

有限混合分布のパラメータ推定とその応用

西岡孝彦・三浦幹彦

(信州大学繊維学部繊維管理学研究室)

1. はじめに

確率変数 X が

$$p(x; \phi) = \alpha_1 f_1(x; \theta_1) + \alpha_2 f_2(x; \theta_2) + \cdots + \alpha_k f_k(x; \theta_k) \quad (1.1)$$

但し,

$$\phi^T = (\alpha^T, \theta_1^T, \dots, \theta_k^T)$$

$$\alpha_j > 0, j=1, 2, \dots, k; \sum_{i=1}^k \alpha_i = 1$$

$$f_j(\cdot) \geq 0, \int_{-\infty}^{\infty} f_j(x) dx = 1$$

という密度関数を持つ時、これを有限混合密度関数と呼び、 X は有限混合分布に従うという。この時、重み α_j を混合比、 $f_j(\cdot)$ を成分密度と呼ぶ。パラメータは大きく成分の数を表す k 、混合比 α 、成分密度のパラメータ θ に分けられる。

Peason (1894) は、分散の異なる二つの一変量正規分布の混合分布をかにの甲らの大きさを表すモデルとして適用し、モーメント法によるパラメータ推定法を提案した。これ以来正規分布のパラメータ推定に関しては数多くの研究が発表され (Titterington, Smith and Makov (1985)), 応用分野も広い範囲に渡っている。また、その過程で最尤法がモーメント法より優れていることが Tan and Chang (1972), Fryer and Robertson (1972) などによって指摘された。Hasselblad (1966) は正規分布の混合分布に対して重みつきくり返し法により最尤推定値を求め、この方法を指数分布族まで拡張している。この方法は Dempster ら (1977) により一般的な EM アルゴリズムとして整理され今日に到っている。この間、Day (1967) などにより一変量から多変量正規分布の混合への拡張もなされている。有限混合分布は、最近ではパラメトリックなパターン認識、クラスター分析のモデルとして重要な地位を築きつつある (MacLachlan and Basford (1988))。

しかしながら、連続密度関数を考えた場合、これまでの研究対象は、ほとんどが正規分布を成分密度とするもので、他の分布に関してはあまりとりあげられていない。しかし、後に示すように、応用にあたり、正規分布以外の分布を成分密度として用いるのが妥当な場合がしばしばある。そこで、本論文では X は一変量とし、(1)二つのガンマ分布の混合分布 (2)指数分布と正規分布の混合分布をとりあげ、そのパラメータ推定に関して考察を加える。そのために、まず、有限混合分布のパラメータ推定に関する一般的な結果を要約する。

2. 有限混合分布のパラメータ推定

2.1 最尤法

グラフやモーメント法に基づく初期のころの推定法は、計算機の発達とともに現在ではほとんど使われなくなり、より良い推定値の得られる最小距離法、最尤法及びベイズ推定法へと変っている。距離測度に何を選ぶかによりいろいろな規準の考えられる最小距離法と多次元パラメータ推定問題の特別な場合となるベイズ推定に関する議論については省略し、ここでは最もよく使われる最尤法に関してまとめる。

未知の混合比 α と k 個の成分密度がそれぞれ持っているパラメータ θ をパラメータとする母集団から n 個の標本

$$x_1, x_2, \dots, x_n$$

が得られたとき、これらの標本は母集団のパラメータ α, θ に従っていて、同時に実現した値と考えられる。すなわち、同時確率密度、

$$p(x_1, x_2, \dots, x_n; \phi)$$

但し、
$$\phi^T = (\alpha^T, \theta_1^T, \dots, \theta_k^T)$$

によって n 個の標本の値が実現する確率密度を与えることができる。特に n 個の標本が母集団からの無作為抽出となっている場合、同時確率密度は母集団の密度関数の積、

$$L(\phi; x_1, x_2, \dots, x_n) = p(x_1; \phi) p(x_2; \phi) \cdots p(x_n; \phi) \quad (2.1)$$

で与えられ、これを尤度関数と呼び、Fisher (1921) 以来多くの研究がなされている。実際にパラメータを推定するには、式 (2.1) の取り扱いを容易にするため対数をとった、

$$\log L(\phi; x_1, x_2, \dots, x_n) = \sum_{i=1}^n \log \left\{ \sum_{j=1}^k \alpha_j f_j(x_i; \phi) \right\} \quad (2.2)$$

が用いられ、この式は対数尤度関数と呼ばれる。混合密度の場合は式 (2.2) に示すように混合比を掛けた成分密度の和の対数となることから、解析は容易でないことがうかがえる。

他方、ベイズ流の考え方から最尤法をみると、標本を得た母集団に対していくつかの考えられるパラメータ $(\phi_1, \phi_2, \dots, \phi_t)$ があり、その内の1組のパラメータ ϕ_r が提案されたとする。今、標本抽出により得られた実現 q と、その他の情報 Q 、例えば母集団分布の型など、が与えられたとき ϕ_r を支持する確率は、

$$P(\phi_r | q, Q) = \frac{P(\phi_r | Q) P(q | \phi_r, Q)}{\sum_T P(\phi_r, q | Q)} \quad (2.3)$$

の条件付き確率で与えられる。式 (2.3) において、パラメータ ϕ_r と母集団分布の型などに関する情報 Q が与えられたとき、標本が実現する確率、 $P(q | \phi_r, Q)$ を $L(q | \phi_r,$

Q) におきかえると,

$$p(\phi_r|q, Q)\alpha p(\phi_r|Q)L(q|\phi_r, Q) \quad (2.4)$$

が得られる。もちろん $L(q|\phi_r, Q)$ はもとの確率の意味は失われるが、 $P(q|\phi_r, Q)$ に比例したものとなっている。さて、式 (2.4) の左辺は標本の実現 q と母集団分布の型などの情報としての Q が与えられたときパラメータ ϕ_r がとられる確率であり、いくつか提案されたパラメータ ($\phi_1, \phi_2, \dots, \phi_t$) の中から母集団分布のパラメータに限りなく近い ϕ_r がとられたとき最大の確率を与える。一般にこの型の確率を事後確率とよぶ。しかしながらこの事後確率を直接求めることは困難なため右辺を評価することで左辺の評価に代えることになる。右辺第1項の確率の型は事前確率とよばれ、未知のパラメータそのものの確率であるため、我々には知る事ができない。最近の研究ではこの事前確率に特定の確率分布を仮定し、事後確率を効率良く推定しようとする方向にあるが、従来の最尤法ではベイズの仮説に基きどのパラメータの生起も同等に確からしいという立場で解析を行い、実際この仮説で十分推定可能なことが知られている。そこで評価可能なのは右辺第2項のみとなる。これは式 (2.1) の尤度関数に当るもので、パラメータ ϕ_r と母集団分布の型などの情報 Q が与えられたときの標本の実現 q に対する評価値を与える。いいかえると、 n 個の標本の実現 q は母集団分布のパラメータや母集団分布の型などの情報を反映した最も可能性の高い実現値であり、逆に母集団の真のパラメータをとったとき尤度は最大の値をとることになる。その時、パラメータ ϕ_r を支持するという意味で対数尤度関数 $\log L(\phi_r; x)$ のことをサポート関数と呼ぶこともある。

