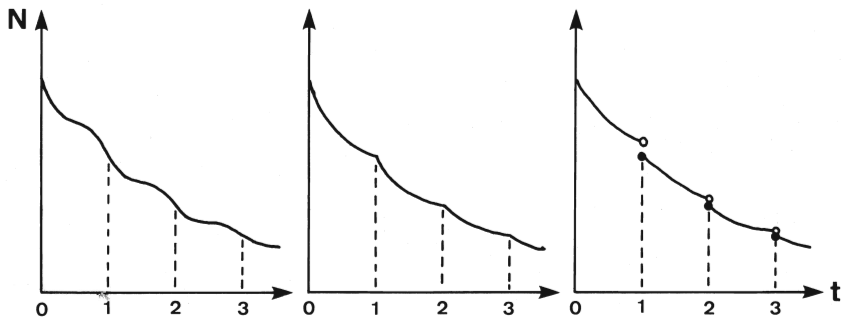


図3 等比数列モデルの一般例

自明な仮定であるが『 N は単調減少』である。つまり

$$x < y \text{ のとき } N_x \geq N_y \quad (1.7)$$

となっている。また s 、 Z は定数だから時間(横軸)の起点のとり方には影響されない。しかし、時間の単位には影響される。例えば1年=365日とすると、

$$s(\text{年}) = s(\text{日})^{365}, \quad Z(\text{年}) = 365 Z(\text{日})$$

である。

(2) 確率と確率密度

統計学においては離散モデルと連続モデルでは扱いが異なる。簡単な例で示すと、『サイコロで6の目が出る確率は $1/6$ 』であるが、『身長が160cmの人の確率は0』である。つまり身長が160.0000……cmという人は存在しない。『身長が160cm以上161cm未満の人の確率は……』というように範囲(区間)を与えなくてはならない。

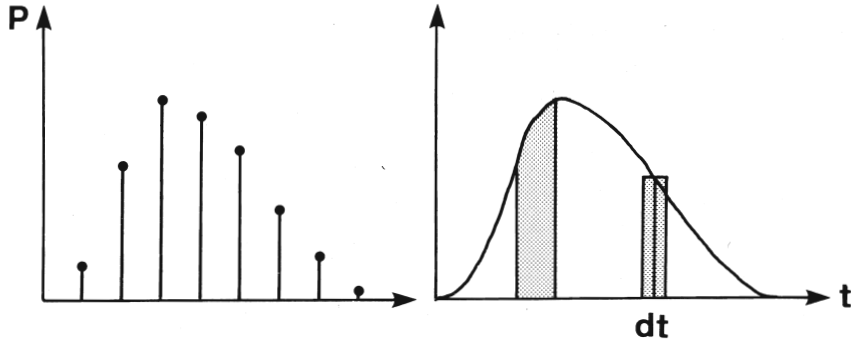


図4 離散モデル（左）と連続モデル（右）

これは分布図で示すと理解しやすい。図4よりわかるように

① 離散モデルでは『高さ』が確率を表わす。

② 連続モデルでは『面積』が確率を表わす。

のである。全体の確率は1であるが、それぞれ

$$\sum_i P_i = 1 \quad (1.8)$$

$$\int_{-\infty}^{\infty} P_i dt = 1 \quad (1.9)$$

と表現される。

連続モデルにおいて微小区間 dt における確率は $P_i dt$ であるが、 P_i を確率密度と呼ぶ。つまり確率密度は『高さ』である。連続モデルでは確率（面積）のかわりに確率密度（高さ）を用いて計算することが多い。二項分布を正規分布で近似する場合には、確率密度→確率という変換が行われている訳である。

(3) ヒストグラム

ヒストグラムは棒グラフとは異なる。棒グラフは高さが度数（確率）を表わすが、ヒストグラムは面積が度数（確率）を表わす。従って棒グラフは離散モデルに、ヒストグラムは連続モデルに適している。

ヒストグラムの各区間の巾は等しくなくてもよいが（図5）、通常は等区間にした方が扱いやすい。ヒストグラムでは各区間（階級）の中央の値（横軸）をもってその区間を代表し、これを階級値と呼ぶ。従って各データに最大で $h/2$ の誤差が加わる（ h は階級の巾）。これは巾 h で一様分布していると考えてよいので、誤差の分散は

$$\langle \Delta x^2 \rangle = \frac{h^2}{12} \quad \langle \rangle : \text{期待値}$$

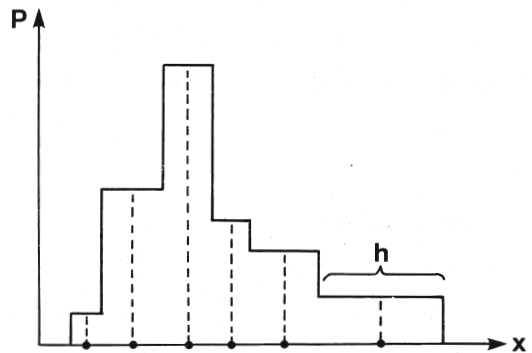


図5 ヒストグラムの例（・は階級値）

となる。等区間でない場合は最大区間の巾を $\max h$ とすれば

$$\langle \Delta x^2 \rangle \leq \frac{(\max h)^2}{12}$$

である。

(4) 最尤法 (さいゆうほう)

パラメーター推定における有力な手法が最尤法である。重要な手法であるにもかかわらず、やさしく解説した教科書が少ない。

いまデータ x_1, x_2, x_3 が与えられたとする [例えば魚の体重 100, 120, 130 g]。これらは A 群か B 群のどちらかに属することがわかっていて、A, B どちらかに決定する必要がある。A, B 群の確率 (密度) 分布は知られていて、それぞれ $P_A(x), P_B(x)$ とする (図 6)。各々のデータ、例え

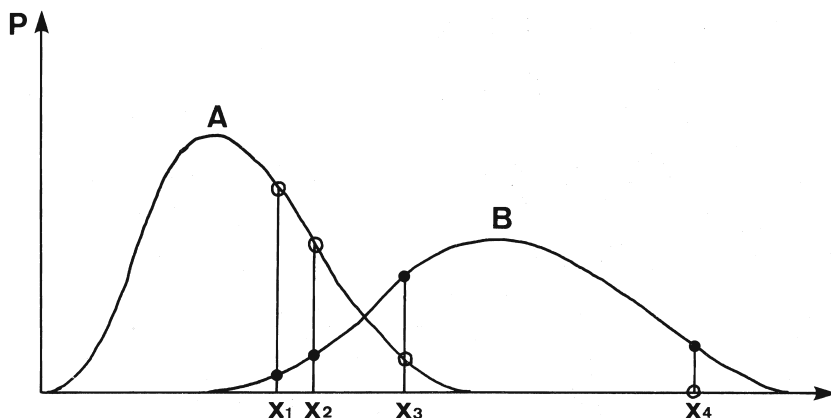


図 6 最尤法の説明図

ば x_1 についてみると $P_A(x_1) > P_B(x_1)$ だから、 x_1 は A に属すると考えた方がよい。しかし、ここで問題にしているのは $x_1 \sim x_3$ が同時に A または B のどちらに属するかという事である。例えば

