

# 系列データに対する単回帰モデルの 適合性を診断する手法とその適用

統計数理研究所 柳 本 武 美  
食品総合研究所 柳 本 正 勝

(1983年9月 受付)

## 1. 序

年度毎に指標が与えられるような系列  $y_i, i=1, \dots, n$  に対しては、通常先ず単回帰モデルが  
あてはめられる。経年的な増加傾向あるいは減少傾向が認められるか否かが第一義的な解析で  
あることが多いからである。単回帰モデルをあてはめた後に更に詳しい解析を行う。その解析  
は入手可能なデータ、データとは別の知見、あるいは解析の目的に依存する。本稿ではデータ  
が  $(i, y_i), i=1, \dots, n$  に限定され、回帰線が直線ではなく曲線であるかもしれないと考えられ  
るような場合を扱う。モデルとしては

$$(1.1) \quad y_i \sim f(i) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

で  $\varepsilon_i$  は互いに独立とする。あるいは制約のあるモデルとして

$$(1.2) \quad y_i \sim \alpha + \beta i + \varepsilon_i$$

をも考える。

単回帰モデル (1.2) の適合性をより一般的なモデル (1.1) に対して診断する方法として最も  
ポピュラーな方法は  $f(i)$  を特定して統計的検定法を適用することである。例えばモデル (1.1)  
において  $f(i) = \alpha + \beta i + \gamma i^2$  として2次式を仮定して  $F$ -検定を行う。 $f(i)$  としてより一般的な  
多項式もしばしば用いられる。関数  $f(i)$  を特定しない方法としては残差プロット (例えば  
Chatterjee・Price(1977)), Durbin-Watson 法が良く知られている。残差プロットは教科書に  
でてくるような典型的な例を除けば、主観的な判断に依存しすぎる。Durbin-Watson 法は元来  
 $\varepsilon_i$  間の相関性に着目した量であり、 $f(i)$  について何らの情報も与えない。逆に  $f(i)$  に多項式を  
仮定する方法は  $f(i)$  についての仮定は実際以上に明示的であり過ぎる。実際単回帰モデルより  
2次回帰モデルの方が適合が良いことが分っても、推定された2次式は受け入れ難いことが多  
い。2次回帰モデルが単回帰モデルの適合の良さを調べるために道具として用いた場合でも、2  
次回帰モデルそのものに特別な意味がないからである。

本稿では  $f(i)$  に関して明示的に関数系を仮定せず、滑らかさに対する制約をおき、 $f(i)$  に対  
する推定量を得ると同時に単回帰モデルの適合性を診断する方法を提案する。この方法は超母  
数を用いた推定 (例えば Dempster ら (1977)) あるいは赤池(1980) による ABIC の延長上に  
位置する。 $f(i)$  に滑らかさの制約をおく方法として Schiller (1973) によるアプローチがある  
が本稿での方法とはその展開が違ふ。後述するようにこのアプローチは Durbin-Watson 法、2  
次回帰モデルを用いた検定にひとつの視点を与える。

提案した方法を食品消費の周年化傾向を示すデータに適用し、満足できる結果を与えること  
を示す。



$$\bar{\mu} \sim N \left( \begin{pmatrix} \alpha \\ \beta \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \frac{\sigma^2}{d^2} \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & \frac{1}{\lambda_3} & & \\ & & & \ddots & \\ & & & & \frac{1}{\lambda_n} \end{pmatrix} \right)$$

とする。このとき尤度  $L=L(\mu, \sigma^2, d^2)=L(\alpha, \beta, \bar{\mu}_3, \dots, \bar{\mu}_1, \sigma^2, d^2)$  は  $\mu_0 = \alpha e_1 + \beta e_2$  として次のように書ける (例えば Rao (1973), pp. 527)。

$$L = \left( \frac{1}{\sqrt{2\pi} \sigma} \right)^n \exp \left( -\frac{1}{2\sigma^2} \|y - \mu\|^2 \right) \left( \frac{d}{\sqrt{2\pi} \sigma} \right)^{n-2} \sqrt{\text{Vol}(D_2)} \exp \left( -\frac{d^2}{2\sigma^2} (\mu - \mu_0)' D_2 (\mu - \mu_0) \right),$$