したがって、最尤法は得られた標本の実現値や母集団分布の型などに関する情報をもとに尤度関数が最大の値をとる様な ϕ を探すことといえる。又、尤度関数を最大にするパラメータは対数尤度関数も最大にすることから、我々は式 (2.2) を最大にするパラメータを求めれば良いことになる。そこで式 (2.2) の各パラメータについて偏微分したものを0とおいた、

$$\frac{\partial \log L(\phi; x)}{\partial \alpha_j} = \sum_{i=1}^n \left\{ \frac{\partial \sum_{j=1}^k \alpha_j f_j(x_i; \phi)}{\partial \alpha_j} \cdot \frac{1}{\sum_{j=1}^k \alpha_j f_j(x_i; \phi)} \right\} = 0$$

$$\frac{\partial \log L(\phi, x)}{\partial \theta_{jl}} = \sum_{i=1}^n \left\{ \frac{\partial \sum_{j=1}^k \alpha_j f_j(x_i; \phi)}{\partial \theta_{jl}} \cdot \frac{1}{\sum_{j=1}^k \alpha_j f_j(x_i; \phi)} \right\} = 0 \quad (2.5)$$

但し、 θ_{jl} は成分密度 $f_j(\cdot)$ のパラメータ θ_j の第 l 要素は最尤方程式と呼ばれ、この解が最尤推定値となる。このとき解が存在するための十分条件は $(\log L(\phi; x))'' < 0$ である。又、式 (2.5) 左辺はスコアと呼ばれる。混合分布の場合、最尤方程式の解析的な解を得ることは非常に困難である。そこで、数値解を得る

幾つかのアルゴリズムが提案されてきた。古典的なものの1つとしてニュートン・ラフソン法がある。これは、スコア $S(\phi)$ をテーラー展開し、その最初の2項を0とおいたときの解から第 ν 回目の反復近似値は、

$$\phi^{(\nu)T} = \phi^{(\nu-1)T} - S^T(\phi^{(\nu-1)}) \left\| \frac{\partial S(\phi^{(\nu-1)})}{\partial \phi_l} \right\|^{-1}$$

$$\text{但し, } \frac{\partial S(\phi^{(\nu-1)})}{\partial \phi_l} = \sum_{i=1}^n \left\{ \frac{\partial^2 \sum_{j=1}^k \alpha_j f_j(x_i; \phi^{(\nu-1)})}{\partial \phi_l \partial \phi_m} \cdot \frac{1}{\left[\sum_{j=1}^k \alpha_j f_j(x_i; \phi^{(\nu-1)}) \right]} \right\}$$

で与えられる。このように、ニュートン・ラフソン法では対数尤度関数の2階微分を必要とするため計算機のパッケージでは1階微分のみを定義すれば2階微分については数値的に得る準ニュートン・ラフソン法が用いられる。

次に、式(2.5)で推定されたパラメータがどの程度の信頼性を持って推定されたか評価できると母集団分布の型などの、モデルの改良に役だてることができる。そこで、パラメータが推定されたときの分散が得られれば一つの評価となるので、分散共分散行列、

$$V(\phi) = -E \left[\sum_{i=1}^n \frac{\partial^2 \sum_{j=1}^k \alpha_j f_j(x_i; \phi)}{\partial \phi_l \partial \phi_m} \cdot \frac{1}{\left\{ \sum_{j=1}^k \alpha_j f_j(x_i; \phi) \right\}^2} \right]^{-1}$$

$$\text{但し, } \phi^T = (\alpha^T, \theta_1^T, \dots, \theta_k^T)$$

を求めれば良い。ここで混合密度の場合、対数尤度関数の2階微分をとると期待値を求めることは非常に困難になるため期待値をとらない分散共分散行列、

$$- \left[\sum_{i=1}^n \frac{\partial^2 \sum_{j=1}^k \alpha_j f_j(x_i; \phi)}{\partial \phi_l \partial \phi_m} \cdot \frac{1}{\left\{ \sum_{j=1}^k \alpha_j f_j(x_i; \phi) \right\}^2} \right]^{-1} \quad (2.6)$$

但し、 $\partial \phi_l$ 、 $\partial \phi_m$ は ϕ の第 l 、第 m 要素。

を用いることができる。分散共分散行列で必要となる対数尤度関数の2階微分はニュートン・ラフソン法を行う際にも必要で、最尤方程式を解く反復の過程で自然に分散共分散行列が得られるが、その他のアルゴリズム、例えば最急降下法、ガウス・ニュートン法などを採用すると新たため式(2.6)を評価しなければならない。

2.2 SAS や BMDP の利用

最尤推定値は非線形回帰を利用して求めることができる。SAS の NLIN プロシージャや BMDP の P3R などがそれに当たり、ともに、

$$s^2 = \frac{1}{n-p} \sum_{i=1}^n (y_i - f(x_i; \phi))^2 \quad (2.7)$$

但し、 p はパラメータの数
を最小にするような ϕ を求めるプログラムである。最尤法は $\log L(\phi; x) = \sum_{i=1}^n \log \left[\sum_{j=1}^k \alpha_j f_j(x_i; \phi) \right]$ を最大にする方法であるから、適当な c に対して、

$$\sum_{i=1}^n \left\{ c - \log \left[\sum_{j=1}^k \alpha_j f_j(x_i; \phi) \right] \right\} = \sum_{i=1}^n \left\{ \sqrt{c - \log \left[\sum_{j=1}^k \alpha_j f_j(x_i; \phi) \right]} \right\}^2 \quad (2.8)$$

を最小にすることと同じである。ここで c は平方根内が正値となる様な定数を与えれば良い。したがって、式 (2.7) で、 $y_i = 0$ とし $f(x_i; \phi)$ を式 (2.8) 左辺大カッコ内の式でおきかえると非線形最小二乗を行うことに等しくなる。又、最小二乗回帰は、

$$\sum_{i=1}^n \frac{\partial f(x_i; \phi)}{\partial \phi_l} [y_i - f(x_i; \phi)] = 0$$

