非正則な統計モデルに基づく林分成長分析
Forest Stand Growth Analysis Based on Irregular Statistical Models

二宮 嘉行・柳原 宏和・吉本　敦
Ninomiya, Y., Yanagihara, H. & Yoshimoto, A.

キーワード:　局所錘モデル, 混合分布モデル, 森林計画学, 成長分析, 変化点モデル

要約:　林分内の間伐効果・競合効果・成長パターンの異質性などを検知する問題では, 非正則モデルに基づく分析が必要となるが, それらの分析手法には通常の統計理論を適用することはできない. 状況や検証したい仮説により使用する非正則モデルは異なり, 今までの非正則モデルに基づく分析では, それらのモデルごとに性質が議論されてきた. しかしながら, 近年それらを統一的に扱う理論が発展されつつある. 本論文では, 非正則モデルの主である局所錘モデルにおける検定問題に注目し, 通常の検定とどう異なるかを説明し, 林分成長分析におけるその必要性について検討する.

Abstract: In growth analysis, irregular statistical models are needed to detect some property in a forest stand, such as thinning effect, self-thinning effect or difference of growth patterns. The irregular models do not permit to use usual statistical theory, and then, the properties of them were discussed by each problem. Recently, however, unified statistical theory for the irregular models has been developed. In this paper, some tests based on the locally conic model, which is representative irregular model, are elaborated. In addition, differences from usual tests are explained, and the need of the irregular models is investigated in growth analysis.

Keywords: change-point model, forest planning, growth analysis, locally conic model, mixture model

1. はじめに

Candy (1997), Fang and Bailey (2001) などが一般化線形混合効果モデルあるいは非線形混合効果モデルを用いて林分成長分析を行っているように, 近年, より正確な予測を得るための複雑なモデリングおよびそれに基づいた分析が必要とされている. 例えば, 林分内の間伐効果の有無, 競合効果の有無, 成長パターンの変化などをモデリングに取り入れる場合も, それらを統計的に検知する必要があり, モデリングが複雑化する要因である. これらの検知の問題は非正則モデルとして取り扱われ, その分析手法には通常の統計理論を適用することはできない.

上記の例のみならず, 通常の統計理論を適用できないようなモデルの総称が非正則モデルである. 例えば正則モデルでの検定問題では, 尤度比検定統計量が漸近的にカイ二乗分布に従うという好ましい性質が通常成り立つが, 非正則モデルではそうならない. 状況や検証したい仮説により使用する非正則モデルは異なるため, それらのモデルごとに尤度比検定統計量の分布論が展開されてきた. 例えば, 二つ以上の分布を混ぜ合わせてできる分布である混合分布において, Ghosh and Sen (1985) やHartigan (1985) は, はたして本当に二つの分布が混在しているのか否かという問題に対する尤度比検定が通常とは異なることを指摘した. 他には, 時系列のARMA (AutoRegressive Moving Average; 自己回帰移動平均) モデルにおけるAR次数やMA次数がいくつなのかという検定問題 (Hannan 1982, Veres 1987), 折れ線を回帰関数とするモデルにおける節点数 (折れている回数) がいくつなのかという検定問題 (Feder 1975), あるタイプの非線形回帰項の有無を調べる検定問題 (Naiman 1986) などが非正則モデルの問題であり, それぞれのモデルにおいて尤度比検定統計量の分布論が議論されている. これに対して比較的最近, Dacunha-Castelle and Gassiat (1997) は局所錘モデルを導入し, 上で挙げた問題の多くを統一的に扱う理論を発展させた. さらに福水・栗木 (2004) は局所錘モデルを含む特異モデルを定義し, その性質を議論している.

本論文では非正則モデルの主である局所錘モデルにおける検定問題に注目し, 通常の検定とどう異なるかを説明し, 林分成長分析におけるその必要性について検討する. まず, 正則なモデルのもとでの尤度比検定の性質を 2 章で紹介し, 次に局所錘モデルの定義を 3 章でおこなう. 4 章では, そのモデルのもとでの検定の性質を紹介する. 5 章では, この局所錘モデルを用いた検定方法が, 間伐等を原因とする林分内の構造変化・林分内の競合あるいは空間相関・林分内の成長パターンの異質性, などを検知する問題に適用できることを示す. 6 章では, 結論をまとめるとともに今後の課題を挙げる.

