

寄稿

魚類の成長式における検定とモデル選択

赤嶺達郎

Statistical Test and Model Selection of Fish Growth Formula

Tatsuro AKAMINE

Many formulae have been presented for fish growth. Although Schnute (1981) formula is widely used in fish population dynamics, it is equivalent to Richards (1959) one. Akamine-Richards formula presented by Akamine (1993) is the simplest and the most practical one which integrates von Bertalanffy, its cubic, Gompertz, logistic, and their seasonal oscillating models expanded by using a periodic function. In this paper, we discuss statistical tests and interval estimations of parameters for fish growth formulae. In particular, the weighted least-squares method for curve fitting, chi-square test for the data with variance, F-test for the data without variance, and model selection using AIC (Akaike Information Criterion) are explained mathematically, and a numerical example of interval estimations for von Bertalanffy formula is presented by using a spreadsheet program.

Key words: growth, interval estimation, model selection, statistical test

はじめに

魚類の成長式について多くの数式が提示されてきた。最近では主要な成長式を内包する統一的な成長式や周期関数によって拡張された成長式が提示され、それらについては既に Akamine (1993) および赤嶺 (1995) 等において解説している。一方、今日では小型計算機用の表計算ソフトによって容易に数理モデルのパラメータが推定できるようになってきており、成長式や網目選択性曲線および混合正規分布についての手法が提示されている (東海, 1997; 相澤・滝口, 1999; 五利江, 2001, 2002)。とりわけ Haddon (2001) ではほとんどすべての資源解析手法が表計算ソフトで処理されている。しかし反面、数式の表現や推定および検定方法について、いくつかの文献で数理的な誤りが散見されるようになった。とりわけ成長式に関するものにおいてこの傾向が顕著である。これは成長式の数理モデルや検定方法等についての議論が不十分であることに起因している。したがってここでは Haddon (2001) を中心としたいくつかの文献の誤りを正すとともに、標準的な成長データの解析手法を提示する。

Richards 式

魚類の成長式として現在もっとも広く用いられているのは von Bertalanffy 式 (以下 VB 式と略す、他も同様) であるが、Gompertz 式 (G 式) や logistic 式 (L 式) もしばしば用いられている。Richards 式 (R 式) はこれらを統一的に記述する成長式であるが、さまざまに表現できるため一部の論文に誤りが認められる。この節ではそれらを整理して解説する。

Richards (1959) は微分方程式:

$$\frac{dw}{dt} = \eta w^m - \kappa w \quad (1)$$

を解いて、

$$w = \left\{ \frac{\eta}{\kappa} - \left(\frac{\eta}{\kappa} - w_0^{1-m} \right) e^{-(1-m)\kappa t} \right\}^{1/(1-m)} \quad (2)$$

という成長式を提示した (w は体サイズ, t は時間, η は同化量の係数, κ は異化量の係数, m は成長曲線の形を決定するパラメータ, w_0 は $t=0$ のときの w の値)。しかし (2) 式は η と κ を含んだままなので扱いにくい。また $m \rightarrow 1$ のとき G 式に近づくが, $m=1$ では (1) 式は指数関数となってしまう。実際, Taylor (1962) は同様の微分方程式を解いて一般化 VB 式 (gVB 式), 指数関数, 一般化 L 式の 3 種類に分類したが, G 式には言及していない。指数関数は成長式として有効ではないので, このまま用いることは問題がある。

2003年6月11日受付, 2003年11月7日受理

独立行政法人水産総合研究センター中央水産研究所
National Research Institute of Fisheries Science, Fisheries Research
Agency, 2-12-4 Fukuura, Kanazawa-ku, Yokohama, Kanagawa
236-8648, Japan
akabe@affrc.go.jp

能勢ほか(1988)では $\eta/\kappa = w_z^{1-m}$ とおいているが (w_z は $t \rightarrow \infty$ のときの w の値), 微分方程式の解法および $m \rightarrow 1$ のとき G 式に近づくことの証明は簡単ではない。

Akamine (1988) は R 式の標準型として, 微分方程式:

$$\frac{dw}{dt} = kw \frac{1 - (w/w_\infty)^r}{r} \quad (3)$$

および, 成長式:

$$w = \frac{w_\infty}{(1 + r e^{-k(t-t_0)})^{1/r}} \quad (4)$$

を提示した (k は成長係数, t_0 は変曲点または $w=0$ となる t の値, $r=m-1$), これらの数式では $r=0$ のとき G 式:

$$\frac{dw}{dt} = kw(\ln w_\infty - \ln w) \quad (5)$$

$$w = w_z \exp(-e^{-k(t-t_0)}) \quad (6)$$

となること, 指数関数と対数関数の定義式:

$$e^x = \lim_{r \rightarrow 0} (1 + rx)^{1/r} \quad (7)$$

$$\ln y = \lim_{r \rightarrow 0} \frac{y^r - 1}{r} \quad (8)$$

から簡単に導かれる (赤嶺, 1994). また (3) 式から (4) 式を簡単に導くには, Akamine (1988) の変数変換を用いればよい. なお計算機では (4) 式に $r=0$ を代入するとエラーが発生するので, その際には (6) 式を用いる. $r=0$ の近傍では桁落ちする危険性があるため, Akamine (1988) はそれを防止するための級数展開を提示している.