ここで ϕ_l はパラメータ ϕ の第 l 要素
を解くことであるから、 $f(x_i; \phi)$ を対数尤度関数に置きかえて非線形回帰問題と考えることができる。この2つの統計解析パッケージともに、対数尤度関数と最尤方程式を与えればパラメータの推定値と漸近標準誤差が得られるが、この場合の残差解析には意味がないので各パラメータの推定時の分散を知るには式 (2.6) を求める必要がある。又、SAS の *NLIN* プロシージャでは解法アルゴリズムのオプション指定ができ、最急降下法、ガウス・ニュートン法、マーカート法、分割法のうちの1つを選ぶことができる。

2.3 EM アルゴリズムによる計算

不完全データ問題を基本として *Dempster* らによって整理された EM アルゴリズムは、簡単に述べれば、EステップとMステップの二つから成っている。欠測値も含めた完全データの対数尤度 $l(\phi)$ を考え、パラメータの初期値 $\phi^{(0)}$ と観測値 $y^T = (y_1, y_2, \dots, y_n)$ が与えられたという条件の下での対数尤度関数の条件付期待値

$$E[l(\phi) | \phi^{(0)}, y]$$

を求めるのがEステップである。

次にこの期待値を最大にする解を求めるMステップを行い、その解を一回目の値 $\phi^{(1)}$ とする。このEとMのステップを ϕ の変化が小さくなるまでくり返していけば、最尤推定値 $\hat{\phi}$ が得られ、しかも、このくり返しにより得られる ϕ の値は

$$\log L(\phi^{(k+1)}) \geq \log L(\phi^{(k)})$$

という望ましい性質を備えている。ただし、 $\log L(\phi^{(k)})$ は観測可能なデータだけから作られる対数尤度関数に k 回のくり返しにより得られた $\phi^{(k)}$ を代入したものを示す。

観測データ $y^T = (y_1, y_2, \dots, y_n)$ は不完全データで、完全データ $x_1^T = (y_1, z_1^T)$, $x_2^T = (y_2, z_2^T)$, \dots , $x_n^T = (y_n, z_n^T)$ のうちのどの成分から抽出されたものかを示すデー

た, z_1, z_2, \dots, z_n が欠測しているものとする。ここで, z_1, z_2, \dots, z_n は互いに独立に多項分布に従うものとする。この時, 完全データの同時分布は

$$\prod_{i=1}^n \prod_{j=1}^k \alpha_j^{z_{ij}} f_j(y_i; \theta_j)$$

となり, 従って対数尤度関数は

$$l(\phi) = \sum_{i=1}^n \sum_{j=1}^k z_{ij} \{ \log \alpha_j + \log f_j(y_i; \theta_j) \} \quad (2.9)$$

で与えられる。 $\phi^{(0)}$ と y が与えられた時の $l(\phi)$ の条件付期待値は (2.9) 式において z_{ij} をその条件付期待値

$$E[z_{ij} | \phi^{(0)}, y] = w_j(y_i, \phi^{(0)}) = \frac{\alpha_j^{(0)} f_j(y_i; \phi^{(0)})}{f(y_i, \phi^{(0)})}$$

すなわち, $\phi^{(0)}$, y_i が与えられた時の, y_i が j 番目の成分に属する事後確率 $w_j(y_i, \phi^{(0)})$ で置き換えたものである。これが E ステップとなる。ここで, 混合比 α と成分密度のパラメータ θ が独立であれば, M ステップにより

$$\frac{\partial l(\phi)}{\partial \alpha} = 0$$

から

$$\alpha_j^{(1)} = \frac{1}{n} \sum_{i=1}^n w_j(y_i, \phi^{(0)})$$

となり, θ については

$$\frac{\partial}{\partial \theta} \sum_{i=1}^n \sum_{j=1}^k w_j(y_i, \phi^{(0)}) \log f_j(y_i; \theta_j) = 0$$

の解を求めることになる。 α と θ が独立でない場合については複雑となるが, 三浦ら (1988) にその例がある。

この EM アルゴリズムを用いる時の問題点は, ニュートン・ラフソン法では情報行列の計算が直接的であるのに対し, 実際の尤度関数とは異なる完全尤度関数を用いているため直接計算できないことであった。この問題は Louis (1982) により解決され, 観測値データによる情報行列の計算を完全尤度関数を用いる EM アルゴリズムの中に組み込むことができるようになった。

$S(x, \phi)$, $S^*(y, \phi)$ を対数完全尤度関数及び対数尤度関数をそれぞれ ϕ で一回微分したものとする。また $B(x, \phi)$, $B^*(y, \phi)$ を ϕ で二回微分した行列の符号を変えたものとする。この時, 不完全データ y の情報行列 $I_Y(\phi)$ は

$$I_Y(\phi) = E_\phi[B(X, \phi)|X \in R] - E_\phi[S(X, \phi)S^T(X, \phi)|X \in R] \\ + S^*(y, \phi)S^{*T}(y, \phi)$$

となり ϕ を最尤推定値 $\hat{\phi}$ で置き換えれば、

$$I_Y(\hat{\phi}) = E_{\hat{\phi}}[B(X, \hat{\phi})|X \in R] - E_{\hat{\phi}}[S(X, \hat{\phi})S^T(X, \hat{\phi})|X \in R]$$

で与えられる。但し、 $R = \{x: y(x) = y\}$ である。ここでデータが互いに独立であれば、観測値に基づく情報行列（期待値をとらないもの）は

$$I_Y(\hat{\phi}) = \sum_{i=1}^n E_{\hat{\phi}}[B_i(X_i, \hat{\phi})|X_i \in R_i] - \sum_{i=1}^n E_{\hat{\phi}}[S_i(X_i, \hat{\phi})S_i^T(X_i, \hat{\phi})|X_i \in R_i] \\ + \sum_{i=1}^n E_{\hat{\phi}}[S_i(X_i, \hat{\phi})|X_i \in R_i]E_{\hat{\phi}}[S_i^T(X_i, \hat{\phi})|X_i \in R_i] \quad (2.10)$$