2. 正則 (通常) なモデルにおける尤度比検定

まず正則 (通常) なモデルにおける尤度比検定の性質を紹介する. $X_1,\ldots,X_n$ をある確率分布 $f(x)$ から発生した独立な標本とする. 今, 確率分布が p 個のパラメータ $\theta_{01},\ldots,\theta_{0p}$ を用いて $f_0(x\,|\,\boldsymbol{\theta}_0)$ ($\boldsymbol{\theta}_0=(\theta_{01},\ldots,\theta_{0p})$) と書けるのか, あるいは q (>p) 個のパラメータ $\theta_{11},\ldots,\theta_{1q}$ を必要として $f_1(x\,|\,\boldsymbol{\theta}_1)$ ($\boldsymbol{\theta}_1=(\theta_{11},\ldots,\theta_{1q})$) となるのかということを検証する問題を考える. つまり, 帰無仮説 H_0 を $f(x)\in\{f_0(x\,|\,\boldsymbol{\theta}_0)\}$, 対立仮説 H_1 を $f(x)\notin\{f_0(x\,|\,\boldsymbol{\theta}_0)\}$ だが $f(x)\in\{f_1(x\,|\,\boldsymbol{\theta}_1)\}$, として統計的仮説検定を考える. ただし, H_1 の確率分布は H_0 の確率分布を含む, すなわち $\{f_0(x\,|\,\boldsymbol{\theta}_0)\}\subset\{f_1(x\,|\,\boldsymbol{\theta}_1)\}$, とする. さて, H_0 対 H_1 の検定統計量としてはいくつか考えられるが, ここでは通常用いられる尤度の比に基づく統計量である尤度比検定統計量,

[1]
$$T_0\equiv 2\log\frac{\prod_{i=1}^{n} f_1(X_i\,|\,\hat{\boldsymbol{\theta}}_1)}{\prod_{i=1}^{n} f_0(X_i\,|\,\hat{\boldsymbol{\theta}}_0)}=2\left\{\sum_{i=1}^{n}\log f_1(X_i\,|\,\hat{\boldsymbol{\theta}}_1)-\sum_{i=1}^{n}\log f_0(X_i\,|\,\hat{\boldsymbol{\theta}}_0)\right\}$$

を考える. ここで $\hat{\boldsymbol{\theta}}_1$, $\hat{\boldsymbol{\theta}}_0$ はそれぞれ $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_0$ の最尤推定量である. H_0 の受容・棄却は, T_0 の実現値 t_0 を用い, $P(T_0>t_0)$ という p 値の大小で決定される. ここで $P(T_0>t_0)$ は H_0 のもとでの確率である. 通常その p 値が 5 % を上回るか否かで受容・棄却が決定される. 実際の計算では $P(T_0>t_0)$ を厳密に求めることは難しいが, その近似のためには以下の定理が有用である (詳細は竹村 1991, 13.2 節参照).

定理 1．通常成り立つような緩い条件 (正則条件) が満たされれば, H_0 のもと [1] 式の T_0 は漸近的に自由度 $q-p$ のカイ二乗分布に従う. つまりその分布関数を $G_{q-p}(x)$ と記せば, $\mathrm{P}(T_0 > t_0)$ は $1-G_{q-p}(t_0)$ に収束する.

上の定理より p 値の近似が $1-G_{q-p}(t_0)$ によって容易に与えられるがゆえに, 尤度比検定は最もよく用いられる検定の一つとなっている. 以下に, それらの一般的な例を二つあげる.

例 1．正規分布モデル. 各々独立な確率変数 $X_1,\ldots,X_n$ は正規分布 $\mathrm{N}(\beta,1)$ に従うという状況において, 平均が 0 か否かの検定, つまり $H_0:\beta=0$ 対 $H_1:\beta\neq 0$ という検定を考える. 真の確率密度関数を $f(x)$ とし $\mathrm{N}(\beta,1)$ の確率密度関数を $\phi(x\,|\,\beta)$ とすれば, H_0 は $f(x)\in\{\phi(x\,|\,0)\}$, H_1 は $f(x)\notin\{\phi(x\,|\,0)\}$ だが $f(x)\in\{\phi(x\,|\,\beta)\}$, と書ける. $\phi(x\,|\,0)$ のパラメータ数は 0, $\phi(x\,|\,\beta)$ のパラメータ数は 1 であり, また当然 $\{\phi(x\,|\,0)\}\subset\{\phi(x\,|\,\beta)\}$ であるので, 定理 1 より尤度比検定統計量

[2] $$T_1 \equiv 2\left\{\sum_{i=1}^{n}\log\phi(X_i\,|\,\hat{\beta})-\sum_{i=1}^{n}\log\phi(X_i\,|\,0)\right\}$$

は漸近的に自由度 1 のカイ二乗分布に従うことがわかる. ここで, $\hat{\beta}$ は β の最尤推定量である.