ここで、 $\text{Vol}(D_2) = \prod_{i=3}^n \lambda_i$  であり、 $\mu$  の台は  $(n-2)$  次元の空間  $\{\mu \mid \bar{\mu}_1 = \alpha, \bar{\mu}_2 = \beta\}$  である。この式は Schiller (1973) の (8) 式に見られるが、 $\bar{\mu}_1$  と  $\bar{\mu}_2$  についてより明示的に与えた。また赤池 (1980) の 4 節例 *b* における  $D_2$  の定義で、原著での記号で、 $\alpha = \beta = 0$  において  $\mu$  の分布を退化させた場合になる。 $L$  を変形して

$$L = \left( \frac{1}{\sqrt{2\pi} \sigma} \right)^{2n-2} d^{n-2} \sqrt{\prod_{i=3}^n \lambda_i} \cdot \exp \left[ -\frac{1}{2\sigma^2} \left\{ y'(I - (I + d^2 D_2)^{-1})y + (\mu - (I + d^2 D_2)^{-1}y)'(I + d^2 D_2)(\mu - (I + d^2 D_2)^{-1}y) \right\} \right]$$

を得る。上式右辺第 2 項について

$$(\mu - (I + d^2 D_2)^{-1}y)'(I + d^2 D_2)(\mu - (I + d^2 D_2)^{-1}y) = (\bar{\mu} - T\bar{y})'(I + d^2 \Lambda)(\bar{\mu} - T\bar{y})$$

ただし

$$\Lambda = \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & & & 0 \\ & & \lambda_3 & & \\ & 0 & & \ddots & \\ & & & & \lambda_n \end{pmatrix}$$

$$T = (I + d^2 \Lambda)^{-1}$$

が成り立つ。また

$$\det(I + d^2 D_2) = \det(I + d^2 \Lambda) = \prod_{i=3}^n (1 + \lambda_i d^2)$$

に注意する。他の母数を固定して  $L$  を最大にする  $\mu$  を求めることと、 $\bar{\mu}(\bar{\mu}_1 = \alpha, \bar{\mu}_2 = \beta)$  を求めることは同値である。求める値は  $\hat{\bar{\mu}} = T\bar{y}$  となる。これは  $\bar{\mu} = (I + d^2 D_2)^{-1}y$  と書き変えることができる。次に  $d^2$  と  $\sigma^2$  の推定量を得るために  $\bar{\mu}$  について周辺尤度を求めると

$$\text{ML} = \left( \frac{1}{\sqrt{2\pi} \sigma} \right)^{n-2} d \cdot n^{-2} \sqrt{\prod_{i=3}^n \lambda_i} \cdot \frac{1}{\sqrt{\prod_{i=3}^n (1 + \lambda_i d^2)}} \exp \left( -\frac{1}{2\sigma^2} y'(I - (I + d^2 D_2)^{-1})y \right)$$

を得る。ML を最大にする  $\sigma^2$  は

$$\sigma^2 = \frac{1}{n-2} y'(I - (I + d^2 D_2)^{-1})y$$

となる。これを ML に代入して  $-2 \log \text{ML}$  を求め、 $d^2$  を  $x$  と書いて定数を無視することに

よって(2.1)式の右辺を得る。 $U(x)$ を最小にする $x$ を求めて $\hat{x}$ を得、これを代入して $U(\hat{x})$ ,  $\hat{\mu}$ ,  $\hat{\sigma}^2$ を推定する。

$U(x)$ は $-2 \log(\text{周辺尤度}) + \text{定数}$ であるから、 $U(x)$ が大きいことはモデルのあてはまりが悪いことを示している。

#### 4. 方法の吟味

2, 3節では方法とその導出を形式的に記述した。本節では具体的に吟味する。根本的な問題として $\mu$ に構造を仮定すること及び周辺尤度を適用することがあるけれども、ここではこれらは前提として論じる。

この方法の適用についての困難は $U(x)$ の数値計算と $U(x)$ の関数としての振舞が不明な点にある。ある $x$ に対して $U(x)$ を計算する方法は種々考えられるが、 $U(x)$ を最小にする $x$ を計算する必要があることを考えれば固有値と固有ベクトルを求めておいた方が良いと思われる。行列 $D_2$ の固有値と固有ベクトルは陽には表現されないので数値計算によらなければならない。高速計算機が使用できる場合はこの計算は容易であるが、小型計算機では実行困難になる恐れがある。