で与えられる。

3. いろいろな成分密度関数の混合分布

3.1 二つのガンマ分布の混合分布

正規分布の混合分布に関しては、推定法及び推定されたパラメータの精度に関して多くの研究がなされているが、ガンマ分布を成分密度とする混合分布については、ほとんど研究されていない。そこで、ここでは2節で述べた一般的結果を利用して混合密度

$$p(x; \alpha, \xi_1, \gamma_1, \xi_2, \gamma_2) = (1-\alpha)f_1(x; \xi_1, \gamma_1) + \alpha f_2(x; \xi_2, \gamma_2) \quad (3.1)$$

但し、成分密度はガンマ分布

$$f_j(x; \xi_j, \gamma_j) = \frac{\xi_j^{\gamma_j}}{\Gamma(\gamma_j)} (\xi_j x)^{\gamma_j - 1} e^{-\xi_j x}, \quad x > 0$$

で表わされる分布のパラメータ推定及び推定パラメータの精度に関する考察を加えた。図1に(3.1)式のパラメータの値を変化させた時に得られる典型的な密度関数の形を示した。(a)は二つの峰をもつ形、(b)は最初に指数的に減少しその後峰ができる形、(c)は大きな峰ができた後に極端に長く右に尾を引く形となっている。従って、二つのガンマ分布の混合分布はこれらに代表されるような変動を示すデータのモデルとして使える可能性を示唆している。この場合のEMアルゴリズムは次のようになる。

観測可能なデータを y_1, y_2, \dots, y_n とすれば、完全データの対数尤度関数は(2.9)式から

$$l(\phi) = \sum_{i=1}^n \sum_{j=1}^2 z_{ij} \{ \log \alpha_j + \gamma_j \log \xi_j - \log \Gamma(\gamma_j) \\ + (\gamma_j - 1) \log y_i - \xi_j y_i \} \quad (3.2)$$

となる。但し $\alpha_1 = 1 - \alpha$ とする。従って、この z_{ij} を事後確率 $w_j(y_i, \phi^{(m)})$ で置き換えたものを最大にするように $\alpha_j, \xi_j, \gamma_j$ を求めるには、

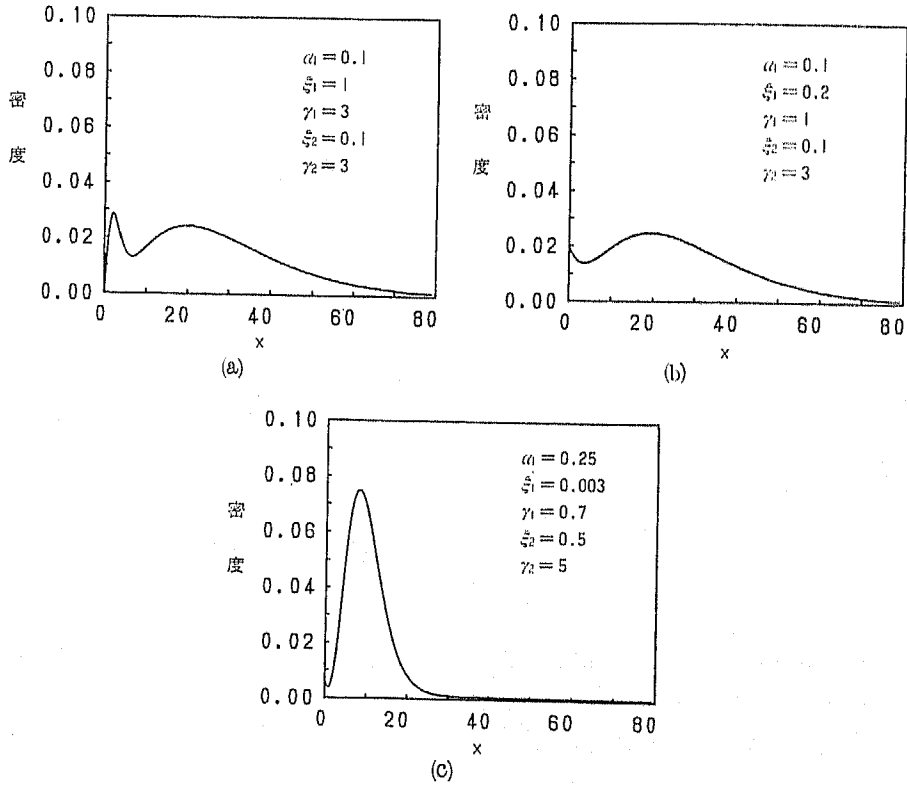


図1 二つのガンマ分布の混合分布

$$\alpha_j^{(m+1)} = \frac{1}{n} \sum_{i=1}^n w_j(y_i, \phi^{(m)}) \quad j=1, 2$$

により α_j を求め

$$\Psi(\gamma_j^{(m+1)}) - \log \xi_j^{(m+1)} = \frac{\sum_{i=1}^n w_j(y_i, \phi^{(m)}) \log y_i}{\sum_{i=1}^n w_j(y_i, \phi^{(m)})}$$

$$\frac{\gamma_j^{(m+1)}}{\xi_j^{(m+1)}} = \frac{\sum_{i=1}^n w_j(y_i, \phi^{(m)}) y_i}{\sum_{i=1}^n w_j(y_i, \phi^{(m)})}$$

を解き $\xi_j^{(m+1)}$, $\gamma_j^{(m+1)}$ を求めればよい。ただし、 $\Psi(\cdot)$ はディガンマ関数を表す。このステップを収束するまでくり返し、得られた値を最尤推定値 $\hat{\phi}$ とする。

また、情報行列を求めるために (2.10) 式の各項を計算すれば、

$$I_1 = \sum_{i=1}^n E_{\hat{\phi}}[B_i(X_i, \hat{\phi}) | X_i \in R_i]$$

$$= \begin{bmatrix} \frac{n}{\hat{\alpha}_1} + \frac{n}{\hat{\alpha}_2} & O \\ O & A_1 \\ & & A_2 \end{bmatrix}$$

但し,

$$A_j = \begin{bmatrix} \frac{n\hat{\alpha}_j \hat{\gamma}_j}{\hat{\xi}_j^2} & -\frac{n\hat{\alpha}_j}{\hat{\xi}_j} \\ -\frac{n\hat{\alpha}_j}{\hat{\xi}_j} & n\alpha_j \Psi'(\gamma_j) \end{bmatrix}$$

$$I_2 = \sum_{i=1}^n E_{\hat{\phi}}[S_i(X_i, \hat{\phi}) S_i^T(X_i, \hat{\phi}) | X_i \in R_i]$$

$$= \begin{bmatrix} \frac{n}{\hat{\alpha}_1} + \frac{n}{\hat{\alpha}_2} & O \\ O & D_1 \\ & & D_2 \end{bmatrix}$$

$$D_j = \begin{bmatrix} \sum_{i=1}^n w_j(y_i, \hat{\phi}) \left(\frac{\hat{\gamma}_j}{\hat{\xi}_j} - y_i \right)^2 & \sum_{i=1}^n w_j(y_i, \hat{\phi}) \left(\frac{\hat{\gamma}_j}{\hat{\xi}_j} - y_i \right) (c_j + \log y_i) \\ \sum_{i=1}^n w_j(y_i, \hat{\phi}) \left(\frac{\hat{\gamma}_j}{\hat{\xi}_j} - y_i \right) (c_j + \log y_i) & \sum_{i=1}^n w_j(y_i, \hat{\phi}) (c_j + \log y_i)^2 \end{bmatrix}$$