微分方程式は「ベクトル場」を意味している (森本, 1987), (3) 式は w の全域で定義されている. しかし (4) 式の曲線は $0 \leq w \leq w_z$ の範囲に限定されたものである. 体サイズの成長式としてはこれで十分である. 一方, Akamine (1988) ではもうひとつの表現型として,

$$w = \frac{w_\infty}{[1 + \{(w_\infty/w_0)^r - 1\} e^{-kt}]^{1/r}} \quad (9)$$

を提示している. この式は個体群サイズの成長式として生態学で用いられているもので, $w < 0$ および $w > w_z$ の部分の曲線も含んでおり, $w_0 > w_z$ の場合には $w > w_z$ の部分における単調減少関数となる. したがって体サイズの場合には (4) 式を, 個体群サイズの場合には (9) 式を用いるのがよい.

ところで gVB 式は VB 式を β 乗したもので,

$$w = w_z (1 - e^{-k(t-t_0)})^\beta \quad (10)$$

と表される. 体長には $\beta=1$ が, 体重には $\beta=3$ が用いられることが多い. $t=t_0$ のとき $w=0$ となるので理解しやすい. しかしこの式は $\beta < 0$ の場合には $w < 0$ および $w > w_z$ の部分

の曲線を意味し, 体サイズの成長式としては不適當となる. したがって gVB 式は $\beta > 0$ の場合に限定して使用すべきで, このとき gVB 式は R 式 (4) における $r < 0$ の部分と一致する.

一方, Schnute (1981) は連立微分方程式:

$$\frac{dw}{dt} = yw, \quad \frac{dy}{dt} = -(k + by)y \quad (11)$$

の解として, 次の成長式 (S 式) を提示した.

$$w = \left\{ w_1^b + (w_2^b - w_1^b) \frac{1 - e^{-k(t-\tau_1)}}{1 - e^{-k(\tau_2 - \tau_1)}} \right\}^{1/b} \quad (12)$$

Schnute は $b < 0$ の場合だけを R 式と呼んでいるが, これは不適切で, S 式は R 式と同一である. (1), (4), (10) 式とは $b=1-m=-r=1/\beta$ という関係にある. (11) 式は R 式の微分方程式 (3) と完全に一致するが, Schnute は理解していないように思われる. Akamine (1993) では (11) 式から (3) 式を導いてこれを証明したが, ここでは逆に (3) 式から (11) 式を導いてみる (この方が簡明なので).

$$y = k \frac{1 - (w/w_\infty)^r}{r} \quad (13)$$

とおくと, (11) の左式が得られる. (13) 式の両辺を t で微分すると,

$$\frac{dy}{dt} = -k \frac{w^{r-1}}{w_\infty^r} \frac{dw}{dt} = -k \left(\frac{w}{w_\infty} \right)^r y = -(k - ry)y$$

となるので, (11) の右式が得られる. S 式は曲線上の 2 点 (τ_1, w_1) , (τ_2, w_2) を用いて R 式を書きかえたもので, $w > w_z$ の部分の曲線も含む. 2 点を $(t_0, 0)$, (∞, w_z) とすれば gVB 式と一致する. また 2 点を $(0, w_z)$, (∞, w_z) とすれば (9) 式と一致する. S 式は同一曲線でも 2 点 τ_1 と τ_2 の設定位置によって, パラメータ w_1 と w_2 の値が異なってしまうため, 成長式の標準型としては扱いにくい. S 式では $b=0$ のときの G 式の表現が分かりづらいが, w_1 をくくりだし, 残りの部分の対数をとって, ロピタルの定理を用いるか, 対数関数のテイラー展開:

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

を用いた上で, 対数関数の定義式 (8) を適用すれば,

$$w = w_1 \left(\frac{w_2}{w_1} \right)^{\psi(t)}, \quad \psi(t) = \frac{1 - e^{-k(t-\tau_1)}}{1 - e^{-k(\tau_2 - \tau_1)}} \quad (14)$$

を得る. したがって (9) 式の場合には,

$$w = w_{\infty} \left(\frac{w_0}{w_{\infty}} \right)^{\psi(t)}, \quad \psi(t) = e^{-kt} \quad (15)$$

と表される。

Schnute and Richards (1990) はさらにパラメータ v を追加して「 $t \rightarrow t^v$ 」と拡張することを提唱しているが (SR式)、これによって成長曲線の当てはまりが格段に良くなったという事例は少ないように思う。むしろ後述のように「 $t \rightarrow F(t)$ 」と拡張して季節変動を表す方が有効だろう。なお実際のSR式は論文中で $\eta = \beta\xi$ と表されており、これはS式の k を w_x に変換した成長式：

$$w^{-b} = w_{\infty}^{-b} + \frac{(w_1^{-b} - w_{\infty}^{-b})(t_2 - t)(t_2 - t_1)}{(w_2^{-b} - w_{\infty}^{-b})(t_1 - t)(t_2 - t_1)} \quad (16)$$