例 2．混合比既知の混合分布モデル. 各々独立な確率変数 $X_1,\ldots,X_n$ は通常は標準正規分布 $\mathrm{N}(0,1)$ に従うが, 確率 α で平均 β をもった正規分布 $\mathrm{N}(\beta,1)$ に従うこともある, という状況を想定する. このとき X は $(1-\alpha)\phi(x\,|\,0)+\alpha\phi(x\,|\,\beta)$ という密度関数をもつ確率分布に従うことになる. なお, この分布は 1 章で述べた混合分布の一つであり, α は混合比とよばれる. 今, この混合比 α が既知で 1/2 であるとする. このとき, 二つの分布が混在しているか否かの検定は $H_0:\beta=0$ 対 $H_1:\beta\neq 0$ で構成される. 一般にこの検定の検定統計量の分布を陽に求めるのは困難であるが, 尤度比検定統計量の漸近分布ならば容易に求めることができる. つまり, パラメータ数が 1 である H_1 の分布 $(1/2)\phi(x\,|\,0)+(1/2)\phi(x\,|\,\beta)$ は, パラメータ数が 0 である H_0 の分布 $\phi(x\,|\,0)$ を含むので, 定理 1 より漸近分布が自由度 1 のカイ二乗分布となることがわかる.

3. 局所錐型パラメトリゼーション

2 章の例 2 では, 混合比が既知である, つまり既に決定されているという状況を仮定している. そのとき, 定理 1 の正則条件は満たされているため, カイ二乗分布を用いて容易に検定を行うことができた. しかしながら, 混合比が未知であるという状況を仮定した途端, 定理 1 の正則条件が満たされなくなり, 自由度 $q-p$ のカイ二乗分布を検定に用いることができなくなる. そのため, 混合比が未知の場合, 定理 1 とは異なる分布を検定に用いる必要がある. このような設定を一般化するために, Dacunha-Castelle and Gassiat (1997) は局所錘モデルを導入した. 今, B をある有界閉集合 (例えば閉区間) として, 確率分布の集合 $\Phi=\{f(x\mid\alpha,\beta)\mid\alpha\geq 0,\beta\in \mathrm{B}\}$ を考える. このとき

(A1)　$\alpha=0$ のとき, $f(x\mid\alpha,\beta)$ は β の値によらず $f_0(x)$ となる. つまり任意の β に対し $f_0(x)=f(x\mid 0,\beta)$ となる.

(A2)　$\alpha\neq 0$ のとき, $f(x\mid\alpha,\beta)$ は $f_0(x)$ とはならない.

が成り立つようなモデル Φ を局所錐モデルという. なお, ここでは簡単のために細かい条件を省略したが, 詳細は Dacunha-Castelle and Gassiat (1997) を参照していただきたい. この局所錐という名称は, $f_0(x)$ を錘の頂点として, α と β がそれぞれその頂点からの「距離」と「方向」を意味するパラメータと解釈されることから付けられている. また, $\beta\neq\beta'$ であっても $f(x\mid 0,\beta)$ と $f(x\mid 0,\beta')$ を識別できないことから, Φ は識別不能性をもつモデルともいう.

真のモデルがこの頂点 $f_0(x)$ となっているケースでは, $X_1,\ldots,X_n$ がその独立な標本であっても, 最尤推定量は漸近的に正規分布に従わず, 尤度比検定統計量も漸近的にカイ二乗分布に従わない. つまり, 局所錘モデルにおける頂点か否かの検定では, 定理 1 の結果を用いることができない. 以下の例 3 はその検定の例である.

例 3 . 混合比未知の混合分布モデル (その 1). 2 章の例 2 での混合分布モデル $\{(1-\alpha)\phi(x\mid 0)+\alpha\phi(x\mid\beta)\}$ の混合比 $\alpha(\geq 0)$ が未知で, かつ $\beta\neq 0$ とする. このモデルにおいて $\alpha=0$ とすると分布は β の値によらず $\phi(x\mid 0)$ となり, 先の条件 (A1) が成り立つ. このとき, $f_0(x)=\phi(x\mid 0)$ である. また, $\alpha\neq 0$ ならば分布は $\phi(x\mid 0)$ とならないため, 条件 (A2) も成り立つ. したがって混合分布モデルは局所錘モデルであり, $\alpha=0$ の分布 $\phi(x\mid 0)$ はその頂点であるといえる. 今,