より困難な問題は $U(x)$ の振舞である。 $x$ の関数として必ずしも単純ではない。先ず実際的に有用な微係数及び $x \rightarrow 0, \infty$ での極限での性質を求める。特に $x$ が十分大きいときの振舞を調べることが重要である。

$$(4.1) \quad \lim_{x \rightarrow 0} U(x) = (n-2) \log \sum \lambda_i \bar{y}_i^2 \quad (= U_0)$$

$$(4.2) \quad U'(x) = (n-2) \frac{-\sum \bar{y}_i^2 \frac{\lambda_i^2}{(1+\lambda_i x)^2}}{\sum \bar{y}_i^2 \frac{\lambda_i}{1+\lambda_i x}} + \sum \frac{\lambda_i}{1+\lambda_i x}$$

$$= (n-2) \frac{\sum \bar{y}_i^2 \frac{\lambda_i}{x(1+\lambda_i x)^2}}{\sum \bar{y}_i^2 \frac{\lambda_i}{1+\lambda_i x}} - \sum \frac{1}{x(1+\lambda_i x)}$$

を得る。(2.3)式と(4.1)式から

$$(4.3) \quad U_0 - U_1 = (n-2) \log \left( \sum \lambda_i \bar{y}_i^2 / \sum_{i=3}^n \bar{y}_i^2 \right) - \sum \log \lambda_i$$

が導かれる。一方 $x$ が十分大きいとき $U'(x)$ は

$$U'(x) \doteq \frac{1}{x^2} \left\{ \frac{(n-2) \sum_{i=3}^n \frac{1}{\lambda_i} \bar{y}_i^2}{\sum_{i=3}^n \bar{y}_i^2} - \sum \frac{1}{\lambda_i} \right\}$$

で近似される。従って $x$ が十分大きいとき、

$$(4.4) \quad \sum \frac{1}{\lambda_i} \bar{y}_i^2 / \sum \bar{y}_i^2 > \sum \frac{1}{\lambda_i} / (n-2)$$

であれば、 $U'(x) > 0$ になる。

$U(x)$ の最小を与える $x$ を求めるためには、 $U(x)$ が上に凹であれば都合が良い。しかしながら数値計算によれば多くの場合に上に凹になるが、このことは常には成立しないことが確かめられる。実際(4.4)式で逆の不等号が成立するが、正の $x$ で $U(x)$ は最小を与えることがある。従って $U(x)$ の最小値を求めるために多くの計算が必要になる。しかし $U(x)$ を求めるための計算量は固有値と固有ベクトルを用いて(2.2)式のように簡単に表現できるので膨大な量には

ならない。

### 5. 方法の拡張と既存の方法

2節で方法を提出したとき、我々は  $D_2$  を天下一りに決めた。提案した方法が示すように  $D_2$  を選択したことはこの方法のひとつの重要な点になっている。 $D_2$  として他の行列を選ぶことができ、これによって提案した方法を拡張することができる。

行列  $D_2$  とその固有値  $\lambda_1, \dots, \lambda_n$ , 固有ベクトル  $e_1, \dots, e_n$  は本質的に一対一に対応する。我々の問題の設定から、候補として行列の固有値  $\tau_1, \dots, \tau_n$ , 固有ベクトル  $\nu_1, \dots, \nu_n$  には  $\tau_3 = \tau_2 = 0, \nu_1 = e_1, \nu_2 = e_2$  なる制約を置いて良い。我々が行列  $D_2$  を選択した根拠は必ずしも確固なものではなく、経験的、直観的である。 $D_2$  は時系列解析に良く用いられポピュラーである。また Cline (1973), 大西(1975) らの正当性の指摘がある。

ところで  $D_2$  を選択したとき、固有値と固有ベクトルを計算することに困難があった。逆に固有値と固有ベクトルを共に、何らかの規準で決められれば計算量が大幅に減少する。固有ベクトルの組を決定することは、直交関数系を決定することに対応する。1つの直交関数系として Legendre の多項式がある。これに対応するベクトルとして、 $u_i = (1^i, 2^i, \dots, n^i)$  として Gram-Schmidt の直交化を行ない、ノルムを1に正規化して  $\omega_1, \dots, \omega_n$  を得る。この固有ベクトルについて  $\omega_1$  と  $\omega_2$  の張る空間は  $e_1$  と  $e_2$  の張る空間と同じである。