において $w_x = 1$ と固定したものである。またSR式ではS式と b の符号を逆に、つまり $b = r$ としている。

以上のようにS式は数学的にR式と同一のもので、しかも成長式の標準型としてはR式 (4) よりも劣る。以前は非線型モデルのパラメータ推定が困難だったため、できるだけ線型に近づくように数式を變形することがよく行われていたが、現在ではそのような必要性はほとんどない。しかしHilborn and Walters (1992), Quinn and Deriso (1999), およびHaddon (2001) では成長式として (4) 式ではなくS式およびSR式を提示している。単に表計算ソフトを用いて成長データに成長曲線を当てはめるだけならS式でも十分であるが、さらに踏み込んで成長式の比較検討およびパラメータの区間推定等を行う場合には、数式が複雑すぎて理解しにくい、パラメータ数が多すぎて自由度を誤りやすい等の問題が生じてくる。R式の標準型としてはもっとも簡潔な (4) 式を用いるべきである。

季節変動モデル

Haddon (2001) は成長率の季節変動を表す成長式として、Pitcher and Macdonald (1973) が最初に提示したPM式：

$$w = w_{\infty}(1 - e^{-\phi(t)}), \quad \phi(t) = C \sin \frac{2\pi(t-s)}{52} + k(t-t_0) \quad (17)$$

を紹介している。しかしこの式はパラメータの設定が不適切である。具体的に述べると、振幅 C の評価が不十分なためマイナス成長が現れるときの C の値が自明でない、 $t = t_0$ のとき $w = 0$ とならない、時間 t の単位が人為的な「週」となっていて一般的でない等の欠点がある。このままでは表計算ソフトを用いて曲線を当てはめることは可能でも、その後の比較検討が難しくなる。Haddon (2001) ではさらに別な周期関数を加えた拡張式を提示しているが、単に当てはめを良くするだけの拡張にすぎない。数理モデルにおいてパラメータを増やす場合にはもっと慎重な議論が必要である。この方面の拡張や比較の方法についてはKiso *et al.*

(1992) を参照されたい。

Quinn and Deriso (1999) では1982年に提示されたHoernig and Hanumara式 (HH式)：

$$w = w_{\infty}(1 - e^{-k[t-t_0 - 10\phi(t) - \phi(t_0)]}), \quad \phi(t) = \frac{\delta\Omega}{2\pi} \sin \frac{2\pi}{\Omega}(t-t_1) \quad (18)$$

を紹介している。この式では上記3つの欠点が修正されていて応用上は問題ないが、関数 ϕ の意味づけが不十分である上に、数式表現として簡潔でない。時間 t の単位を基準化する Ω は、データをそのように変更すればよいだけだから不要である。なお多くの論文で周期関数による成長式の拡張がVB式だけに限定されているのは奇妙な話である。Quinn and Deriso (1999) 等ではR式 (4) ではなくてS式を用いているので、拡張すると数式が複雑になりすぎるためかもしれない。

Akamine (1993) は季節変動を示す成長式の標準型として、

$$w = \frac{w_{\infty}}{(1 + r e^{-k(F(t) - F(t_0))})^{1/r}}, \quad F(t) = t + \frac{A}{2\pi} \sin 2\pi(t-t_1) \quad (19)$$

を提示した (この式の基本はR式なので、AR式と略す)。ここで $F(t)$ は時間 t を拡張したもので、具体的には周期関数：

$$f(t) = 1 + A \cos 2\pi(t-t_1) \quad (20)$$

を積分したものである (A は振幅、 t_1 は f が最大となる t の値)。 $f(t)$ が水温に対応する場合には、 $F(t)$ は積算水温に対応する。S式を同様に拡張する場合には、S式において $t \rightarrow F(t)$ 、 $t_1 \rightarrow F(t_1)$ 、 $t_2 \rightarrow F(t_2)$ とすればよい。なおこれより前のAkamine (1986, 1988) で用いた周期関数は $a \leq f(t) \leq 1$ と設定したので、 $a < 0$ のときマイナス成長が現れるが、 a の推定値が変化すると成長係数 k の推定値も変化するため、モデル間の比較には不向きである。

McGarvey and Fowler (2002) ではHH式をgVB式に拡張して用いている。ただし時間の単位を月としているため $\Omega = 12$ と固定し、 $u = \delta\Omega$ とおいている。これは振幅 δ を u と誤りやすいため、不適切な変形である。この式をAR式と比較しているが、gVB式はAR式の $r < 0$ の部分と一致するから、成長曲線の当てはめは完全に一致するはずで、結果に差が出るのは不合理である。体長組成から直接に成長式を推定しているが、誤差分布に通常の正規分布ではなく、端の切れた正規分布を使用している。このような複雑な推定方法を行っているため、パラメータ推定を誤ったものと推察される。季節変動を示す成長式の標準型としては、もっとも簡潔で意味づけの明確なAR式を採用すべきである。

成長曲線の当てはめと検定

曲線の当てはめの基本的手法は最小2乗法であるが、Haddon (2001) の解説には不適切な部分がある。最小2乗法では分散が既知の場合と未知の場合で取扱いが異なる。成長式の解析では年齢査定データや体長組成データ等を用いることが多いため、分散が既知の場合が多い。ここでは最初に分散が既知の場合について解説し、その後で未知の場合について述べる。