分布は混合分布ではない，つまり $\alpha=0$ である，という帰無仮説 H_0 と，分布は混合分布モデルである，つまり $\alpha\neq 0$ である，という対立仮説 H_1 を用いた検定問題を考える．$\hat{\alpha}$ と $\hat{\beta}$ をそれぞれ α と β の最尤推定量とすると，H_0 対 H_1 の尤度比検定統計量は

[3] $$T_2 \equiv 2\left[\sum_{i=1}^{n}\log\{(1-\hat{\alpha})\phi(X_i\,|\,0)+\hat{\alpha}\phi(X_i\,|\,\hat{\beta})\}-\sum_{i=1}^{n}\log\phi(X_i\,|\,0)\right]$$

となる．しかし，H_0 の分布 $\phi(x\,|\,0)$ が局所錐モデルの頂点となっているため，H_0 のもとでこの T_2 は漸近的にカイ二乗分布に従わない．

4. 局所錐モデルにおける検定問題

3 章の例 3 以外でも，例えば ARMA モデルの次数の検定や折れ線回帰モデルの次数の検定は，局所錘モデルの頂点か否かの検定である．次に，この検定の分布論に関する二つの理論的結果を記す．

定理 2 . (Dacunha-Castelle and Gassiat 1997). 前述した条件 (A1), (A2) を満たす局所錘モデル $\{f(x\,|\,\alpha,\beta)\,|\,\alpha\geq 0,\beta\in\mathrm{B}\}$ が微分可能性などのいくつかの正則条件を満たすとき，尤度比検定統計量の平方根はある正規確率過程の最大値に (分布) 収束する．

正規確率過程の最大値の分布は Hotelling (1939) と Weyl (1939) により開発されたチューブ法 (例えば Sun 1993) による公式を用いて評価でき，p 値はその公式を用いて評価できる．

例 4 . 混合比未知の混合分布モデル (その 2). 定理 2 より，[3] 式の $H_0:\alpha=0$ 対 $H_1:\alpha\neq 0$ の尤度比検定統計量 T_2 の平方根は $\max_{\beta\in\mathrm{B}}\xi_\beta$ に (分布) 収束する．ただし $\{\xi_\beta\}$ は平均 0，分散 1 の正規過程で，自己相関は

[4] $$r_{\beta\beta'}\equiv \mathrm{E}(\xi_\beta\xi_{\beta'})=\frac{e^{\beta\beta'}-1}{\sqrt{(e^{\beta^2}-1)(e^{\beta'^2}-1)}}$$

である．今，$J_r(\beta)=-\partial^2 r_{\beta\beta'}/\partial\beta^2\,|_{\beta'=\beta}$ とおく．このとき，チューブ法を用いると，$\mathrm{P}(\max_{\beta\in\mathrm{B}}\xi_\beta>t)$ の評価が自己相関 [4] を用いて

[5] $$F(t\,|\,\mathrm{B})\equiv\int_t^\infty\phi(x\,|\,0)\mathrm{d}x+\frac{1}{\sqrt{2\pi}}\phi(t\,|\,0)\int_{\beta\in\mathrm{B}}\sqrt{J_r(\beta)}\mathrm{d}\beta$$

で与えられる. ここで $\phi(x\,|\,0)$ は N(0,1) の確率密度関数である. つまり, T_2 の実現値を t_2 としたとき, p 値の評価は $F(t_2^{1/2}\,|\,\mathrm{B})$ で与えられる.

上の方法において, 自己相関 [4] を導くため, およびその後のチューブ法に適用するためには, 複雑な計算が必要となる. また, 混合する分布を少し変えるだけでも自己相関は [4] 式と異なるものとなり, 新たな計算が必要となる. それに対し, Chen *et al*. (2001) は, 混合分布モデルの検定において尤度比検定を修正し, 問題設定ごとに計算を必要としない検定法を導入した. 彼らは, 最尤推定量の代わりに罰則付対数尤度を最大化する推定量を用いた, 修正尤度比検定を提案した. 3 章の例 3 の問題でいえば, その罰則付対数尤度は例えば,

[6] $$\sum_{i=1}^{n}\log\{(1-\alpha)\phi(X_i\,|\,0)+\alpha\phi(X_i\,|\,\beta)\}+\log\alpha$$

と書くこともできる. ここでは $\log\alpha$ が罰則の項であり, $\alpha=0$ となることを避けている. これを最大化する α と β をそれぞれ $\tilde{\alpha}$ と $\tilde{\beta}$ と記せば, 修正尤度比検定統計量は

[7] $$T_3\equiv 2\left[\sum_{i=1}^{n}\log\{(1-\tilde{\alpha})\phi(X_i\,|\,0)+\tilde{\alpha}\phi(X_i\,|\,\tilde{\beta})\}-\sum_{i=1}^{n}\log\phi(X_i\,|\,0)\right]$$

となる. このとき, この T_3 の漸近分布は以下のように容易に求まる.