例1. 上記の  $\omega_1, \dots, \omega_n$  に対して固有値を  $\lambda_1 = \lambda_2 = 0, \lambda_3 = 1, \lambda_4 = \dots = \lambda_n = K (> 1)$  とする。2次の項を表すベクトル  $\omega_3$  に対する固有値  $\lambda_3$  のみ中間の値とする。行列  $D_2$  の代りに  $\sum \lambda_i \omega_i \omega_i'$  を入れると、 $U(x)$  は

$$U(x) = (n-2) \log [z_3^2 \{(1+Kx)/(1+x)\}^{n-1} + S_3 K \{(1+x)/(1+Kx)\}^{n-1}]$$

ただし、

$$z_i = \omega_i' y_i \quad i=1, \dots, n, \quad S_3 = \sum_{i=4}^n z_i^2.$$

$U_1 = (n-2) \log(z_3 + S_3) + (n-3) \log K$  だから  $U(x) > U_1 - c$  は適当な  $c'$  に対して  $z_3/S_3 > c'$  と書き換えられる。一方通常の検定問題  $H_0: f(i) = \alpha + \beta i, H_1: f(i) = \alpha + \beta i + \gamma i^2$  に対して水準を  $\alpha$  とすると  $z_3^2 / (S_3/n-3) > F_{n-3,1;\alpha}$  が棄却域となる。

次に回帰線の推定に関しては、 $\hat{x}$  で  $U(x)$  が最小になるとき (2.5) 式によって  $\hat{\mu}$  が推定される。これは

$$\hat{\mu} = W' \Lambda_1 W y$$

ただし、 $W = (\omega_1, \dots, \omega_n)'$ ,  $\Lambda_1$  は対角要素が  $\delta_1 = \delta_2 = 1, \delta_3 = 1/(1+\hat{x}), \delta_4 = \dots = \delta_n = 1/(1+K\hat{x})$  なる対角行列、で表わされる。一方、2次式の回帰の場合には

$$\hat{\mu} = W' \Lambda_2 W y$$

ただし  $\Lambda_2$  は対角要素が  $\delta_1 = \delta_2 = \delta_3 = 1, \delta_4 = \dots = \delta_n = 0$  で推定する。

さて (2.1) 式の  $U(x)$  の定義で  $D_2$  の代りに  $y' D_1 y = \sum_{i=1}^{n-1} (y_i - y_{i+1})^2$  を満たす行列  $D_1$  を用い、 $(n-2)$  の代りに  $(n-1)$  として  $y_i$  の代りに  $g_i = y_i - \hat{\alpha} - \hat{\beta} i$  を代入することを考える。行列  $D_1$  の固有値と固有ベクトルとは陽に表現されるので計算は容易になる。 $D_1$  では0の固有値は1つでその固有ベクトルは  $(1, \dots, 1)'$  である。そのため  $y_i$  を用いると直線性の適合ではなくて、 $y_i$  の同等性の診断になる。そこで  $y_i$  の代りに  $g_i$  を用いる。著者らの計算例では直線性の診断