$$Y_i = P_i(x_1) + P_i(x_2) + P_i(x_3)$$

とおいてみる。このとき $Y_A > Y_B$ だから $x_1 \sim x_3$ は A に属すると判断してよさそうである。しかし、 x_4 が図 6 のように与えられた場合でも $Y_A > Y_B$ となって $x_1 \sim x_4$ は A に属すると判断されるが、これは不合理である。そこで

$$L_i = P_i(x_1) \cdot P_i(x_2) \cdot P_i(x_3)$$

を考えてみると、 $x_1 \sim x_3$ については $L_A > L_B$ となって A に属すると判断され、 x_4 を加えた場合には $P_A(x_4) \doteq 0$ だから $L_A < L_B$ となるため、 $x_1 \sim x_4$ は B に属すると判断され合理的である。 $x_1 \sim x_3$ がそれぞれ独立の場合 [x_1 と x_2 は夫婦で、♀の方が♂より必ず重いというような場合は独立ではない]、 L は同時確率 (密度) であるが、これを尤度 (ゆうど) と呼ぶ。つまり

$$\text{尤度} = \begin{cases} \text{確率の積 (離散モデル)} \\ \text{確率密度の積 (連続モデル)} \end{cases}$$

と考えて差し支えない。

一般には確率（密度）分布のタイプ（例えば正規分布）がわかっている、パラメータの値が不明な場合が多い。従って L を最大にするパラメータの値を採用する。これが最尤法である。

$L = \prod P_i$ は 0 に近い値になって計算しにくいので、対数尤度 $\ln L = \sum \ln P_i$ を使うことが多い（これは負の大きな値になる）。つまりパラメータを θ とすると

$$\frac{\partial (\ln L)}{\partial \theta} = \frac{1}{L} \frac{\partial L}{\partial \theta} = 0 \quad (1.10)$$

を満足する θ を真の θ と推定する。

正規分布の例を下に示す。確率密度分布は

$$P_i = N(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right\}$$

で与えられる。データが $x_1 \sim x_m$ のとき対数尤度は

$$\begin{aligned} Y = \ln L &= \sum_{i=1}^m \ln P_i \\ &= \sum_{i=1}^m \left\{ -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \left(\frac{x_i - \mu}{\sigma}\right)^2 \right\} \end{aligned}$$

となる。パラメータは μ と σ の 2 つだから

$$\frac{\partial Y}{\partial \mu} = \sum_{i=1}^m \frac{x_i - \mu}{\sigma^2} = 0 \quad (1.11)$$

$$\frac{\partial Y}{\partial \sigma} = \sum_{i=1}^m \left\{ -\frac{1}{\sigma} + \frac{(x_i - \mu)^2}{\sigma^3} \right\} = 0 \quad (1.12)$$

である。これより

$$\mu = \frac{\sum x_i}{m} = \bar{x} \quad (\text{算術平均}) \quad (1.13)$$

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{m} \quad (2 \text{ 次モーメント}) \quad (1.14)$$

を得る。ここで (1.13) を (1.14) に代入すれば

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{m} \quad (1.15)$$

となる。(1.13) を標本平均、(1.15) を標本分散と呼ぶ。これが最尤法の解である。

ここで多少面倒な（理解しにくい）問題が生じる。

$$\begin{aligned} \sum (x_i - \mu)^2 &= \sum (x_i - \bar{x} + \bar{x} - \mu)^2 \\ &= \sum (x_i - \bar{x})^2 + 2(\bar{x} - \mu) \sum (x_i - \bar{x}) + \sum (\bar{x} - \mu)^2 \\ &= \sum (x_i - \bar{x})^2 + m(\bar{x} - \mu)^2 \end{aligned}$$

であるが、両辺の期待値を計算すると

$$\langle \sum (x_i - \mu)^2 \rangle = \sum \langle (x_i - \mu)^2 \rangle = m \sigma^2$$

$$\langle m (\bar{x} - \mu)^2 \rangle = m \left\langle \frac{\sum (x_i - \mu)^2}{m^2} \right\rangle = \sigma^2$$

となるので、結局

$$\langle \sum (x_i - \bar{x})^2 \rangle = m \sigma^2 - \sigma^2 = (m - 1) \sigma^2$$

$$\therefore \sigma^2 = \frac{\sum (x_i - \bar{x})^2}{m - 1} \quad (1.16)$$

を得る。これが不偏分散である。

真の μ , σ は母集団のパラメータなので推定不能である。(1.14) は推定不能な μ を右辺に含んでいる。(1.15) は μ の代わりに推定値 \bar{x} を用いた訳であるが、そのために偏りが生じ、結局 (1.16) のように補正が必要となると解釈される。

これは立場上の問題であって、現代統計学は最尤性よりも不偏性の方をより重要視するため、(1.15) ではなくて (1.16) を採用する訳である。つまり最尤法は推定量を構成する 1 つの手法であると位置づけるのが適当で、最尤法によって得られた推定量が望ましい性質をもっているかどうかは個別に調べるという立場をとらなくてはならない (鷲尾1978) ということになる。

しかし、最尤法より有力な推定法はまだ知られていないし、例えば多重正規分布

$$P_i = \prod_{k=1}^n K_k N(\mu_k, \sigma_k, x_i), \quad \sum_k K_k = 1$$

のようなモデルにおいて、最尤法で得られたパラメータの推定値について期待値を計算することは困難であり、また無理に行うだけの意義もない。[$m \rightarrow$ 大のとき (1.15) \simeq (1.16) であり、応用上は最尤法で十分である]。