定理 3 . (Chen *et al*. 2001). 混合分布モデルが微分可能性などのいくつかの正則条件を満たすとき, H_0 のもと [7] 式の修正尤度比検定統計量 T_3 は漸近的に自由度 1 のカイ二乗分布に従う. つまりその分布関数を $G_1(x)$ と記せば, $\mathrm{P}(T_3 > t)$ は $1-G_1(t)$ に収束する.

定理 3 を 3 章の例 3 に適用すると以下のようになる.

例 5 . 混合比未知の混合分布モデル (その 3). 定理 3 より, $H_0:\alpha=0$ 対 $H_1:\alpha\neq 0$ の修正尤度比検定の p 値は $1-G_1(t_3)$ で与えられる. ここで, t_3 は [7] 式の修正尤度比検定統計量 T_3 の実現値である.

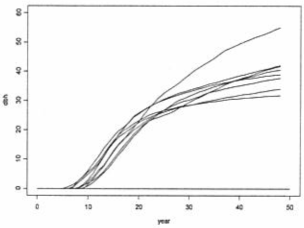
5. 林分成長分析における局所錐モデル

5.1. 林分における構造変化探索

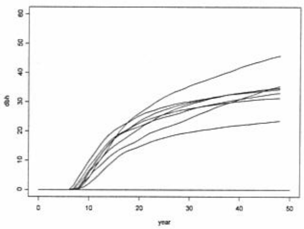
ここでは, 局所錘モデルを林分成長分析に適用し, 実際の解析に応用することを考え, その例を示す. 林分内における各林木の成長は, 例えば隣接木との競合や除間伐, 自然枯死などによる周辺環境の変化の影響を受けている. そして, その成長過程はそうした事象が発生した後, ある時点 (変化点) を経て過去の成長構造とは別のものに変化するものと考えられる. 二宮・吉本 (2004) では, 一本一本の林木に対して構造変化があるか否かを統計的に判断する方法が与えられている. しかしながら, 林分 (複数の林木) 全体の経時成長を把握することも重要であるので, ここでは, 林分に対する構造変化について考える. 林分における構造変化探索では, 周辺環境の変化の影響を受ける林木と受けない林木の二種類が混在するときが問題となる. もし混在しないのであれば, 帰無仮説・対立仮説として構造変化なし・ありのモデルを考え, すべての林木がその独立な標本とすればよい. ここでは, 混在の可能性を考慮し, 混合分布によるモデリングを行うこととする.

以下では, 記録はないものの12年目あるいは18年目くらいに間伐されたと推測される, 福岡県星野村の林分にあるスギ31本の48年分の直径成長データに対し, 12・18・24・30・36年目に構造変化があるか否かの検定をおこなう. 図1は, 個々の林木の成長パターンを見られるよう, 直径成長データを 7 , 8 本ずつまとめて描いたものである. ここで, 二宮・吉本 (2004) で扱った変化検知のための検定を考え, i 番目の林木に対する検定統計量を X_i とする. このとき, X_i は変化がないならば平均 0 , 変化があれば平均 $\beta(\neq 0)$ の正規分布に従うことになる. 変化しているか否かはもちろん未知なので, 各林木が変化しているか否かの確率を導入し, これを α とする. これより, X_i は混合分布の密度関数 $(1-\alpha)\phi(x\mid 0)+\alpha\phi(x\mid\beta)$ をもつとして考えることができる.