成長曲線の当てはめにおける理想的な解析手順は以下のようになる。(1) データとして各測定時刻 t_i における同一コホートの個体数 n_i 、平均 μ_i および分散 σ_i^2 がデータとして得られる ($i=1\sim\xi$)。 (2) これから「平均 μ_i の分散」 $\sigma_i^{2*}=\sigma_i^2/n_i$ を計算する。(3) したがって代表データとして ξ 個の $(t_i, \mu_i, \sigma_i^{2*})$ が得られる。(4) 目的関数は重みづけ最小2乗法:

$$Y = \sum \frac{\{w(t_i) - \mu_i\}^2}{\sigma_i^{2*}} = \sum z_i^2 \quad (21)$$

を用いる。ここで \sum は $i=1\sim\xi$ の和を意味している。これを最小にするパラメータ値を計算機で求めればよい。非線型回帰モデルであるが、残差は正規分布に従うと仮定し、推定や検定には残差平方和を用いるので、通常の線型回帰モデルと同様になる。 Y は $\chi^2(\xi)$ 分布に従うから、この値からデータとモデルの適合度を判定できるし、さらに各パラメータの検定や区間推定が行える。McGarvey and Fowler (2002) は体サイズが小さい場合に、残差が端の切れた正規分布に従うと仮定しているが、そのような場合でも中心極限定理によって平均値は正規分布に従う。したがって平均値とその分散を用いるのであれば、残差は正規分布に従うと仮定してよい。また赤嶺 (1997) に示したように、最小2乗法においては個々のデータを分散 σ_i^2 で重みづけして用いた場合と、平均値とその分散 σ_i^{2*} で代表させて用いた場合とは結果は異ならない。以下に主要な事項について解説する。

1) パラメータの区間推定

パラメータの真値を θ_0 、 Y_{\min} を与えるパラメータ値を θ_1 とする。真値 θ_0 は推定不能であるが、点推定値 θ_1 は表計算ソフトを用いれば簡単に求まる (五利江, 2001)。データ数 ξ 、パラメータ数 η の場合、 Y_{\min} は $\chi^2(\xi - \eta)$ 分布に従う。 η 個のパラメータが固定されるため、自由度が η だけ減少するからである。VB式の場合、パラメータは (k, w_x, t_0) の3つなので、 $\eta=3$ である。帰無仮説:

$$H_0: \theta_j = \theta_{j0} \quad (j=1\sim\eta) \quad (22)$$

のとき Y_0 とすると、 Y_0 は $\chi^2(\xi)$ 分布に従う。 χ^2 分布の加法性から、

$$Y_0 - Y_{\min} \sim \chi^2(\eta) \quad (23)$$

となる。これを用いればパラメータの区間推定ができる。分散が未知の場合には $\sigma^2=$ 一定と仮定して、F検定を行えばよい。残差平方和を

$$S = \sum \{w(t_i) - \mu_i\}^2 \quad (24)$$

とすると、F分布では分子と分母が独立なので、

$$\frac{(S_0 - S_{\min})/\eta}{S_{\min}/(\xi - \eta)} \sim F(\eta, \xi - \eta) \quad (25)$$

を用いればよい。(23)式と(25)式は線型モデルの場合と同じである。線型モデルでは信頼域が楕円形であるが、成長式等の非線型モデルでは信頼域がバナナ状に曲がっているため、信頼域の推定はやや難しくなる。また $\sigma^2=$ 一定という仮定は線型回帰モデルでは妥当な場合が多いが、成長式等の非線型回帰モデルの場合には不適切な場合が多いので、他のデータ等から分散を推定して、できるだけ重みづけした方がよい。

2) 2曲線の比較

成長解析では複数の成長式を比較して検定することがある。典型的な例として雄と雌の成長式を比較する場合について解説する。雄のデータとして $(t_m, \mu_m, \sigma_m^{2*})$ が ξ_m 個、雌のデータとして $(t_f, \mu_f, \sigma_f^{2*})$ が ξ_f 個与えられたとする。帰無仮説は

$$H_0: \theta_{jm} = \theta_{jf} \quad (j=1\sim\eta) \quad (26)$$

なので、全データを込みにして1本の成長曲線を求めた場合を Y_0 、雄のデータと雌のデータを別々にして2本の成長曲線を求めた場合を Y_m および Y_f とする。このとき Y_0 は $\chi^2(\xi_m + \xi_f - \eta)$ に従い、 $(Y_m + Y_f)$ は $\chi^2(\xi_m + \xi_f - 2\eta)$ に従うから、

$$Y_0 - Y_m - Y_f \sim \chi^2(\eta) \quad (27)$$

となるので検定できる。分散が未知の場合には、

$$\frac{(S_0 - S_m - S_f)/\eta}{(S_m + S_f)/(\xi_m + \xi_f - 2\eta)} \sim F(\eta, \xi_m + \xi_f - 2\eta) \quad (28)$$

を用いればよい。赤嶺 (2001a) のp49の(3.4.28)式は誤りなので修正されたい。

通常は雌雄が同時に測定されることが多い ($t_m = t_f$, $\xi_m = \xi_f$)。この場合、各測定時ごとに $H_0: \mu_{jm} = \mu_{jf}$ のように平均値の差の検定を行って、雌雄のデータが同一母集団かどうかを検定する方法も考えられる。しかし測定時ごとに結果が異なる場合には解釈が難しくなる。「多重比較」問題における分散分析とt検定の関係に似ている。成長式の同一性を検定する場合には、(27)式または(28)式を用いる方が無難だろう。