但し, $C_j = \log \hat{\xi}_j - \Psi(\gamma_j)$

$$I_3 = \sum_{i=1}^n E_{\hat{\phi}}[S_i(X_i, \hat{\phi}) | X_i \in R_i] E_{\hat{\phi}}[S_i^T(X_i, \hat{\phi}) | X_i \in R_i] = \sum_{i=1}^n d_i d_i^T$$

但し,

$$d_i = \begin{bmatrix} \frac{w_1(y_i, \hat{\phi})}{\hat{\alpha}_1} - \frac{w_2(y_i, \hat{\phi})}{\hat{\alpha}_2} \\ w_1(y_i, \hat{\phi}) \left(\frac{\hat{\gamma}_1}{\hat{\xi}_1} - y_i \right) \\ w_1(y_i, \hat{\phi}) (c_1 + \log y_i) \\ w_2(y_i, \hat{\phi}) \left(\frac{\hat{\gamma}_2}{\hat{\xi}_2} - y_i \right) \\ w_2(y_i, \hat{\phi}) (c_2 + \log y_i) \end{bmatrix}$$

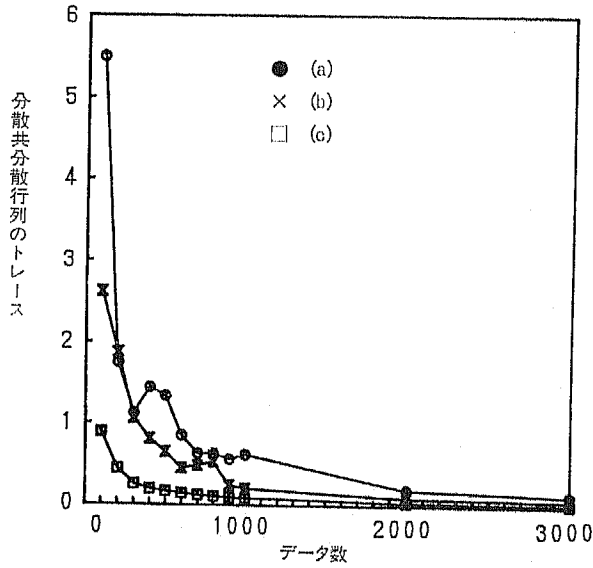


図2 データ数と全分散との関係

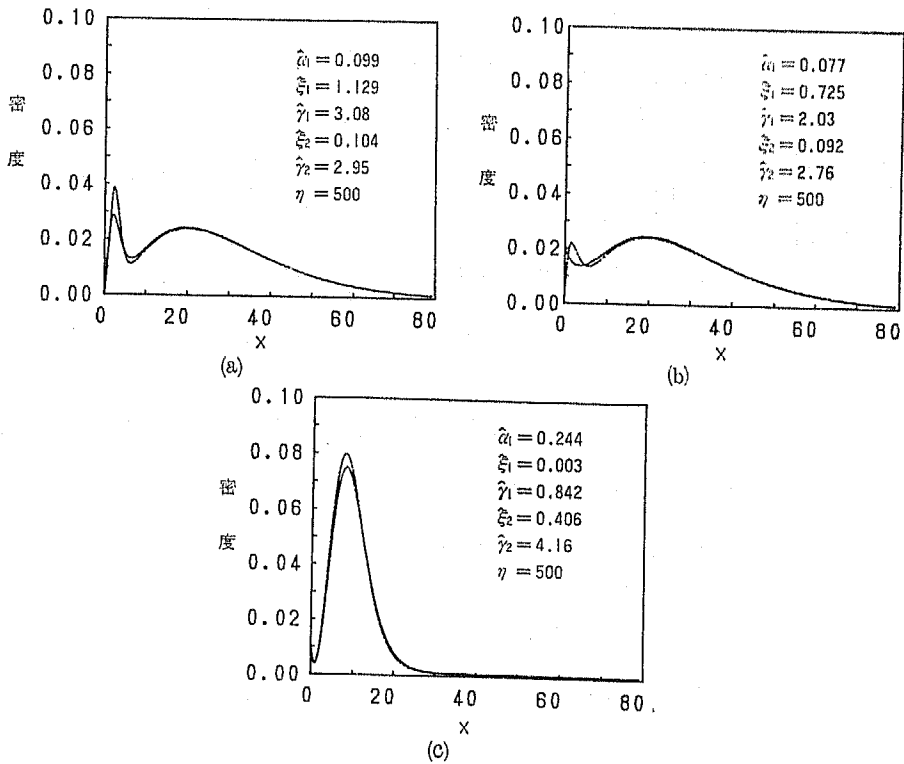


図3 母集団密度と推定密度の比較

— : 母集団密度 : 推定密度

となり、最尤推定値 $\hat{\phi}$ の観測値に基づく分散共分散行列は、情報行列 $I_Y(\hat{\phi}) = I_1 - I_2 + I_3$ の逆行列として得られる。

図1に示した三種類の分布に対してシミュレーションによりデータ数を変化させた時の観測値に基づく分散共分散行列の変化を調査した。図2はその結果を表し、縦軸は全分散 ($I_Y^{-1}(\hat{\phi})$ のトレース) を示している。この調査から、データ数が100~500程度の時は(c)型の分布の推定精度が(a), (b)型に比べてかなり良いことが知られた。これを確かめるために500個のデータ(シミュレーションによる)から推定したパラメータの値を持つ密度関数のグラフと比較して図3に示した。(c)では母集団分布とデータから検定した密度関数は、かなり近似しているが、(a), (b)では混合比が0.1と小さいため x の小さい所での不一致が目立ち推定が困難となっている。特に(b)では、型パラメータが1を境にして分布が大きく変化する影響をうけて困難さが増していることが知られた。これをさらに検討するために、同じデータに対してプロファイル対数尤度

$$\sup_{\theta} \log L(\alpha_1)$$

を計算しその結果を図4に示した。(a), (b)の場合は平坦な変化を示しパラメータ推定が困難であることがわかる。しかし、(c)の場合は(a), (b)に比べて大きな変化を示し、比較的推定が容易であることがうかがえる。