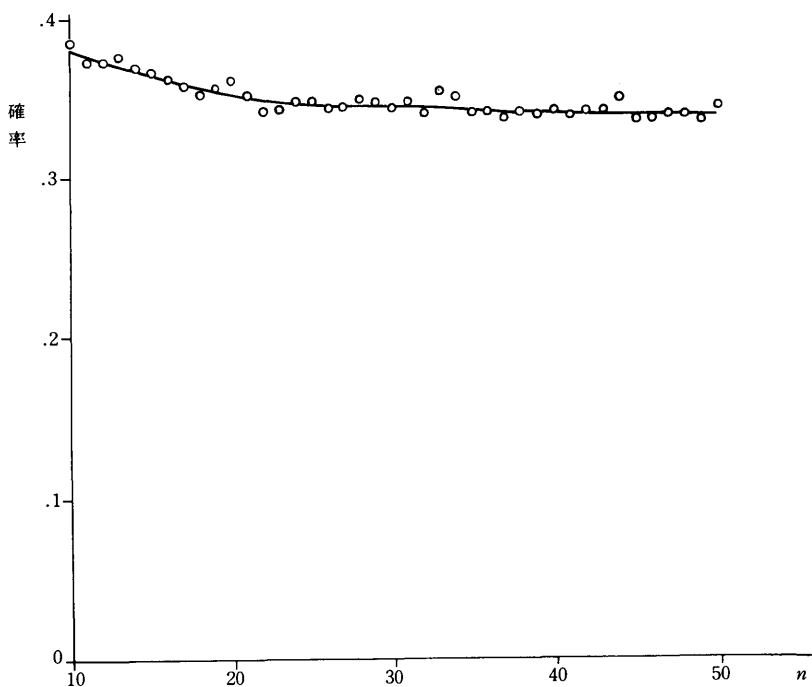


図 1.-1. 長さ  $n$  を 10(1)50 として単回帰の適合の良さを表わす統計量  $U_1-U_m$  が  $c$  より大きくなる確率のシミュレーション結果, 図中折線は提案した方法によるスムーズ化した回帰線;  $c=0$  の場合

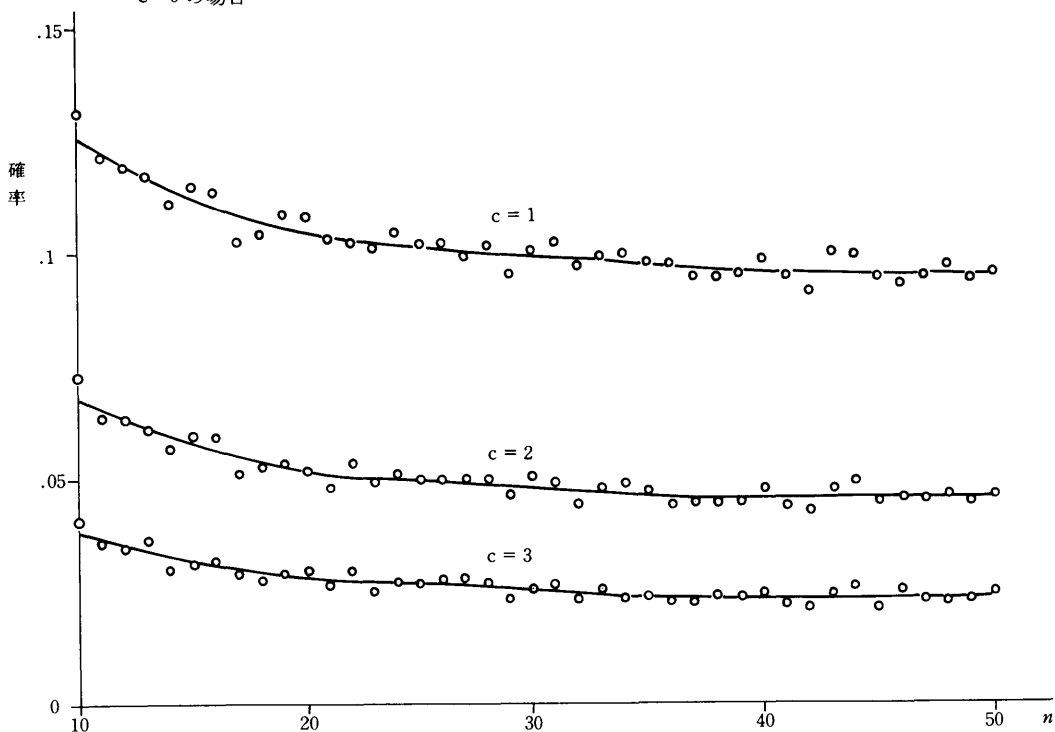


図 1.-2.  $c=1, 2, 3$  の場合

としてはほぼ同様の結果を得るが、得られる  $\hat{\mu}$  の推定を図示すると起伏がはげしくなる。

この方法は理論的に不明瞭であるが、任意のパラメトリックな回帰モデル  $f(i)$  について  $g_i = y_i - \hat{f}(i)$  が用いることができ適用範囲が広い。

例2. 上記の仮定の下で形式的に演算を行えば、(2.3)より、

$$U_1 = (n-1) \log \sum (g_i - \bar{g})^2 + \sum_{i=1}^{n-1} \log (2(1 - \cos i\pi/n)),$$

また(4.1)より

$$U_0 = (n-1) \log g' Dg = (n-1) \log \sum_{i=1}^{n-1} (g_i - g_{i+1})^2,$$

ただし  $g = (g_1, \dots, g_n)'$  とする。従って

$$U_0 - U_1 = (n-1) \log \left\{ \frac{\sum (g_i - g_{i+1})^2}{\sum (g_i - \bar{g})^2} \right\} - \sum \log (2(1 - \cos i\pi/n))$$