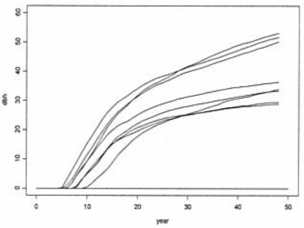
(a) 1番目から8番目の林木



(b) 9番目から16番目の林木



(c) 17番目から24番目の林木



(d) 25番目から31番目の林木

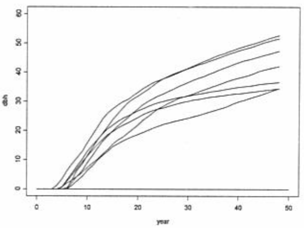


図1. 福岡県星野村の林分にあるスギ31本の48年分の直径成長データ

解析 1. 混合分布モデルを用いない尤度比検定. まず, 混合分布モデルを用いない検定を考えておく. つまり, 2 章の例 1 における $H_0: \beta = 0$ 対 $H_1: \beta \neq 0$ の尤度比検定により変化検知をおこなう. 表1の上段に, β の最尤推定量 $\hat{\beta}$, 尤度比検定統計量 [2] の実現値 t_1, および p 値を示した. 12年目, 18年目の p 値が 5 % を大きく下回ることから, これらの年に変化があったことが示唆される.

表1. k 年目の変化検知問題に対するパラメータの推定値, 検定統計量の実現値, およびp 値

		k=12	k=18	k=24	k=30	k=36
解析 1	$\hat{\beta}$	1.05	1.28	−0.115	0.346	0.242
	t_1	17.0	25.5	0.206	1.85	0.909
	p値	5.54×10^{-9}	9.14×10^{-13}	0.521	0.0544	0.178
解析 2	$\hat{\alpha}$	0.359	0.673	0.395	0.997	0.975
	$\hat{\beta}$	2.69	1.91	−0.195	0.346	0.244
	t_2	27.3	29.5	0.137	1.85	0.893
	p値	8.22×10^{-13}	9.00×10^{-14}	0.757	0.110	0.305
解析 3	$\tilde{\alpha}$	0.375	0.647	0.518	0.655	0.598
	$\tilde{\beta}$	2.66	1.94	−0.170	0.414	0.275
	t_3	27.3	29.5	0.158	1.47	0.622
	p値	1.53×10^{-13}	1.64×10^{-14}	0.574	0.0865	0.265

解析 2．混合分布モデルを用いた尤度比検定. 次に, 3 章の例 3 での $H_0 : \alpha = 0$ 対 $H_1 : \alpha \neq 0$ に関する尤度比検定により変化検知を考え, その統計量の p 値を 4 章の例 4 にしたがって与える方法を考える. 今, B を適当に $[-3, -0.1] \cup [0.1, 3]$ とし, 表1の中段に, α ・ β の最尤推定量 $\hat{\alpha}$ ・ $\hat{\beta}$, 尤度比検定統計量 [3] の実現値 t_2, および p 値を挙げている. 結果から, 12年目の p 値が解析 1 のそれよりずっと小さくなっていることがわかる. これは, この検定が変化を検知するための検定として解析 1 の検定よりすぐれている, つまり検出力がある, ということを意味している. 一方, 18・24・30・36年目の p 値は解析 1 のそれより少し大きくなっている. そこで図2のような x_i の頻度分布を描いて調べたところ12年目の頻度分布が二峰となっているのに対し, それ以外は単峰となっていた. 二つの分布における混合分布は基本的に二峰となるため, 12年目のデータの分布には混合分布が適していることがわかる. つまり, 実際に変化があるものとないものとが混在しているときは混合分布モデルを用いる方がよく, 混在していないときは用いない方がよい, ということがいえる.

解析 3．混合分布モデルを用いた修正尤度比検定. 最後に, 2 章の例 3 での設定で変化検知を考え, p 値を 4 章の例 5 にしたがって与える方法を考える. このとき B を決めたりチューブ法のための計算をしたりせずに p 値を求められる. 表1の下段に, α ・ β の罰則付最尤推定量 $\tilde{\alpha}$ ・ $\tilde{\beta}$, 修正尤度比検定統計量 [7] の実現値 t_3, および p 値を挙げている. 結果から, 解析 2 より容易に p 値を求められるだけでなく, 検出力の高い検定となっていることがわかる.

5.2. その他の分析問題

局所錘モデルは上記以外にも様々な問題に対応できる. 次にその問題例を二つ記す.