なお、最小 2 乗法は残差

$$d_i = \frac{y - y_i}{\sigma_i}, \quad d_i^* = y - y_i$$

がそれぞれ正規分布 $N(0, 1)$, $N(0, \sigma)$ に従うという仮定で、最尤法を用いて推定する手法である。

(5) 漁獲方程式

BARANOV (1918) によって導入された漁獲方程式は以下のように定義できる。

$$\left\{ \begin{array}{l} \frac{dN}{dt} = -ZN \end{array} \right. \quad (1.17)$$

$$\left\{ \begin{array}{l} \frac{dC}{dt} = FN \end{array} \right. \quad (1.18)$$

N は資源尾数, C は累積漁獲尾数である。 Z, F は定数でそれぞれ全死亡係数, 漁獲死亡係数と呼ばれる。一般に

$$Z = F + M \quad (1.19)$$

として, M を自然死亡係数と呼ぶが, 前述のように種苗放流実験では大部分を逸散が占める場合が多い。(1.18) ÷ (1.17) より

$$\frac{dC}{dN} = -\frac{F}{Z}$$

従って

$$dC = -\frac{F}{Z}dN \quad (1.20)$$

近似して

$$\Delta C = -\frac{F}{Z}\Delta N \quad (1.21)$$

が成立する。つまり資源尾数の減少した量 ($-\Delta N$) のうち, 一定の比率 (F/Z) の量が常に漁獲される訳である。

(1.17), (1.18) は連立微分方程式 (自励系) であるが, 簡単に解ける。初期条件を

$$t = 0 \text{ のとき } N = N_0, C = C_0 = 0 \quad (1.22)$$

とすると

$$N_t = N_0 e^{-Zt} \quad (1.23)$$

$$C_t = \frac{F}{Z} N_0 (1 - e^{-Zt}) \quad (1.24)$$

となる。

ここで (1.4) と同じ拡張を考えてみる。つまり (1.17), (1.18) において

$$Z \rightarrow Zf(t), F \rightarrow Fg(t)$$

$$f(t+1) = f(t), g(t+1) = g(t)$$

$$\int_0^1 f(t) dt = 1, \int_0^1 g(t) dt = 1$$

とおく。このとき

$$N_t = N_0 e^{-Z \int_0^t (f(t) - F(0))} \quad (1.25)$$

$$C_a = F \int_0^a g(t) N_t dt \quad (1.26)$$

となる。

$$F(t+1) = F(t) + 1$$

であるから

$$N_{t+1} = e^{-Z} N_t \quad (1.27)$$

となる。ここで

$$\Delta N_a = \int_a^{a+b} \left(\frac{dN}{dt} \right) dt, \quad \Delta C_a = \int_a^{a+b} \left(\frac{dC}{dt} \right) dt$$

とすれば、(1.27) より

$$\Delta N_{a+1} = e^{-Z} \Delta N_a, \quad \Delta C_{a+1} = e^{-Z} \Delta C_a \quad (1.28)$$

となる。これはほとんど自明であるが、図7及びカヴァリエリの原理(図8)より容易に理解できる。

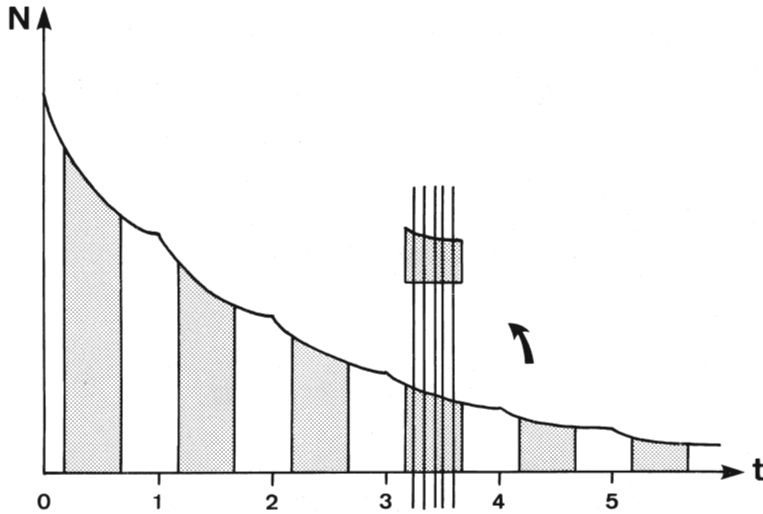


図7 拡張した漁獲モデル

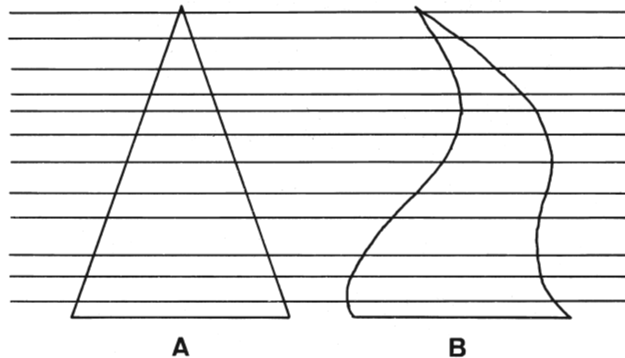


図8 カヴァリエリの原理(面積はA=B)

(1.28) より $s = e^{-Z}$ が直ちに求まる。これはかなり重要な注意事項である。つまり本来の漁獲方程式(1.17), (1.18) は Z , F が常に一定と仮定しているの、漁業が常に同じ状態で行

われていないと適用できない。しかし、(1. 28) より例えば1年間における漁業の変化が毎年同じであれば s (年) を求めることができる。また

$$\frac{\Delta C_{a+1}}{\Delta N_{a+1}} = \frac{\Delta C_a}{\Delta N_a} = \text{一定}$$

であるが、(1. 21) と違って $(-F/Z)$ となるとは限らないため、 F と M の分離はそれぞれの場合ごとに (1. 26) をきちんと計算しなくてはならない。

以上のように s が容易に推定できるのは、指数関数モデルにおいて区間を等間隔にとった場合、高さも面積も同じ比の等比数列になるという性質によっている。

2. 各 論

(1) TANAKA (1967) の方法

GULLAND (1955) が既に同じ結果を導いているが、導き方は TANAKA (1967) の方がずっとスマートである。両者とも $t \rightarrow \infty$ としているが、ここでは $t \rightarrow u$ として有限区間として扱う。まず (1. 24) より総面積は

$$C_u = \frac{F}{Z} N_0 (1 - e^{-Zu}) = \frac{F}{Z} N_0 (1 - s^u) \quad (2. 1)$$

となるから、確率密度関数は

$$P(t) = \frac{dC/dt}{C_u} = \frac{Ze^{-Zt}}{1 - e^{-Zu}} \quad (2. 2)$$

となる。

生データが $t_1 \sim t_m$ で与えられたとき、対数尤度は (2. 2) より

$$\begin{aligned} Y &= \sum_{i=1}^m \ln P(t_i) \\ &= m \ln Z - Z \sum t_i - m \ln (1 - e^{-Zu}) \end{aligned}$$

となる。従って Z の推定量は

$$\frac{\partial Y}{\partial Z} = m \left(\frac{1}{Z} - \frac{u}{e^{Zu} - 1} \right) - \sum t_i = 0 \quad (2. 3)$$

を Z について解けばよい。

一方、データがヒストグラム (t_i, F_i) の組として与えられた場合は

$$\begin{aligned} Y &= \sum F_i \ln P(t_i) \\ &= (\sum F_i) \ln Z - Z \sum F_i t_i - (\sum F_i) (1 - e^{-Zu}) \end{aligned}$$

となる。従って Z の推定量は

$$\frac{\partial Y}{\partial Z} = (\sum F_i) \left(\frac{1}{Z} - \frac{u}{e^{Zu} - 1} \right) - \sum F_i t_i = 0 \quad (2. 4)$$