となる。ここで  $DW = \frac{\sum (g_i - g_{i+1})^2}{\sum (g_i - \bar{g})^2}$  は Durbin-Watson 検定量として知られていることに注意する。

### 6. 規準値 $c$ の選定

2節で提案した手法では(2.4)式の  $c$  を具体的に定めていない。 $U_m$  は  $-2 \log$ (周辺尤度)であるが、通常対数尤度で成り立つ漸近理論は成立しない。そこでシミュレーションによって  $c$  を設定する。

規準値  $c$  の設定は検定論に基づいて行う。即ち  $y_i$  が線型であるという仮定(1.2)の下で  $U_1 - U_m$  の分布を求めた。シミュレーションは  $n=10 \sim 50$  について行った。 $x = e^s$ ,  $s = -15(0.5)15$  について  $U(x)$  を求め、これらと  $U_1$  の最小値を  $U_m$  とし、更に  $U_1$  が最小でないとき最小な  $U(x)$  を与える  $\hat{s}$  に対し、 $s$  は  $\hat{s} - 0.5(0.05)\hat{s} + 0.5$  について最小な  $U(x)$  を求めた。

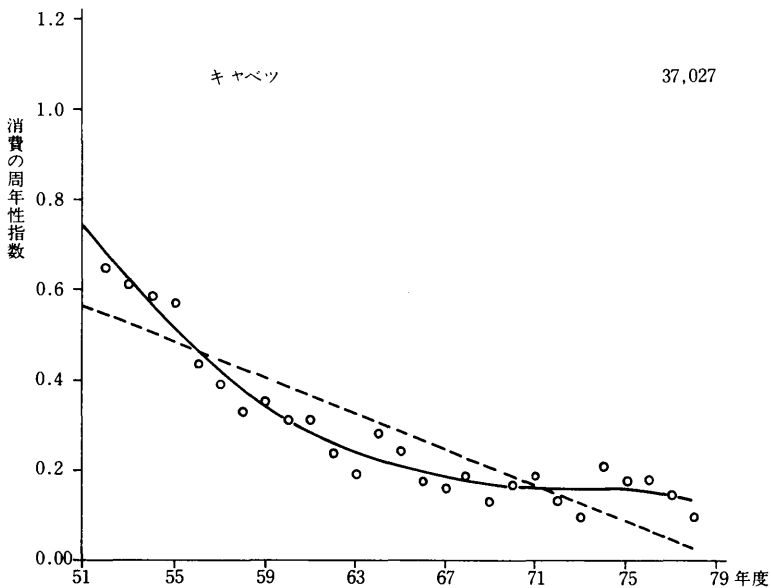


図2.-1. 提案した方法による回帰を単回帰, 図中の数字は,  $U_1 - U_m$  キャベツの場合

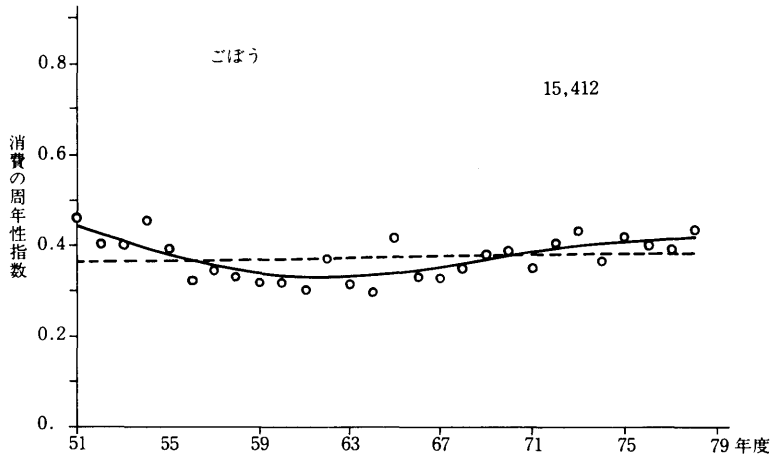


図 2.-2. ごぼうの場合

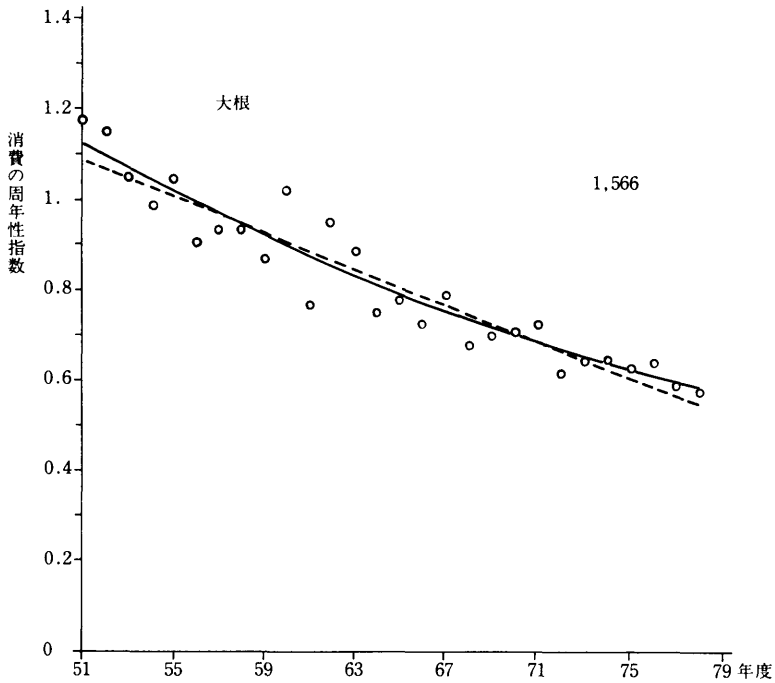


図 2.-3. 大根の場合

各  $n$  について 1 万回の繰り返しで、 $U_1 - U_m > c$  となる回数を  $T_{n,c}$ 、 $c=0, 1, 2, 3$  としたときの値を図 1 に示す。

$T_{n,c}$  は  $n$  について減少する傾向が窺える。 $c$  としては区切りが良く、且つ水準を 5% としたときの棄却限界に近い値として  $c=2$  が選ばれる。



## 7. 適用

本節では提案した方法を具体的な問題に適用する。ここで扱う問題は食品消費の周年化傾向を扱う。近年の生産技術の進歩、輸送の保存技術の向上から高値となる端境期を狙った出荷が可能になり、収穫期に依存した季節性の消費が周年消費に移ってきた。これを定量的にとらえることがこの研究の目的である。