復帰糸長データへの応用:

蚕が繭を作る時、繭の中央を離れてから再び繭の中央に復帰するまでに吐く糸長を測

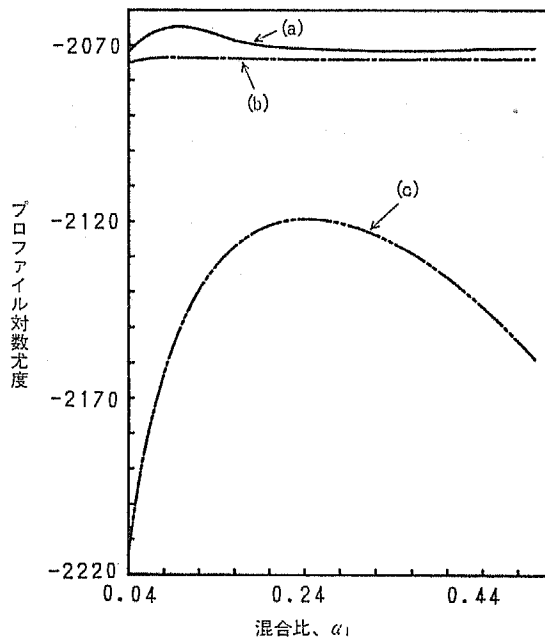


図4 プロファイル対数尤度関数

定して作成したのが図5である。このデータの特徴は 0～25mm の付近に大半が集中し、しかもかなり長い糸長のデータが無視できない頻度で出現し、図1(o)と似た様子を示すことであった。そこで、このデータの分布モデルとして二つのガンマ分布を成分密度とする混合分布を考え、最尤法によってパラメータの推定を行った。EM アルゴリズムによって得られた推定値を図5に、推定値の漸近的分散共分散行列を表1に示した。また、この推定値を用いて理論確率を計算し図中に示した。0～5mm の部分で実測相対度数と理論確率との間に違いがあるが、このデータの母集団モデルとして二つのガンマ分布の混合分布が十分利用できることを示している。

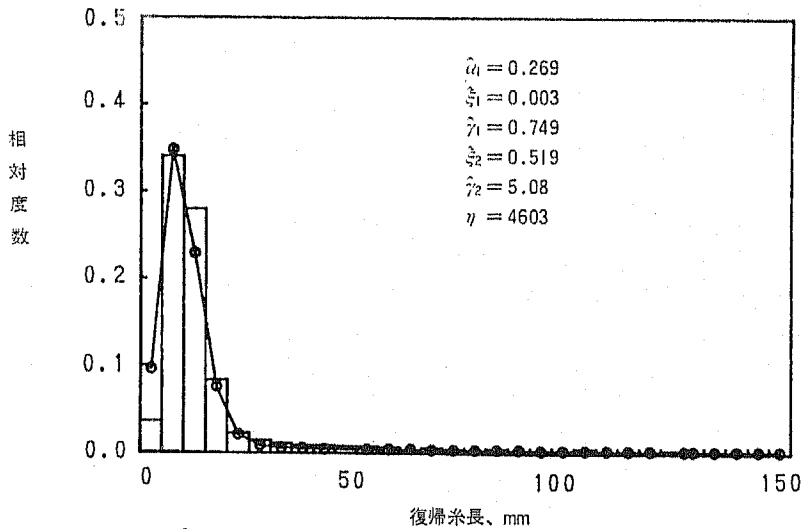


図5 復帰糸長の実測相対度数と推定確率との比較

表1 漸近分散共分散行列

	α_1	ξ_1	γ_1	ξ_2	γ_2
α_1	0.63955×10^{-4}	-0.14761×10^{-6}	-0.87218×10^{-4}	0.36179×10^{-4}	0.30596×10^{-3}
ξ_1		0.29368×10^{-7}	0.39144×10^{-5}	-0.17055×10^{-6}	-0.18712×10^{-5}
γ_1			0.10733×10^{-2}	-0.12964×10^{-3}	-0.11929×10^{-2}
ξ_2				0.27997×10^{-3}	0.24145×10^{-2}
γ_2					0.22411×10^{-1}

3.2 指数分布と正規分布あるいはガンマ分布、指数分布と正規分布の混合分布

これらの成分密度の組合せによる分布は、一般に寿命に関するものに多く見られる。物理的な説明として、指数分布、ガンマ分布で初期故障するまでの時間を、又摩擦により寿命が尽きるまでの時間を正規分布であてはめていて、指数分布と正規分布による混合密度は、

$$p_{EN}(x; \alpha, \lambda, \mu, \sigma) = (1-\alpha)f_1(x; \lambda) + \alpha f_2(x; \mu, \sigma) \quad (3.3)$$

但し, $f_1(x; \lambda) = \lambda e^{-\lambda x}$

$$f_2(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

で与えられる。ここで $\mu \gg a$ なら、この混合密度関数は2つの峰をもち、標本平均値 $\frac{1}{n} \sum_{i=1}^n x_i$ はこれらの峰の間に位置し、平均値から得られる情報は稀薄なものとなる。そのようすを図6に示す。しかし式(3.3)の4つのパラメータが推定できれば多くの情報が得られることになる。例えば初期故障率 $1-\alpha$ 、初期故障を起すまでの平均時間 $1/\lambda$ 、摩耗により寿命が尽きるまでの平均値 μ とその分散 σ^2 などである。