問題 1．林木成長の競合・空間相関の検知

ここでは, 林分内の競合がその成長に影響を与えているか否かを統計的にみるために, 競合効果を検知するための検定を構成する. そのためには, 競合効果を取り入れたモデルを構築する必要があるが, 競合の仕方にはさまざまな可能性があり, それを固定することは誤ったモデリングにつながる危険を

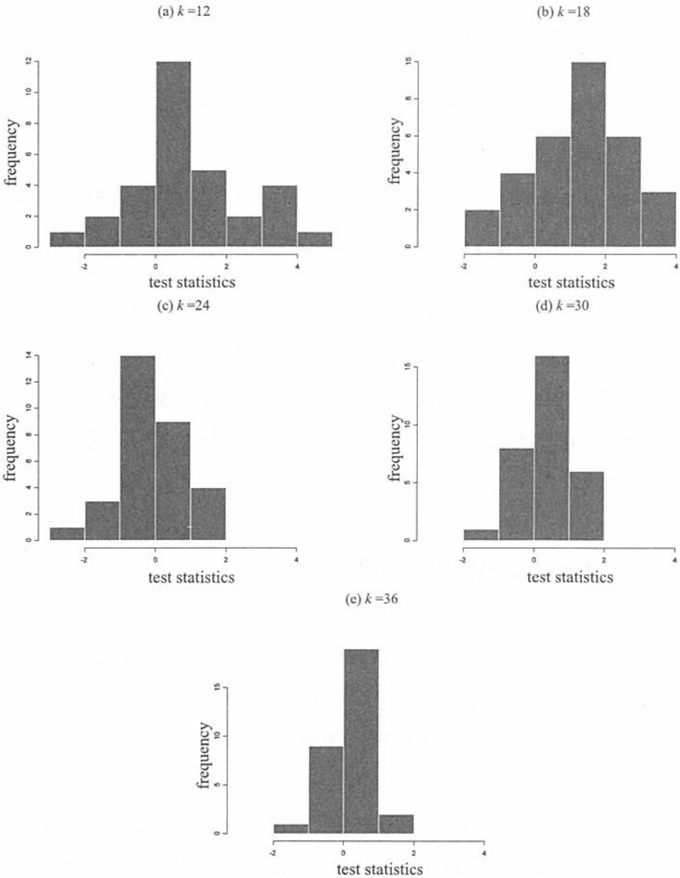


図2. k 年目の変化を検知するための検定統計量 X_i の頻度分布

はらむ. そこで, 競合の形に関するパラメータを導入したモデルを扱うことにする.

この問題の設定としては様々なパターンが考えられるが, 例として各林木の位置と胸高直径のデータが得られているものとし, もっとも大きな影響をもつ競合の要因候補として隣の林木の位置を考えることにする. $D_i\,(i=1,\ldots,n)$ を i 番目の林木の胸高直径, R_i を i 番目の林木とそこから最も近い林木との距離とし, $R_i=r_i$ が与えられたもとでの D_i の (条件付) 分布を $\mathrm{N}(\mu-\alpha\exp(-\beta r_i),\sigma^2)$ とするモデルを考える. なお, $\alpha\exp(-\beta r_i)$ が競合効果の

項であり, 仮に効果がなければ D_i は平均 μ になるものとしている. また, このモデルでは, 距離 r_i が大きいと指数的に競合効果が小さくなり, その程度を $\beta(>0)$ が定めている. つまり, β が大きいと隣接木が影響を及ぼす距離は短く, 小さいと距離は長くなる.

$g(r)$ を R_i の適当な確率密度関数とすると, (D_i, R_i) の同時確率密度関数は

[8] $$f(d,r\,|\,\alpha,\beta,\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}\left(d-\mu+\alpha e^{-\beta r}\right)^2\right\}g(r)$$

となる. ここで $\alpha=0$ とすると, [8] 式は β に依存せず, 局所錘モデルと同様の性質をもつものとなる. すなわち, μ と σ^2 という別のパラメータがさらに存在するため, 局所錘モデルを拡張したものとなる. 競合効果を検知するための検定は $\alpha=0$ 対 $\alpha>0$ で構成され, 局所錘モデル (を拡張したもの) の頂点か否かの検定として対応できることがわかる.

問題 2 . 林木成長パターンの異質性の検知

5.1節や問題 1 では, 成長に影響を与える因子を具体的に想定し, その因子を検知するためのモデリングおよび検定の構成を扱った. ここでは, 因子を想定する前の段階として, そもそも林分内の成長パターンに違いがあるのかといった問題を扱う. 具体的には, 柳原ら (2006) でおこなっている成長パターンのクラスタリングを検定の枠組みで考える.