### 7-1. 周年化傾向 (柳本ら(1982))

総理府統計局による「家計調査年報」(1951,2年は「消費実態調査年報」)より1951~1978年での各年各月の消費量(調査では購入量)を得る。第*i*年*j*月の消費量を $x_{ij}$ としたとき、第*i*年度の消費の周年性度を表わす指数 $r_i$ を各月日数で補正した変動係数で定義する。即ち、

$$r_i = \sqrt{\sum_{j=1}^{12} (X_{ij} - X_i./12)^2 / 11}$$

ただし、

$$X_{ij} = (x_{ij}/d_{ij}) / (x_i./d_i)$$

$d_{ij}$ : *i*年*j*月の日数。

データ( $i, r_i$ ) $i=1, \dots, n$ について単回帰分析を行なった。多くの食品で $\hat{\beta} < 0$ が得られ、それは統計的に有意であった。勿論傾向は区々であり、 $\hat{\beta}$ が大きい食品としてトマト、ピーマン、減少傾向の認められない食品としてごぼうがあった。

### 7-2. 回帰曲線

大筋において減少傾向が認められたとしたら、次にはより詳しい解析が望まれる。ひとつの関心は周年化が達成された状態にあるか否かの資料にしたい。周年性指数は常に正であり、ある程度周年化が進めば頭打ち状態になると思われる。

この目的には2次式のような多項式をあてはめることは適切でない。常識的に考えればロジスティック型の回帰曲線

$$f(i) = \gamma + \delta \exp(\alpha + \beta i) / (1 + \exp(\alpha + \beta i)) \quad (\delta > 0)$$

このモデルをあてはめて最尤推定量 $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}$ を得たとすると、 $\hat{\beta} < 0$ であれば $\hat{\gamma}$ が達成される周年性度の推定量になる。このモデルの単回帰モデルに対する適合は尤度比検定によって判断できる。しかしながらこのようなパラメトリックモデルでは仮定した回帰曲線は強く影響される。したがって各種の食品項目については全く不適当な場合がある。現実問題として母数の推定量を得ることが大変難しい場合がある。このようなことは仮定したモデルの適合が良くないときに起り易い。