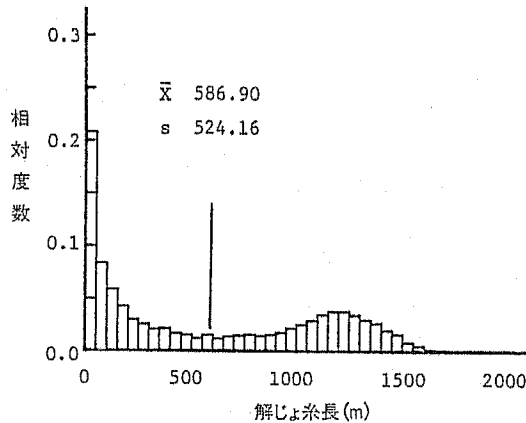


図6 解じよ糸長分布

最尤法でパラメータを推定するため、まず式(3.3)をもとに対数尤度関数、

$$\log L_{EN}(\phi; x) = \sum_{i=1}^n \log[(1-\alpha)f_1(x_i; \lambda) + \alpha f_2(x_i; \mu, \sigma)] \quad (3.4)$$

n はデータ数

を得る。次に、この対数尤度関数からパラメータごとに偏微分したものを0とおいた最尤方程式、

$$\frac{\partial \log L_{EN}(\phi; x)}{\partial \alpha} = \sum_{i=1}^n \frac{1}{p_{EN}(x_i; \phi)} [-f_1(x_i; \lambda) + f_2(x_i; \mu, \sigma)] = 0$$

$$\frac{\partial \log L_{EN}(\phi; x)}{\partial \lambda} = (1-\alpha) \sum_{i=1}^n \frac{1}{p_{EN}(x_i; \phi)} \left(\frac{1}{\lambda} - x_i \right) f_1(x_i; \lambda) = 0$$

$$\begin{aligned}\frac{\partial \log L_{EN}(\phi; x)}{\partial \mu} &= \alpha \sum_{i=1}^n \frac{1}{p_{EN}(x_i; \phi)} \left(\frac{x_i - \mu}{\sigma^2} \right) f_2(x_i; \mu, \sigma) = 0 \\ \frac{\partial \log L_{EN}(\phi; x)}{\partial \sigma} &= \alpha \sum_{i=1}^n \frac{1}{p_{EN}(x_i; \phi)} \frac{1}{\alpha} \left[-1 + \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_2(x_i; \mu, \sigma) = 0\end{aligned}$$

$$\text{但し, } p_{EN}(x_i; \phi) = (1 - \alpha) f_1(x_i; \lambda) + \alpha f_2(x_i; \mu, \sigma)$$

を得る。この解が最尤推定値となるが、ニュートン・ラフソン法で解を得る場合には先に示したように対数尤度関数の2階微分まで必要とする。その他のアルゴリズムを用いた場合にもパラメータの推定時の分散共分散行列を得るには2階微分が必要となるのでパラメータごとの偏2階微分を求めておく。その各要素は、

$$\begin{aligned}\frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \alpha^2} &= - \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} [-f_1(x_i; \lambda) + f_2(x_i; \mu, \sigma)]^2 \\ \frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \alpha \partial \lambda} &= - \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \left(\frac{1}{\lambda} - x_i \right) f_1(x_i; \lambda) f_2(x_i; \mu, \sigma) \\ \frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \alpha \partial \mu} &= \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \left(\frac{x_i - \mu}{\sigma^2} \right) f_1(x_i; \lambda) f_2(x_i; \mu, \sigma) \\ \frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \alpha \partial \sigma} &= - \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \left(\frac{1}{\sigma} \right) \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_1(x_i; \lambda) f_2(x_i; \mu, \sigma) \\ \frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \lambda^2} &= - (1 - \alpha) \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \left\{ \alpha \left(\frac{2}{\lambda} - x_i \right) x_i f_2(x_i; \mu, \sigma) \right. \\ &\quad \left. + (1 - \alpha) \frac{1}{\lambda^2} f_1(x_i; \lambda) \right\} f_1(x_i; \lambda) \\ \frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \lambda \partial \mu} &= - \alpha (1 - \alpha) \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \left(\frac{1}{\lambda} - x_i \right) f_1(x_i; \lambda) \\ &\quad \left(\frac{x_i - \mu}{\sigma^2} \right) f_2(x_i; \mu, \sigma) \\ \frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \lambda \partial \sigma} &= \alpha (1 - \alpha) \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \left(\frac{1}{\lambda} - x_i \right) f_1(x_i; \lambda) \\ &\quad \frac{1}{\sigma} \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_2(x_i; \mu, \sigma) \\ \frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \mu^2} &= - \alpha \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \frac{1}{\sigma^2} \left\{ (1 - \alpha) \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_1(x_i; \lambda) \right. \\ &\quad \left. + \alpha f_2(x_i; \mu, \sigma) \right\} f_2(x_i; \mu, \sigma) \\ \frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \mu \partial \alpha} &= - \alpha \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \frac{1}{\sigma^2} \left\{ (1 - \alpha) \left[2 \frac{x_i - \mu}{\sigma} \right. \right. \\ &\quad \left. \left. - \left(\frac{x_i - \mu}{\sigma} \right)^3 \right] f_1(x_i; \lambda) + \alpha \frac{x_i - \mu}{\sigma} f_2(x_i; \mu, \sigma) \right\} f_2(x_i; \mu, \sigma)\end{aligned}$$

$$\frac{\partial^2 \log L_{EN}(\phi; x)}{\partial \sigma^2} = \alpha \sum_{i=1}^n \frac{1}{[p_{EN}(x_i; \phi)]^2} \frac{1}{\sigma^2} \left\{ (1-\alpha) \left[2 - 5 \left(\frac{x_i - \mu}{\sigma} \right)^2 + \left(\frac{x_i - \mu}{\sigma} \right)^4 \right] f_1(x_i; \lambda) + \alpha \left[1 - 3 \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_2(x_i; \mu, \sigma) \right\} f_2(x_i; \mu, \sigma)$$

但し, $p_{EN}(x_i; \phi) = (1-\alpha)f_1(x_i; \lambda) + \alpha f_2(x_i; \mu, \sigma)$
 $f_1(x_i; \lambda) = \lambda e^{-\lambda x_i}$

$$f_2(x_i; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}$$

となる。以上の式を要素とする行列の逆行列の負が分散共分散行列となる。

最尤方程式の解は SAS の NLIN プロシージャを用い、解法アルゴリズムはガウス・ニュートン法で行った。したがって、対数尤度関数と最尤方程式を SAS 言語で記述す