3) F検定と尤度比検定

Haddon (2001) はF検定と尤度比検定を論じているが、両

者の違いを認識していない。ここでは両者の違いについて解説する。最小2乗法では残差が正規分布に従うと仮定している。個々のデータの確率密度 P_i は

$$P_i = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2} z_i^2\right) \quad (29)$$

となる。これより全データの尤度（同時確率密度）はデータ数を ξ として、

$$L = \prod P_i = \frac{1}{\prod \sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2} \sum z_i^2\right) \quad (30)$$

となる。ここで \prod は $i=1\sim\xi$ の積を意味している。これより Y_{\min} のとき L_{\max} となることがわかる。つまり回帰モデルでは最尤法と最小2乗法は一致する。後者の方が計算量が少なく、収束も安定している。後者を用いるべきである。真の解の尤度を L_0 、そのときの残差を z_0 、 L_{\max} の残差を z_1 とする。このとき尤度比 λ を用いると、

$$-2 \ln \lambda = -2 \ln \frac{L_0}{L_{\max}} = \sum z_{0i}^2 - \sum z_{1i}^2 \sim \chi^2(\eta) \quad (31)$$

となる。上式において「 $-2 \ln$ 」という操作は、正規分布の指数部分に含まれている z^2 を引き出すためのものである。

分散が未知の場合にはすべて同一の σ^2 を仮定する。このとき確率密度は

$$P_i = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} z_i^2\right) \quad (32)$$

となるので、

$$-2 \ln \lambda = \sum \ln \frac{\sigma_0^2}{\sigma_1^2} + \sum z_{0i}^2 - \sum z_{1i}^2 \quad (33)$$

となる。最尤法で σ^2 を推定すると、 $\sigma_0^2 = S_0/\xi$ 、 $\sigma_1^2 = S_{\min}/\xi$ となるから、これを(33)式に代入する。 $\sum z^2 = S/\sigma^2$ だから右辺第2~3項は $\xi - \xi = 0$ となるので、

$$-2 \ln \lambda = \xi \ln \left(\frac{S_0}{S_{\min}} \right) = \xi \ln \left(1 + \frac{S_0 - S_{\min}}{S_{\min}} \right) \quad (34)$$

を得る。Haddon (2001) の p231 ではこの式の途中までで終わっていて、これが近似的に χ^2 分布に従うとしているが、ここで正式に (34) 式から χ^2 分布を導いてみる。最右辺に $\ln(1+x)$ のテイラー展開を用いて、 $\xi \rightarrow \infty$ とすれば、 S_0 と S_{\min} は ξ に比例して大きくなるから、

$$-2 \ln \lambda \rightarrow \frac{S_0 - S_{\min}}{S_{\min}/\xi} \sim \frac{\chi^2(\xi) - \chi^2(\xi - \eta)}{\chi^2(\xi - \eta)/\xi} \quad (35)$$

となる。分子は $\chi^2(\eta)$ 分布に従うが、分母の期待値は $\xi \rightarrow \infty$ のとき $(\xi - \eta)/\xi \rightarrow 1$ となるから、 $(-2 \ln \lambda)$ は $\chi^2(\eta)$ 分

布に従う。

しかしながら (34) 式は同時に、F検定：

$$\frac{S_0 - S_{\min}}{S_{\min}} \sim \frac{\eta}{\xi - \eta} F(\eta, \xi - \eta) \quad (36)$$

が適用できることを意味している。(35) 式の尤度比検定 (χ^2 検定) は大標本でないといけないが、(36) 式の F 検定は小標本でも適用できる。したがって分散が未知の場合には大標本の尤度比検定ではなくて、小標本の F 検定を行うべきである。なおこの節の議論は鷲尾 (1978) を参考とした。また除去法における確率モデル (Akamine *et al.*, 1992) においても (33) 式と類似した式が導かれているが、本質的に異なるモデルなので、その後の扱い方が異なっている。

4) AIC について

赤池の情報量規準 (AIC) は水産分野でも広く用いられるようになってきている。赤嶺 (1987, 2001b) で簡単に紹介したが、尤度との関係についてここで補足説明する。厳密な導出や解説は坂元ほか (1983)、坂元 (1985)、石黒 (1993)、および岸野 (1993, 1999) を参照されたい。