i 番目の林木の, 時点 t_{ij} の胸高直径を X_{ij} とし $(1\le i\le n,\ 1\le j\le p_i)$, これに Richards (1958) の成長関数 $\exp(\theta_{i1})[1-\exp\{-\exp(\theta_{i2})t_{ij}\}]^{\exp(\theta_{i3})}$ をあてはめる. なお, ここでは曲線型をシグモイド型に制約するために, 従来の Richards の成長関数のパラメータを指数関数により変換している. さて, 推定量 $(\hat\theta_{i1},\hat\theta_{i2},\hat\theta_{i3})$ は成長パターンが一つならば $\mathrm{N}(\boldsymbol{\gamma},\Sigma)$, 成長パターンが二つならば $\mathrm{N}(\boldsymbol{\gamma},\Sigma)$ と $\mathrm{N}(\boldsymbol{\beta},\Sigma)$ が混ざった分布にしたがうとする. ここで, $\boldsymbol{\gamma}=(\gamma_1,\gamma_2,\gamma_3)$, $\boldsymbol{\beta}=(\beta_1,\beta_2,\beta_3)$ であり, Σ は 3 × 3 行列である. 5.1節と同様に考え, $\boldsymbol{x}=(x_1,x_2,x_3)$ でかつ $\mathrm{N}(\boldsymbol{\beta},\Sigma)$ の確率密度関数を $\phi(\boldsymbol{x}\,|\,\boldsymbol{\beta},\Sigma)$ として $(\hat\theta_{i1},\hat\theta_{i2},\hat\theta_{i3})$ は混合分布の密度関数 $(1-\alpha)\phi(\boldsymbol{x}\,|\,\boldsymbol{\gamma},\Sigma)+\alpha\phi(\boldsymbol{x}\,|\,\boldsymbol{\beta},\Sigma)$ をもつとすれば, 成長パターンが一つか否かの検定は $\alpha=0$ 対 $\alpha\ne0$ で与えられる.

6. 結論および考察

林分成長分析において, 局所錘モデルの頂点か否かの検定は様々な場面で適用することができる. また, その検定問題において, p 値を評価するためには通常の理論が適用できないものの, 修正尤度比検定は容易に評価できる. しかしながら局所錘モデルの方向パラメータ (β) が複数となるようなケース, 例えば成長パターンの異質性検知の検定のケースに対しては, 修正尤度比検定の理論を拡張することが必要となってくる.

引用文献

Candy, S. G. 1997. Estimation in forest yield models using composite link functions with random effects. *Biometrics* 53: 146-160

Chen, H., Chen, J. and Kalbfleisch, J. D. 2001. A modified likelihood ratio test for homogeneity finite mixture models. *J. of the Royal Statistical Society, Series B* 63: 19-29

Dacunha-Castelle, D. and Gassiat, E. 1997. Testing in locally conic models and application to mixture models. *ESAIM Probability and Statistics* 1: 285-317

Fang, Z. and Bailey, R. L. 2001. Nonlinear mixed effects modeling for Slash Pine dominant height growth following intensive silvicultural treatments. *Forest Science* 47: 287-300

Feder, P. I. 1975. The log likelihood ratio in segmented regression. *The Annals of Statistics* 3: 84-97

福水健次・栗木　哲. 2004. 特異モデルの統計学. *統計科学のフロンティア* Vol.7: 1-230

Ghosh, J. K. and Sen, P. K. 1985. On the asymptotic performance of the log-likelihood ratio statistic for the mixture model and related results. In *Proc. of the Berkeley conference in honor of Jarzy Neyman and Jack Kiefer* Vol.2. Wadsworth, Belmont, CA. pp.789-806

Hannan, J. 1980. The estimation of the order of an ARMA process. *The Annals of Statistics* 8: 1071-1081

Hartigan, J. A. 1985. A failure of likelihood asymptotics for normal mixtures. In *Proc. of the Berkeley conference in honor of Jarzy Neyman and Jack Kiefer* Vol.2. Wadsworth, Belmont, CA. pp.807-810

Naiman, D. Q. 1986. Conservative confidence bands in curvilinear regression. *The Annals of Statistics* 14: 896-906

二宮嘉行・吉本　敦. 2004. 林木直径成長データを用いた成長傾向に対する変化点探索法の構築. *森林資源管理と数理モデル* Vol.3: 69-80

Richards, F. J. 1958. A flexible growth function to empirical use. *J. of Experimental Botany* 10: 290-300

竹村彰通. 1991. 現代数理統計学. 創文社, 東京. 347p.

Veres, S. 1987. Asymptotic distribution of likelihood ratios for overparametrized ARMA processes. *J. of Time Series Analysis* 8: 345-357

栁原宏和・吉本　敦・二宮嘉行. 2006. 複数の成長パターンを持つスギ単純同齢林における炭素固定量予測. *森林資源管理と数理モデル* Vol.5: 63-84