を Z について解けばよい。

(2. 4) は (2. 3) を形式上含んでいる。

$$F_i \equiv 1 \text{ のとき } \sum F_i = m$$

となるからである。 $u \rightarrow \infty$ のときは通常の TANAKA の方法となる。このとき (2. 3), (2. 4) においてテイラー展開より

$$\frac{u}{e^{Zu} - 1} = \frac{1}{Z + \frac{Z^2}{2}u + \dots} \xrightarrow{u \rightarrow \infty} 0$$

となるため、それぞれ

$$\frac{1}{Z} = \frac{\sum t_i}{m} = \bar{t} \quad (2. 3')$$

$$\frac{1}{Z} = \frac{\sum F_i t_i}{\sum F_i} = \bar{t} \quad (2. 4')$$

となる。つまり Z は平均時間 \bar{t} の逆数で与えられる。平均時間が長くなれば Z は小さくなる訳で、(1. 17) より当然である。

ところでこれらの式は \bar{t} の計算からも導ける。(2. 2) より

$$\bar{t} = \int_0^u t P(t) dt = \frac{Z}{1 - e^{-Zu}} \int_0^u t e^{-Zt} dt$$

ここで右辺の積分を部分積分で行うと

$$\begin{aligned} \int t e^{-Zt} dt &= -\frac{1}{Z} t e^{-Zt} + \frac{1}{Z} \int e^{-Zt} dt \\ &= -\frac{1}{Z} e^{-Zt} \left(t + \frac{1}{Z} \right) \end{aligned}$$

となる。従って

$$\bar{t} = \frac{1}{Z} - \frac{u}{e^{Zu} - 1} \quad (2. 5)$$

を得る。これは (2. 3), (2. 4) と一致する。

(2. 5) を Z について解くのは、変数が Z ただ 1 つなので簡単である。ニュートン法がよいが、二分法等でも十分であろう。

(2) PAULIK (1963) の方法

この手法は CHAPMAN and ROBSON (1960), ROBSON and CHAPMAN (1961) と同じである。さらに結論を先に言うと、TANAKA の方法を区間データ (ヒストグラム) にあてはめたものである。なお一般論は田中 (1985) の P180~ に詳しい。

まず s を求める等比数列モデルを考える。区間 $[i, i+1]$ での漁獲尾数を ΔC_i とおく (図9)。

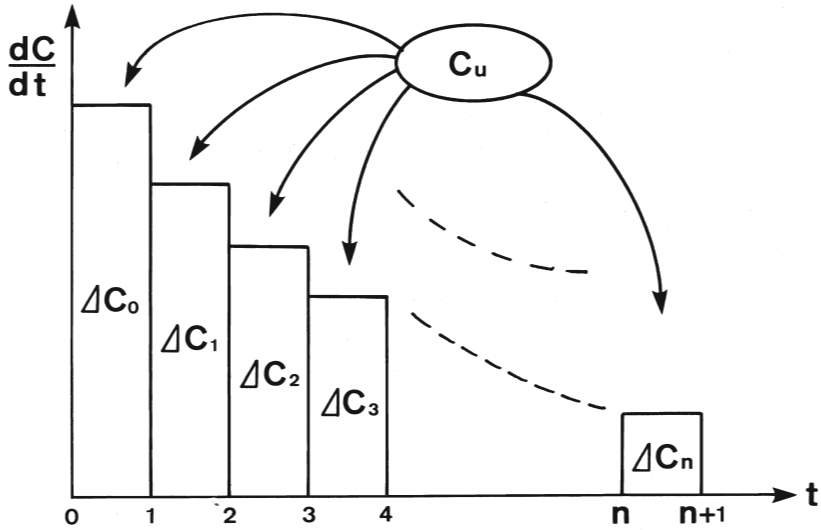


図9 PAULIK のモデル

このとき (1. 24) より

$$\Delta C_i = \frac{F}{Z} N_0 s^i (1-s) \quad (2. 6)$$

とおく。 $\Delta C_i = C_{i-1} - C_i$ と比較すると i が1つずれているが、 s は時間の起点のとり方には影響されないので問題ない。

$t = u = n + 1$ までの総漁獲尾数は

$$C_u = \sum_{i=0}^n \Delta C_i = \frac{F}{Z} N_0 (1-s) (1+s+\dots+s^n) \quad (2. 7)$$

となる。従って ΔC_i にはいる確率 P_i は

$$P_i = \frac{\Delta C_i}{C_u} = \frac{s^i}{1+s+\dots+s^n} \quad (2. 8)$$

である。

各区間ごとの漁獲尾数が F_0, F_1, \dots, F_n (ヒストグラム) で与えられたとき、尤度は、

$$L = \prod_{i=0}^n P_i^{F_i} \quad (2. 9)$$

である。従ってこれを最大にする s が最尤推定値である。(2. 8), (2. 9) より

$$\begin{aligned} \frac{\partial Y}{\partial s} &= \frac{\partial (\ln L)}{\partial s} \\ &= \frac{\partial}{\partial s} \sum_{i=0}^n F_i \left\{ \ln s^i - \ln (1+s+\dots+s^n) \right\} \\ &= \sum F_i \left(\frac{i}{s} - \frac{1+2s+\dots+ns^{n-1}}{1+s+s^2+\dots+s^n} \right) \end{aligned}$$

$$= -\frac{\sum F_i i}{s} - (\sum F_i) \frac{1 + 2s + \dots + ns^{n-1}}{1 + s + \dots + s^n} = 0 \quad (2.10)$$

とおくと、結局

$$\frac{s + 2s^2 + \dots + ns^n}{1 + s + \dots + s^n} = \frac{\sum F_i i}{\sum F_i} = \bar{i} \quad (2.11)$$

を得る。(2.4')と同様に右辺は時間(i)の漁獲尾数 F_i で重み付けした平均となっている。つまりこの方法はTANAKAの方法と本質的に同等であることがわかる。

ところで

$$\begin{aligned} T_n &= 1 + s + \dots + s^n \\ \left. \begin{aligned} sT_n &= s + \dots + s^n + s^{n+1} \\ (1-s)T_n &= 1 \qquad \qquad \qquad -s^{n+1} \end{aligned} \right\} \end{aligned}$$

だから

$$T_n = \frac{1 - s^{n+1}}{1 - s} \quad (2.12)$$

一方、

$$\begin{aligned} U_n &= s + 2s^2 + \dots + ns^n \\ \left. \begin{aligned} sU_n &= s^2 + \dots + (n-1)s^n + ns^{n+1} \\ (1-s)U_n &= s + s^2 + \dots + s^n - ns^{n+1} \end{aligned} \right\} \\ &= \frac{s(1-s^n)}{1-s} - ns^{n+1} \end{aligned}$$