$L_0 = L(\theta_0)$ 、 $L_{\max} = L(\theta_1)$ とする。このとき一般に期待値として、

$$\ln L(\theta_1) = \ln L(\theta_0) + \eta/2 \quad (37)$$

という関係が成立する。この式は (31) 式と同一で、(31) 式では正規分布を仮定しているが、それを一般化したものである。一方、真の対数尤度は不明であるが、それに代わるものとして期待平均対数尤度 $MELL(\theta)$ というものが計算できて、期待値として

$$MELL(\theta_0) = MELL(\theta_1) + \eta/2 \quad (38)$$

という関係が成立する。 $MELL(\theta_0)$ と $\ln L(\theta_0)$ の期待値は一致するので、両式から消去すると、

$$MELL(\theta_1) = \ln L(\theta_1) - \eta \quad (39)$$

は期待平均対数尤度の不偏推定量になっている。したがって

$$AIC = -2 MELL(\theta_1) = -2 \ln L_{\max} + 2\eta \quad (40)$$

と定義する。 (-2) を掛けるのは歴史的経緯による。モデル選択では AIC を最小にするモデルを採用すればよい。分散が既知の場合には

$$AIC = Y_{\min} + 2\eta + \text{定数項} \quad (41)$$

となる。分散が未知の場合には、分散に最尤推定値 (S_{\min}/ξ) を代入して、

$$AIC = \xi \ln S_{\min} + 2(\eta + 1) + \text{定数項} \quad (42)$$

Table 1. The artificial data for von Bertalanffy growth formula (Akamine, 1986).

i	t_i	μ_i	σ_i	i	t_i	μ_i	σ_i	i	t_i	μ_i	σ_i
1	0.5	5	3	8	2.0	47	2	15	3.5	80	3
2	0.8	12	3	9	2.2	54	3	16	4.0	82	2
3	1.0	18	2	10	2.4	63	3	17	4.5	87	3
4	1.2	30	4	11	2.5	66	3	18	5.0	88	3
5	1.3	36	3	12	2.8	69	3	19	7.5	99	5
6	1.5	42	3	13	3.0	68	6	20	10.0	99	2
7	1.7	45	3	14	3.2	74	3				

という式を得る。これは帰無仮説：

$$H_0: \theta_j = \theta_{j0} \quad (j=1 \sim \eta), \quad \sigma^2 = \sigma_0^2$$

を立てて、(33) 式の尤度比検定を行った場合に相当する。(42) 式は (41) 式と比較してパラメータ数が1だけ増えているが、これは定数項に含めてもよい。AICの値自体はそれほど重要ではないので、使用する際には定数項は無視して、AICの値の差だけをみるのが賢明である。(39) 式から、最適なモデルではパラメータが1つ増えるごとに最大対数尤度も1だけ上がる必要がある。赤嶺 (1987) および Akamine (1988) では (42) 式を導く際に不偏分散を用いて近似しているが、これは誤りである。

表計算ソフトによる実例

実際の成長データに成長式を当てはめる場合には、散布図からいくつかの成長式を求め、比較検討することになる。成長式のパラメータを推定するには表計算ソフトを用いて五利江 (2001)、Haddon (2001)、東海 (1997) 等を参照すればよい。次にモデル選択の方法であるが、パラメータ数が同一の場合には残差平方和 Y または S が最小のものを、同一でない場合には AIC (41) 式または (42) 式が最小のものを選択すればよい。これについては Kiso *et al.* (1992) を参照されたい。

検定および推定についてはパラメータ推定が非線型になるということ以外は通常の線型モデルとまったく同じである。ただし検定については問題ないが、区間推定については信頼域が楕円形ではなく、細長いバナナ状となるため注意が必要である。したがってここでは区間推定についてだけ実例を示すことにする。なお特定の1変数についての区間推定は、周辺尤度を用いて通常の信頼区間を求めればよいので、ここでは多変数のパラメータ空間における信頼域を求めることにする。成長式のパラメータ推定を行う場合には、特定の1つのパラメータではなく、すべてのパラメータに興味がある場合が多いからである。この2つの信頼区間の違いについては岸野 (1999) の図4.11を参照されたい。

それでは五利江 (2001) に従って、表計算ソフトを用い

てVB式：

$$w = w_\infty (1 - e^{-k(t-t_0)}) \quad (43)$$

のパラメータを推定してみよう。データは Akamine (1986) の Table 1 の人工データをそのまま使用する (Table 1)。ただしこのデータは σ を含んでいるので、 σ^2 で重みづけする必要がある。したがって目的関数は S ではなくて Y である。表計算ソフトの最適化プログラムを適用すると、Akamine (1986) と同一の解： $Y_{\min} = 19.998$, $(k, w_\infty, t_0) = (0.47850, 100.916, 0.49554)$ が得られた。Akamine (1986) では Marquardt 法の BASIC プログラムを使用した。この程度のモデルであれば表計算ソフトの最適化プログラムでも十分な精度が得られる。

次に各パラメータの95%信頼区間を推定してみる。この場合の信頼域は、

$$Y_0 \leq Y_{\min} + \chi^2(3)_{0.05} \quad (44)$$

を満たす Y_0 を与えるパラメータ値の集合である。 $\chi^2(3)_{0.05} = 7.815$ であるから、結局、「 $Y_0 \leq 27.813$ 」という制限条件を与えて、パラメータ値を検索すればよい。この制限条件下で $(k - 0.47850)$ が最大となるように、3つのパラメータを最適化プログラムで求める。初期値として $(k, w_\infty, t_0) = (0.5, 95, 0.5)$ を与えると、 $(0.55505, 96.803, 0.55806)$ という解が得られた ($Y_0 = 27.813$)。収束の精度が悪いので必ず2、3回反復させる必要がある。次に $(k - 0.47850)$ が最小となるように、初期値として $(0.45, 105, 0.45)$ を与えると、 $(0.41041, 105.417, 0.41377)$ という解が得られた。つまり k の信頼区間は $k = 0.41041 \sim 0.55505$ と推定された (Fig. 1)。他のパラメータについても同様である。 w_∞ については $(0.42225, 106.500, 0.45792)$ 、および $(0.53714, 95.796, 0.52035)$ という解が得られたので、 $w_\infty = 95.796 \sim 106.500$ と推定された。 t_0 については $(0.52298, 99.575, 0.59540)$ 、および $(0.42975, 102.720, 0.37515)$ という解が得られたので、 $t_0 = 0.37515 \sim 0.59540$ と推定された。