そこでパラメトリックなモデルをあてはめる前処理として、また大まかな傾向を掴む解析として提案した方が役立つ。

### 7-3. 提案した方法の適用

野菜を解析した中からキャベツ、ごぼう、大根を例示する。単回帰モデルと提案した方法で推定した $\hat{\mu}$ を図2-1~3に示す。図で $\hat{\mu}_i, i=1, \dots, n$ については( $i, \hat{\mu}_i$ )を折線で結んで図示した。 $U_1 - U_m$ は各々37.027, 15.412, 1.566であった。キャベツの例ではロジスティック型の回帰曲線に似た形をしており、 $U_1 - U_m$ も大きい。これにロジスティック型モデルをあてはめると、

$$.1362 - 1.343 / (1 + \exp(-.9753 + .19463t))$$

ただし、 $t$  は西暦年で世紀分を除いた年度が得られる。これは推定された折線と殆ど一致する。ごぼうの例では単回帰モデル  $\hat{\beta}$  が非零性の検定が有意でないが、 $U_1 - U_m$  は大きい。推定された曲線は  $U$  型でロジスティック型にはほど遠い。大根の例では推定された折線はロジスティック型曲線に近いが、 $U_1 - U_m$  は我々が設定した 2 よりも小さい。シミュレーションの結果から判断してこの程度の傾向は直線性の仮定の下でも十分起る。

本研究において必要なプログラムの作製は小山晴生氏（東京工業大学生）によって手際良くなされた。本稿で提案した手法のコンピュータ・プログラム（フォートラン）は農林計算センター・ライブラリーに登録（柳本ら）の予定である。

### 参 考 文 献

- Akaike, H. (1980). Likelihood and the Bayes procedure, *Bayesian Statistics* (eds. J.M. Bernardo, M. H. DeGroot, D.V. Lindley and A.F.M. Smith), Univ. Press, Valencia, Spain, 143-166.
- Chatterjee, S. and Price, B. (1977). *Regression Analysis by Example*, John Wiley, New York.
- Cline, A.K. (1973). Curve fitting using splines under tension. *Atmospheric Technology*, NCAR, **3**, 60-65.
- Dempster, A.P. et al (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion), *J. Roy. Statist. Soc.*, B, 1-38.
- Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression I, *Biometrika*, **37**, 409-428.
- 大西行雄 (1975). スプライン法を用いた 2 次元補間について, *J. Oceanographical. Soc. Japan*, **31**, 259-264.
- Rao, C.R. (1973). *Linear Statistical Inference and its Application*, John Wiley & Sons, New York.
- Schiller, R.J. (1973). A distributed lag estimator derived from smoothing priors, *Econometrica*, **41**, 775-788.
- 柳本正勝ら (1982). 野菜の消費における周年化傾向, *食総研報*, **41**, 1-8.
- 柳本正勝ら. 系列データの平滑化と単回帰モデルの適合性の診断法, 農林水産研究計算センター・ライブラリー (準備中).

#### 校正時訂正:

119 頁の ML を導出した式で右辺に

$$\frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\{(\alpha - \bar{y}_1)^2 + (\beta - \bar{y}_2)^2\}\right)$$

を掛ける。この修正項から  $\hat{\alpha} = \bar{y}_1$ ,  $\hat{\beta} = \bar{y}_2$  を得る。以下の推論は本文の右辺のみに基づいて行う。3 節の以降の部分で ML は本文中の右辺を指すとみなす。

A Procedure to Diagnose the Goodness of Fit of  
the Simple Regression Model for Serial  
Data with Applications

Takemi Yanagimoto

(The Institute of Statistical Mathematics)

and Masakatsu Yanagimoto

(National Food Research Institute)

Under the assumption of smoothness of the regression curve, a novel procedure to diagnose the goodness of fit of a simple regression model is proposed by applying the marginal likelihood. Here we do not assume any specific form (e.g., polynomial) of the alternative regression function. A marvellous advantage of the procedure is that it provides an estimated regression curve (spline) as a byproduct. The existing procedures including the Durbin-Watson test statistic and residual plots are explored in relation to the proposed procedure.

The procedure is applied to data of trends in year-round consumption of the fresh vegetables in Japan.