だから

$$U_n = \frac{s(1-s^n)}{(1-s)^2} - \frac{ns^{n+1}}{1-s} \quad (2.13)$$

である。 $n \rightarrow \infty$ のとき、 $0 < s < 1$ だから

$$\bar{i} = \frac{U_n}{T_n} \rightarrow \frac{s}{1-s} \quad (2.14)$$

となる。TANAKAの方法では $n \rightarrow \infty$ のとき

$$\bar{i} = \frac{1}{Z} = -\frac{1}{\ln s} \quad (2.4')$$

であった。前者は区間データ、後者は連続データであるから式が異なる。しかし、 $s \doteq 1$ の場合には両式は近似する。これには以下のテイラー展開を用いる。

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 \dots \quad (A)$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \dots \quad (B)$$

なお(B)は(A)を積分すれば得られる。

$$\frac{s}{1-s} = \frac{1}{\frac{1}{s}-1} = -\frac{1}{1-\frac{1}{s}}$$

ここで

$$\begin{aligned} 1 - \frac{1}{s} &= 1 - \frac{1}{1-(1-s)} \\ &= -(1-s) - (1-s)^2 - (1-s)^3 - \dots \end{aligned}$$

一方,

$$\begin{aligned} \ln s &= \ln \{1 - (1-s)\} \\ &= -(1-s) - \frac{(1-s)^2}{2} - \frac{(1-s)^3}{3} - \dots \end{aligned}$$

となるからである。

(3) 応用例

図10に示す。単純な例で明らかに $s = 0.1$ である。
まず PAULIK の方法で計算してみる(図10の上)。(2. 11) の右側より

$$\begin{aligned} \frac{-}{i} &= \frac{0 \times 1000 + 1 \times 100 + 2 \times 10 + 3 \times 1}{1000 + 100 + 10 + 1} = \frac{123}{1111} \\ &= 0.111 \end{aligned}$$

$n \rightarrow \infty$ (打ち切りなし) の場合なので (2. 11) より

$$\frac{s}{1-s} = 0.111$$

これを解くと $s = 0.1$ が求まる。

次に TANAKA の方法で計算してみる。ヒストグラムの階級値(各区間の中央の値)を使用してみると (2. 4') より

$$\begin{aligned} \frac{-}{t} &= \frac{0.5 \times 1000 + 1.5 \times 100 + 2.5 \times 10 + 3.5 \times 1}{1000 + 100 + 10 + 1} \\ &= \frac{678.5}{1111} = 0.611 \end{aligned}$$

従って

$$-\frac{1}{\ln s} = 0.611$$

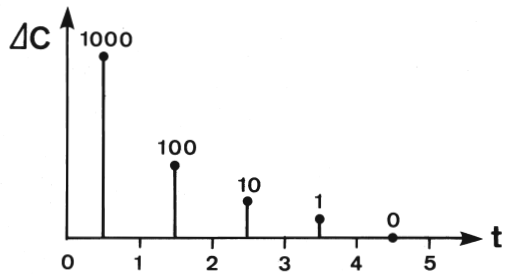
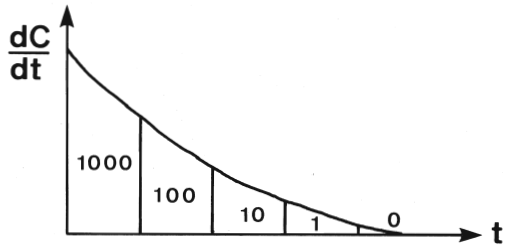
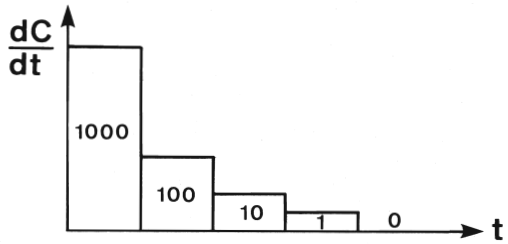


図10 応用例

これを解くと $s = 0.1945$ となり全く間違った値である。

これは \bar{t} の計算を間違えたため、原因はヒストグラムの階級値を使用した点である。TANAKA の方法は連続モデルだから図10の中のように考えなくてはならない。区間 $[0, 1]$ における平均時間は (2. 5) より

$$\bar{t}_1 = \frac{1}{Z} + \frac{1}{1 - e^{-Z}} \quad (2. 15)$$

となる。($s = 0.1$ のとき $\bar{t}_1 = 0.3232$ であるが、前の失敗は $\bar{t}_1 = 0.5$ としたためである) 後の区間 $[i - 1, i]$ での平均時間は

$$\bar{t}_i = \bar{t}_1 + (i - 1) \quad (2. 16)$$

となるから、結局全体の平均時間は

$$\begin{aligned} \bar{t} &= \frac{1000\bar{t}_1 + 100\bar{t}_2 + 10\bar{t}_3 + \bar{t}_4}{1000 + 100 + 10 + 1} \\ &= \frac{123}{1111} + \frac{1}{Z} + \frac{1}{1 - e^{-Z}} \end{aligned}$$

となる。これが $1/Z$ に等しい訳だから

$$\frac{1}{e^{-Z} - 1} = \frac{s}{1 - s} = 0.111$$

で $s = 0.1$ が求まるが、これは PAULIK の方法そのものである。

また、次のようにも考えられる。各区間のある時点で瞬間的にすべて漁獲されると考える (図10の下)。この時点をとりにあえず $t = 0.5, 1.5, 2.5, 3.5, 4.5$ とおく。このとき全漁獲尾数は

$$C_u = \Delta C_{0.5} + \Delta C_{1.5} + \Delta C_{2.5} + \Delta C_{3.5}$$

であるから、各区間の確率は

$$P_{0.5} = \frac{\Delta C_{0.5}}{C_u}, \quad P_{1.5} = \dots$$

となる。尤度は

$$L = P_{0.5}^{1000} \cdot P_{1.5}^{100} \cdot P_{2.5}^{10} \cdot P_{3.5}$$

だから

$$\frac{\partial Y}{\partial s} = \frac{\partial (\ln L)}{\partial s} = 0$$

を満たす s を求めればよい。

ところで瞬間 dt における漁獲は dC で (1. 24) より

$$\frac{dC}{dt} = FN_0 e^{-Zt} = FN_0 s^t$$

であったから、 P の理論値はそれぞれ

$$P_{0.5} = \frac{s^{0.5} dt}{s^{0.5} dt + s^{1.5} dt + s^{2.5} dt + s^{3.5} dt}$$

$$= \frac{1}{1 + s + s^2 + s^3}$$

$$P_{1.5} = \frac{s}{1 + s + s^2 + s^3}$$

$$P_{2.5} = \dots\dots$$

となるが、これは(2. 8)であり、つまり PAULIK の方法をそのままなぞっている訳である。 s は時間の起点のとり方に影響されないから、 $t=0.5, 1.5, 2.5, \dots\dots$ として大丈夫である。