パラメータ空間における信頼域は直方体 $(0.41041 \leq k \leq 0.55505, 95.796 \leq w_\infty \leq 106.500, 0.37515 \leq t_0 \leq 0.59540)$ の全体ではなくて、その中の一部分である。線型モデルの場合に

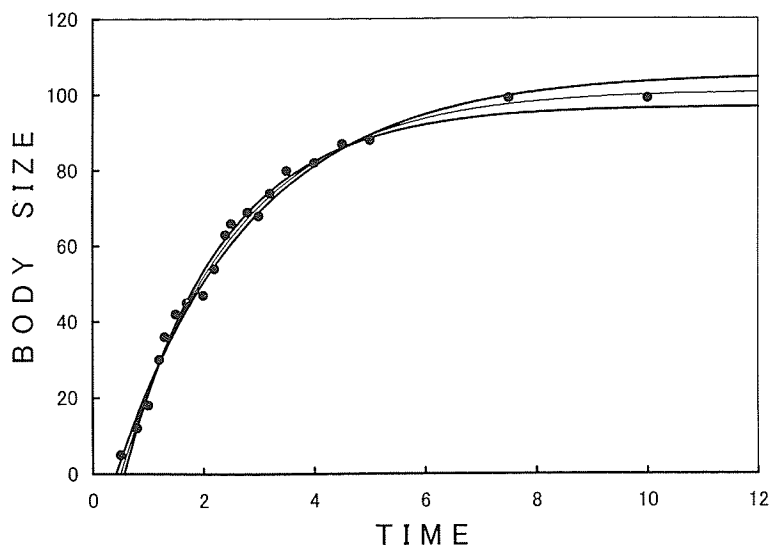


Figure 1. The thin curve shows the optimum solution of von Bertalanffy growth formula for Table 1 data, and the two thick curves show k limits of its 95% confidence domain in the parameter space (k, w_z, t_0).

は直方体に内接する楕円体（信頼域楕円）であるが、非線型モデルの場合には細長いバナナ状の空間となる。さらに一部のパラメータに制限がある場合には、信頼域は狭くなるから、各パラメータの信頼区間は短くなる。例えばこのデータでは $t=0.5$ で既に体長データが得られているのであるから、「 $t_0 \leq 0.5$ 」という制限は妥当と考えられる。つまり t_0 の上限を 0.5 に下げることができる。このことから $t_0=0.5$ と固定して k の最大値を求めると、(0.53876, 96.802, 0.5) という解が得られるので、 k の信頼区間の上限も 0.53876 に下げることができる。同様にして w_z の最小値を求めると、(0.52860, 95.926, 0.5) という解が得られるので、 w_z の下限を 95.926 に上げることができる。なお「 $t_0=0.49554$ 」のようにパラメータを完全に固定した場合には、(44) 式における χ^2 分布の自由度が減少するので注意する必要がある。

Akamine (1986) では F 検定を便宜的に使用したが、前述のように分散が既知の場合には χ^2 検定の方がよい。また当時は制限条件付きの最適化プログラムを手軽に利用できなかったため、ヘッセ行列の逆行列（分散共分散行列）の最大固有ベクトル上において区間推定を行った。この方法は非線型性の強いモデルにはあまり有効でないが、非線型性の弱いモデルやパラメータ数が多いモデルでは有効である（赤嶺、1985；Akamine, 1987）。なおここで示した方法で区間推定を行う場合には、パラメータの初期値が重要となる。Fig. 1 から明らかなように、一般に VB 式をデータに当てはめる場合には、 k と w_z とは負の相関を、 k と t_0 とは正の相関を持つ。このような特性から適切な初期値を設定できたわけであるが、一般の成長式の場合ではこのような相関が自明でない。前述の分散共分散行列を計算する方法

もあるが、むしろ表計算ソフトを用いて Fig. 1 のような作図をしながら、初期値を推定する方が効率的だろう。

まとめ

従来の主要な成長式を包含する式として Schnute (1981) 式が提唱されているが、これは Richards (1959) 式とまったく同一のものである。季節変動を示す成長式についても数学的に同一のものが様々な数式として提示されている。これらの中でもっとも簡潔な Akamine (1993) の標準型 (AR 式) を使用すべきである。また検定や区間推定においては、分散が既知の場合には χ^2 検定を、分散が未知の場合には F 検定を用いるべきである。

謝辞

この論文の初期原稿について遠洋水産研究所の平松一彦室長、および兵庫県立農林水産技術総合センター水産技術センターの五利江重昭主任研究員にはいくつかの誤りを指摘していただいた。また匿名の 2 名の査読者には提出原稿の不備な部分を指摘していただいた。ここに記して深謝いたします。