(4) 解の一意性の証明

s を求めるためには(2. 11)を解かなくてはならないが、既に数表が与えられている(北田1985)。パラメータ数が1だから簡単に解け、二分法等で十分であろう。

問題は解の一意性である。(2. 11)は s の n 次方程式だから、複素数の範囲で n 個の根をもつ。 $0 < s < 1$ でただ1つの根しか存在しないことを証明する必要がある。

$$f(s) = \frac{s + 2s + \dots\dots + ns^n}{1 + s + \dots\dots + s^n} \tag{2. 17}$$

とおく。 $f(0) = 0$,

$$f(1) = \frac{n(n+1)/2}{n+1} = \frac{n}{2} > \bar{i} \tag{2. 18}$$

のとき、 $f(s)$ が単調増加であることが示せばよい(図11)。

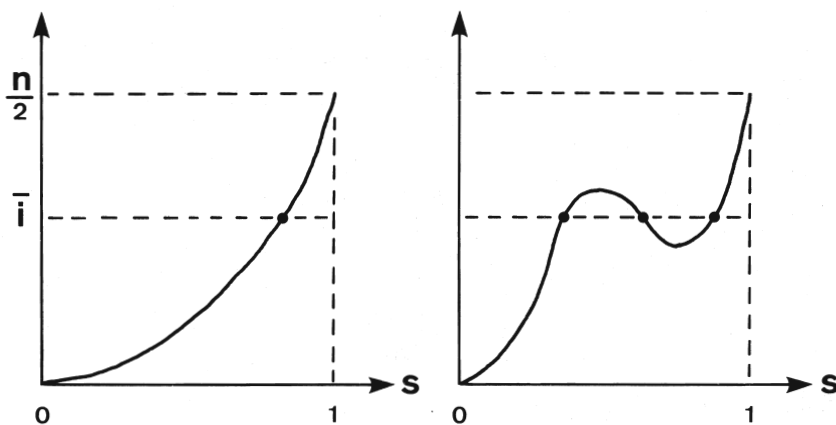


図11 解の一意性の説明図

これは実に簡単に示せる。

$$f(s) = \sum_{i=1}^n \frac{s^i + \dots + s^n}{1 + s + \dots + s^n} \quad (2.19)$$

と分解できるが、このとき

$$\begin{aligned} \frac{s^i + \dots + s^n}{1 + s + \dots + s^n} &= \frac{1}{\frac{1 + s + \dots + s^{i-1}}{s^i + \dots + s^n} + 1} \\ &= \frac{1}{\frac{\frac{1}{s^i} + \dots + \frac{1}{s}}{1 + s + \dots + s^{n-i}} + 1} \end{aligned} \quad (2.20)$$

となるからすべて単調増加で、 $f(s)$ も単調増加となる。

(5) KULLDORFF (1961) の方法

PAULIK の方法は等区間のモデルであったが、KULLDORFF の方法は不等区間のモデルへ拡張されている。

図12において区間 $[a, b]$ における漁獲尾数は (1.24) より

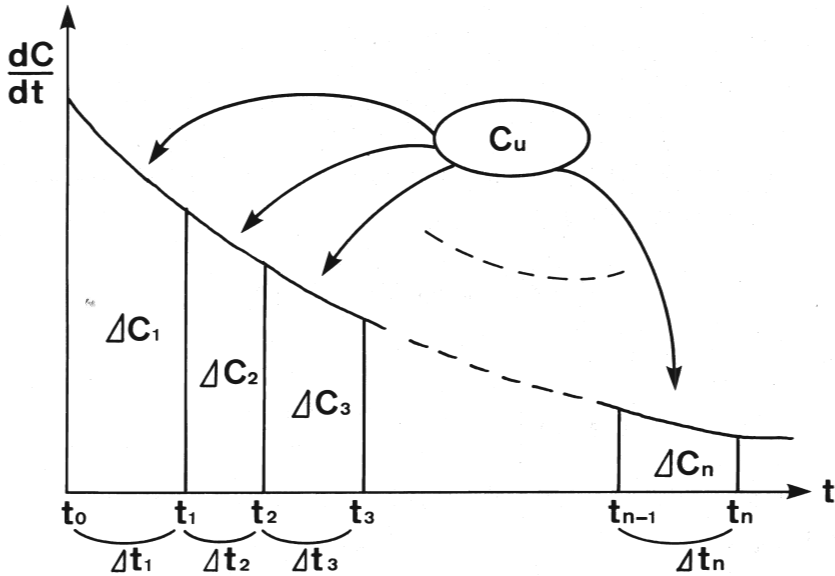


図12 KULLDORFF のモデル

$$\Delta C = \frac{F}{Z} N_0 (s^a - s^b)$$

である。 $t_0 = 0$, $t_n = u$ において一般性を失わない。従って

$$C_u = \frac{F}{Z} N_0 (1 - s^u) \quad (2.1)$$

$$\Delta C_i = \frac{F}{Z} N_0 (s^{t_{i-1}} - s^{t_i})$$

$$= \frac{F}{Z} N_0 s^{t_i-1} (1 - s^{\Delta t_i}) \quad (2.21)$$

$$\Delta t_i = t_i - t_{i-1} \quad (2.22)$$

となる。故にこの区間の確率は

$$P_i = \frac{\Delta C_i}{C_u} = \frac{s^{t_i-1} - s^{t_i}}{1 - s^u} \quad (2.23)$$

となる。

各区間の漁獲尾数が F_1, F_2, \dots, F_n (ヒストグラム) で与えられたとき、尤度は

$$L = \sum_{i=1}^n P_i^{F_i} \quad (2.24)$$

となる。(2.9) とは $i=0$ が含まれていない点で異なる。 s は時間の起点のとり方に影響されないから、 s の推定に差は生じない。このとき

$$\frac{\partial Y}{\partial s} = \frac{\partial (\ln L)}{\partial s} = \sum_{i=1}^n \frac{F_i}{P_i} \frac{\partial P_i}{\partial s} = 0 \quad (2.25)$$

を求めればよい。($\partial P_i / \partial Z$ を計算した方が楽。) 結局

$$\sum_{i=1}^n F_i t_{i-1} - \sum_{i=1}^n F_i \frac{\Delta t_i s^{\Delta t_i}}{1 - s^{\Delta t_i}} + \left(\sum_{i=1}^n F_i \right) \frac{us^u}{1 - s^u} = 0 \quad (2.26)$$

を得る。これを s について解けばよい。 Δt_i はすべてデータとして与える訳だから大変です。

等区間にすると計算は大部楽になる。

$$\Delta t_i \equiv 1, \quad u \rightarrow n+1, \quad \sum_{i=1}^n \rightarrow \sum_{i=0}^n \quad (2.27)$$