引用文献

- 相澤 康・滝口直之 (1999) MS-Excel を用いたサイズ度数分布から年齢組成を推定する方法の検討。水産海洋研究, 63, 205-214。
 赤嶺達郎 (1985) Polymodal な度数分布を正規分布へ分解する BASIC プログラムの検討。日本研報, 35, 129-159。
 Akamine, T. (1986) Expansion of growth curves using a periodic function and BASIC programs by Marquardt method. Bull. Jap. Sea Reg. Fish. Res. Lab., 36, 77-107。
 Akamine, T. (1987) A solution of the multi-cohort model by Marquardt's method. Bull. Jap. Sea Reg. Fish. Res. Lab., 37, 225-257。

- 赤嶺達郎 (1987) 最小二乗法と AIC. 日本海区水産試験研究連絡ニュース, **339**, 9-11.
- Akamine, T. (1988) Estimation of parameters for Richards model. Bull. Jap. Sea Reg. Fish. Res. Lab., **38**, 187-200.
- Akamine, T. (1993) A new standard formula for seasonal growth of fish in population dynamics. Nippon Suisan Gakkaishi, **59**, 1857-1863.
- 赤嶺達郎 (1994) 指数関数の定義式と資源解析学への応用. 水産海洋研究, **58**, 317-320.
- 赤嶺達郎 (1995) 水産資源学における成長式に関する数理的な研究. 中央水研報, **7**, 189-263.
- 赤嶺達郎 (1997) 成長式の一般化. 水産動物の成長解析, 赤嶺達郎・麦谷泰雄編, 恒星社厚生閣, 東京, 52-61.
- 赤嶺達郎 (2001a) 成長式. 資源解析手法教科書, 日本水産資源保護協会, 44-50.
- 赤嶺達郎 (2001b) 尤度と AIC. 資源解析手法教科書, 日本水産資源保護協会, 90-91.
- Akamine, T., H. Kishino and K. Hiramatsu (1992) Non-biased interval estimation of Leslie's removal method. Bull. Jap. Sea Reg. Fish. Res. Lab., **42**, 25-39.
- 五利江重昭 (2001) MS-Excel を用いた成長式のパラメータ推定. 水産増殖, **49**, 519-527.
- 五利江重昭 (2002) MS-Excel を用いた混合正規分布のパラメータ推定. 水産増殖, **50**, 243-249.
- Haddon, M. (2001) Modelling and quantitative methods in fisheries. Chapman and Hall, New York, 406 pp.
- Hilborn, R. and C. J. Walters (1992) Quantitative fisheries stock assessment. Chapman and Hall, New York, 570 pp.
- 石黒真木夫 (1993) 統計モデルと情報量規準. 水産資源解析と統計モデル, 松宮義晴編, 恒星社厚生閣, 東京, 72-81.
- 岸野洋久 (1993) 調査データと統計モデル. 水産資源解析と統計モデル, 松宮義晴編, 恒星社厚生閣, 東京, 90-103.
- 岸野洋久 (1999) 生のデータを料理する. 日本評論社, 東京, 227 pp.
- Kiso, K., T. Akamine, S. Ohnishi, and Y. Matsumiya (1992) Mathematical examination of the growth of sea-run and fluvial forms of female masu salmon *Oncorhynchus masou* in river of the Sanriku district, Honshu, Japan. Nippon Suisan Gakkaishi, **58**, 1779-1784.
- McGarvey, R. and A. J. Fowler (2002) Seasonal growth of King George whiting (*Sillaginodes punctata*) estimated from length-at-age samples of the legal-size harvest. Fish. Bull., **100**, 545-558.
- 森本光生 (1987) パソコンによる微分方程式. 朝倉書店, 東京, 196 pp.
- 能勢幸雄・石井丈夫・清水誠 (1988) 水産資源学. 東京大学出版会, 東京, 217 pp.
- Pitcher, T. J. and P. D. M. Macdonald (1973) Two models for seasonal growth in fishes. J. Appl. Ecol., **10**, 599-606.
- Quinn, T. J. and R. B. Deriso (1999) Quantitative fish dynamics. Oxford Univ. Press, New York, 542 pp.
- Richards, F. J. (1959) A flexible growth function for empirical use. J. Exp. Bot., **10**, 290-300.
- 坂元慶行 (1985) カテゴリカルデータのモデル分析. 共立出版, 東京, 221 pp.
- 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, 233 pp.
- Schnute, J. (1981) A versatile growth model with statistically stable parameters. Can. J. Fish. Aquat. Sci., **38**, 1128-1140.
- Schnute, J. T. and L. J. Richards (1990) A unified approach to the analysis of fish growth, maturity, and survivorship data. Can. J. Fish. Aquat. Sci., **47**, 24-40.
- Taylor, C. C. (1962) Growth equations with metabolic parameters. J. Cons. Int. Explor. Mer., **27**, 270-286.
- 東海 正 (1997) MS-Excel のソルバーによる曳網の網目選択性 Logistic 式パラメータの最尤推定. 水産海洋研究, **61**, 288-298.
- 鷺尾泰俊 (1978) 推定と検定. 共立出版, 東京, 123 pp.