と変更してみると、右側の2項は

$$\left(\sum_{i=0}^n F_i \right) \left\{ -\frac{s}{1-s} + \frac{(n+1)s^{n+1}}{1-s^{n+1}} \right\}$$

となるが、ここで (2.12), (2.13) を使えば完全に (2.10) と一致する。

(6) 解の一意性の証明

(2.26) の解を与える BASIC プログラムは北田 (1987b) にある。一意性の証明を以下に示す。

$$\sum_{i=1}^n F_i t_{i-1} = \sum_{i=1}^n F_i \left(\frac{\Delta t_i s^{\Delta t_i}}{1 - s^{\Delta t_i}} - \frac{us^u}{1 - s^u} \right) \quad (2.26')$$

とする。ここで

$$G(x) = \frac{xs^x}{1-s^x} = \frac{xe^{-Zx}}{1-e^{-Zx}} = \frac{x}{e^{Zx}-1}$$

$$\sum_{i=1}^n \Delta t_i = u \text{ であるから } 0 < a = \Delta t_i < u$$

$$f_i(Z) = G(a) - G(u)$$

とおく。このとき $f_i(Z)$ が Z について単調減少であることを示す。

$$g(x) = \frac{\partial G(x)}{\partial Z} = \frac{-x^2 e^{Zx}}{(e^{Zx} - 1)^2} = -\frac{x^2}{e^{Zx} + 2 + e^{-Zx}}$$

$$= -\frac{1}{Z^2 + \frac{Z^4}{12}x^2 + \dots} < 0$$

ここでテイラー展開

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots \quad (C)$$

を用いた。従って $g(x)$ は x について単調増加である (図13の中)。故に

$$\frac{\partial f_i(Z)}{\partial Z} = g(a) - g(u) < 0$$

となるので $f_i(Z)$ は Z について単調減少である。

一方,

$$G(x) = \frac{1}{Z + \frac{Z^2}{2}x + \frac{Z^3}{6}x^2 + \dots} > 0$$

となるから $G(x)$ は x について単調減少である (図13の上)

$$f_i(Z) = G(a) - G(u) > 0$$

であるが、このとき

$$\begin{aligned} f_i(Z) &= \frac{1}{Z + \frac{Z^2}{2}a + \frac{Z^3}{6}a^2 + \dots} - \frac{1}{Z + \frac{Z^2}{2}u + \frac{Z^3}{6}u^2 + \dots} \\ &= \frac{\frac{u-a}{2} + \frac{Z}{6}(u^2 - a^2) + \dots}{(1 + \frac{Z}{2}a + \frac{Z^2}{6}a^2 + \dots)(1 + \frac{Z}{2}u + \frac{Z^2}{6}u^2 + \dots)} \\ &\xrightarrow{Z \rightarrow 0} \frac{u-a}{2} > 0 \end{aligned}$$

である。同様にして

$$\frac{\partial f_i(Z)}{\partial Z} = g(a) - g(u) \xrightarrow{Z \rightarrow 0} -\frac{u^2 - a^2}{12}$$

を得る。結局 $f_i(Z)$ は図13の下のようなになる。

(2. 26') において、左辺は定数であり、右辺は $Z : 0 \rightarrow \infty$ ($s : 1 \rightarrow 0$) のとき単調減少となるからただ1つの解を持つが、その条件は

$$\sum_{i=1}^n F_i t_{i-1} < \sum_{i=1}^n F_i \frac{u - \Delta t_i}{2}$$

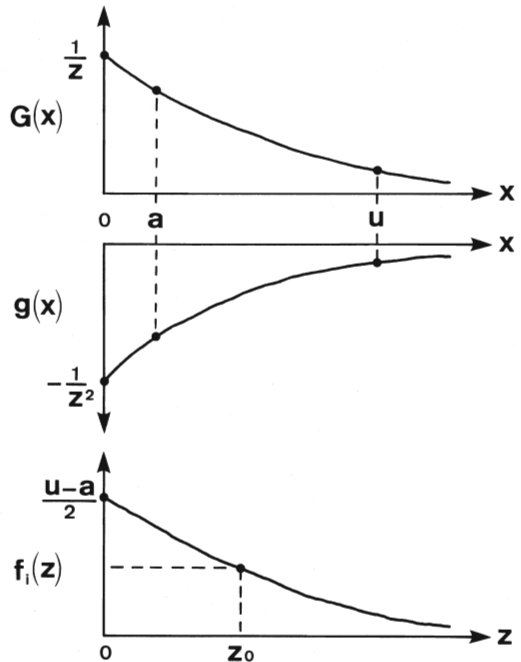


図13 解の一意性の説明図

となる。これを整理すると

$$\sum_{i=1}^n F_i (t_{i-1} - t_i) < u \sum_{i=1}^n F_i \quad (2.27)$$

を得る。

(7) PAULIK と KULLDORFF の方法のまとめ

(5)で示したように KULLDORFF の方法は PAULIK の方法をその中に含んでいる。従って(6)の証明は(4)の証明をその中に含んでいる。等区間のデータに両手法を用いれば、その解は完全に一致するはずである。KULLDORFF の方法の方が不等区間のデータにも使用可能で一般的であるが、応用上は等区間のデータの方が素直で、等区間の手法のみで十分であろう。

結局、同一モデルに最尤法で推定したのであるから、結果はすべて一致する訳である。PAULIK は離散モデルから、KULLDORFF は連続モデルから出発したのであるが、指数関数では高さと同面積が共に同じ性質（等比数列）を有していたため、同一モデルになってしまった訳である。

なお、周期関数で拡張した漁獲方程式モデルについては等区間（区間巾は周期と一致）の手法しか使えない点に注意しておく。

(8) KITAHARA *et al* (1986) の方法

このモデルは休漁期を含んでいることが特徴である。 $F \rightarrow Fg(t)$ と考えると

$$g(t) = \begin{cases} 1 & \text{： 漁期} \\ 0 & \text{： 休漁期} \end{cases} \quad (2.28)$$

となる。したがって

$$Z = \begin{cases} F + M & \text{： 漁期} \\ M & \text{： 休漁期} \end{cases} \quad (2.29)$$

である。

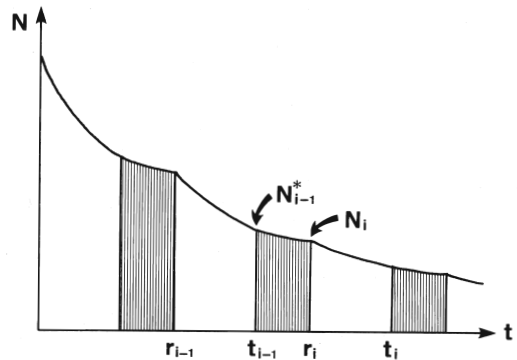


図14 KITAHARA *et al* の説明図（縦線：休漁期）

図14のようにすると

$$\begin{cases} N^*_{i-1} = N_{i-1} e^{-(F+M)(t_{i-1} - r_{i-1})} \\ N_i = N^*_{i-1} e^{-M(r_i - t_{i-1})} \end{cases} \quad (2.30)$$

という関係が成立している。漁獲（C）は漁期にしか得られないから、今までのモデルをつなぎ合わせることになる。rやtのデータを入力せねばならず、長くて複雑な式となるが、原理的には今まで述べた手法で十分に対応可能であろう。むしろ各漁期・休漁期ごとに個別に解析して総合的に判定した方がよいように思われる。

3. 実際上の問題点

以上みてきたように、ここでとりあげた手法は基本的にすべて同等で、きちんとしたデータに正しく用いるなら、結果にほとんど差は生じない。

実際の種苗放流実験では初期減耗が非常に大きく（逸散が原因と思われる）、直接にこれらの手法

を用いると、実際とかけはなれた値になりやすい(図15)。

やはり基本的には、まず対数回帰モデルへのあてはめを行うべきであろう。これは散布図やヒストグラムに相当するもので、記述統計の最も基本的な手法である。安定期であればきれいな直線が得られ、傾きより Z が直ちに求まるはずである。最もそのようなきれいなデータであれば誰も苦労しないのであるが……。

本論を終えるにあたり、貴重な資料や助言をいただいた日本栽培漁業協会北田修一氏に深謝します。

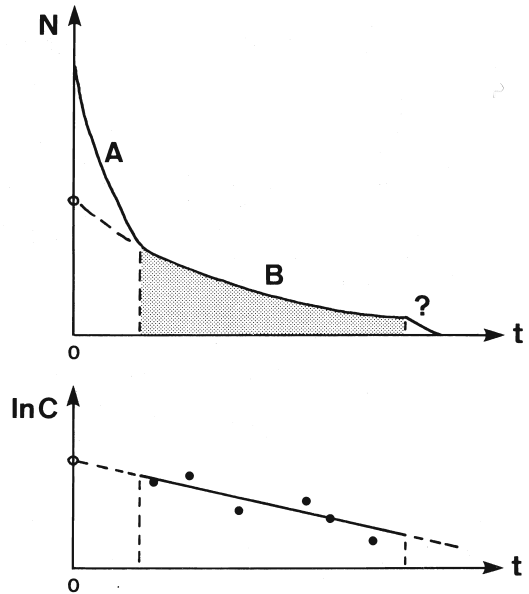


図15 種苗放流モデル(上)と対数回帰モデル(下) A:初期減耗期 B:安定期

引用文献

- AKAMINE, T. (1986) Expansion of growth curves using a periodic function and BASIC programs by MARQUARDT'S method. 日水研報 (36), 77-107.
- BARANOV, F. I. (1918) 漁業における生物学的基礎の問題 (笠原昊・深滝弘訳1951). 水産庁漁業科学叢書 1, 48 pp.
- BEVERTON, R. J. H. and S. J. HOLT (1957) On the dynamics of exploited fish populations. *Fish. Invest.*, 19, 1-533.
- CHAPMAN, D. G. and D. S. ROBSON (1960) The analysis of catch curves. *Biometrics*, 16, 354-368.
- GULLAND, J. A. (1955) On the estimation of population parameters from marked members. *Biometrika*, 42, 269-270.
- 石岡清英 (1987) 質疑応答. pp 137, In 資源評価のための数値解析, 嶋津靖彦編, 恒星社厚生閣, 東京, 162 pp.
- 加藤史彦 (1987) 個体識別をする多回放流・再捕調査における留意点, 栽培技研, 16(1), 47-61.
- 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——. 日裁協研究資料, 28, 20 pp.
- 北田修一 (1987 a) 標識再捕による放流効果の評価. pp 102-117, In 資源評価のための数値解析, 嶋津靖彦編, 恒星社厚生閣, 東京, 162 pp.
- 北田修一 (1987 b) 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) 期間ごとの標識再捕データに基づく死亡係数の推定について. 日水誌, 53(10), 1765-1770.
- KITAHARA, T., K. TAKAHASHI, Y. HUIKAWA, K. ISHIDA and J. YOSHIO (1986) Estimation of fishing mortality coefficient from tag recoveries in a finite period with a few closed seasons. 日水誌, 52(4), 579-584.
- KULLDORFF, G. (1961) Estimation from grouped and partially grouped samples. pp 55-63, John Wiley & Sons, Inc., New York.
- 久野英二 (1986) 動物の個体群動態研究法 I——個体数推定法——. 共立出版, 東京, 114 pp.
- PAULIK, G. J. (1963) Estimation of mortality rates from tag recoveries. *Biometrics*, 19, 28-57.
- ROBSON, D. S. and D. G. CHAPMAN (1961) Catch curve and mortality rate. *Trans. Ame. Fish. Soc.*, 90, 181-189.
- SEBER, G. A. F. (1982) The estimation of animal abundance and related parameters, 2nd edition. Griffin, London, xvii + 654 pp.
- TANAKA, S. (1967) Estimation of fishing coefficient of Mojako by tagging experiments on drifting Seaweeds—I, Method and an example. 日水誌, 33, 1108-1115.
- 田中昌一 (1985) 水産資源学総論. pp 180-193, 恒星社厚生閣, 東京, 381 pp.
- 鷲尾泰俊 (1978) 推定と検定. pp 31-36, 共立出版, 東京, 123 pp.

[質疑応答]

- 藤井徹 (日水研) KULLDORFF の方法よりも PAULIK の方法を勧めているが, 実際に得られるデータがバラバラな時は, KULLDORFF の方法の方が有効なのではないか。
- 赤嶺 オリジナルなデータが不等区間で与えられている場合には KULLDORFF の方法がよい。しかし, 通常のデータを不等区間で区切る方法は何通りもあるため, 同一のデータでも解析者によって解が異なる危険性がある。等区間の手法に統一した方が解析者による差が生じない。
- 大塚 (新潟水試) 追跡調査の回数, 範囲は当初より多めに計画した方がよいか。
- 赤嶺 原則として追跡調査の回数, 範囲は多ければ多いほど良い。労力とのかねあいでも少なくせざるを得ないが……。
- 戸田 (山形水試) $\ln N_t = \ln N_0 - Zt$ の Z を推定する時, F/Z が一定として, C_t で代用してもかまわないか, こうして求めた Z を F と M に分離する簡単な方法は。
- 赤嶺 (1. 17), (1. 20) 式が常に成立していれば $dC = FNdt$ となるから, ΔC (区間 Δt における漁獲尾数) で N (その時の資源尾数) を代用できる。通常は (1. 24) 式で F を求める。しかし, F/Z が一定でない場合 [たとえば (1. 26) 式] には計算がかなり面倒になる。大雑把な推定でよければ F/Z を一定と仮定してよいと思う。