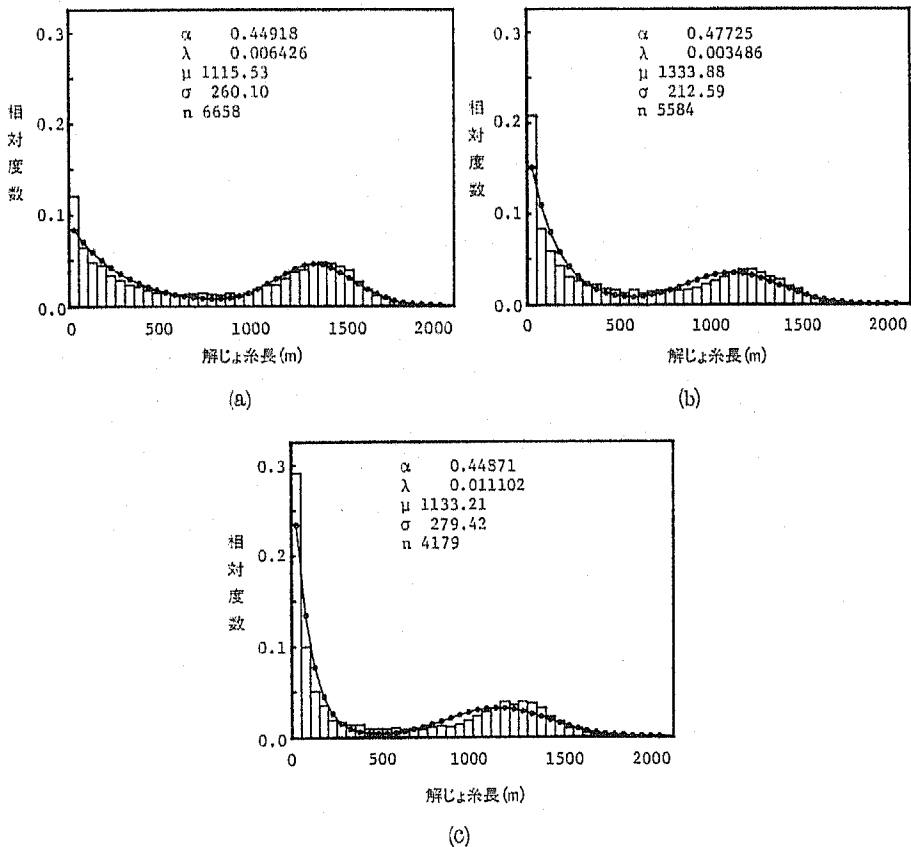


図7 指数分布と正規分布による解じょ糸長へのあてはめ

れば最尤推定値が得られるが、分散共分散行列については別に求める必要がある。

データ解析は昭和54年産春蘭（春月×宝鐘）、初秋蘭（錦秋×鐘和）、晩秋蘭（錦秋×鐘和）のそれぞれの解じょ糸長の調査したものについて行った。図7(a), (b), (c)にそれぞれの解じょ糸長の相対度数と推定されたパラメータによって得られた混合密度による出現確率を示す。3組のデータとも推定された出現確率の0 mから50mの区間の相対度数への適合があまりよくなく、又正規分布をあてはめた摩耗による寿命の部分のモードが解じょ糸長の短い側に移っていることがわかる。すなわち、初期故障の平均値は過大推定され、摩耗寿命の平均値は過小推定されたことになる。しかしながら指数分布と正規分布の混合密度というあらかじめ与えられた情報のもとでは妥当なパラメータが推定されたものと考えられる。分散共分散行列を表2(a), (b), (c)に示す。

表2(a) 春蘭，指数分布と正規分布の漸近分散共分散行列

	α	λ	μ	a
α	0.702830×10^{-4}	0.429499×10^{-6}	-0.173954×10^{-1}	0.164491×10^{-1}
λ		0.116257×10^{-7}	-0.269482×10^{-3}	0.256299×10^{-3}
μ			0.308252×10^0	-0.129737×10^2
a				0.216453×10^2

表2(b) 初秋蘭，指数分布と正規分布の漸近分散共分散行列

	α	λ	μ	a
α	0.610716×10^{-4}	0.802484×10^{-6}	-0.272348×10^{-1}	0.230990×10^{-1}
λ		0.389208×10^{-7}	-0.893796×10^{-3}	0.744168×10^{-3}
μ			0.542915×10^2	-0.272766×10^2
a				0.352843×10^2

表2(c) 晩秋蘭，指数分布と正規分布の漸近分散共分散行列

	α	λ	μ	a
α	0.717699×10^{-4}	0.101478×10^{-5}	-0.190826×10^{-1}	0.199224×10^{-1}
λ		0.136999×10^{-6}	-0.152706×10^{-2}	0.157442×10^{-2}
μ			0.707955×10^2	-0.306312×10^2
a				0.532846×10^2

指数分布と正規分布による混合密度によるモデルにはさらに改良すべき点がある。そこで解じょ糸長の最も短い階級の相対度数をおぎなうようにガンマ分布を加え3つの異なる分布から成る混合密度をモデルとする。この場合の混合密度は、

$$p_{GEN}(x; \alpha, \beta, \xi, \gamma, \lambda, \mu, \sigma) = (1 - \alpha - \beta) f_1(x; \xi, \gamma) + \alpha f_2(x; \lambda) + \beta f_3(x; \mu, \sigma)$$

$$\text{但し, } f_1(x; \xi, \gamma) = \frac{\xi}{\Gamma(\gamma)} (\xi x)^{\gamma-1} e^{-\xi x}$$

$$f_2(x; \lambda) = \lambda e^{-\lambda x}$$

$$f_3(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

となり, 7つのパラメータをもつ。このときの対数尤度関数は,

$$\log L_{GEN}(\phi; x) = \sum_{i=1}^n \log \{ (1-\alpha-\beta) f_1(x_i; \xi, \gamma) + \alpha f_2(x_i; \lambda) + \beta f_3(x_i; \mu, \sigma) \}$$

で与えられる。

次に, 最尤方程式を求めると,

$$\frac{\partial \log L_{GEN}(\phi; x)}{\partial \alpha} = \sum_{i=1}^n \frac{1}{p_{GEN}(x_i; \phi)} [-f_1(x_i; \xi, \gamma) + f_2(x_i; \lambda)] = 0$$

$$\frac{\partial \log L_{GEN}(\phi; x)}{\partial \beta} = \sum_{i=1}^n \frac{1}{p_{GEN}(x_i; \phi)} [-f_1(x_i; \xi, \gamma) + f_3(x_i; \mu, \sigma)] = 0$$

$$\frac{\partial \log L_{GEN}(\phi; x)}{\partial \xi} = (1-\alpha-\beta) \sum_{i=1}^n \frac{1}{p_{GEN}(x_i; \phi)} \left(\frac{\gamma}{\xi} - x_i \right) f_1(x_i; \xi, \gamma) = 0$$

$$\frac{\partial \log L_{GEN}(\phi; x)}{\partial \gamma} = (1-\alpha-\beta) \sum_{i=1}^n \frac{1}{p_{GEN}(x_i; \phi)} [\log(\xi x_i) - \Psi(\gamma)] f_1(x_i; \xi, \gamma) = 0$$

$$\frac{\partial \log L_{GEN}(\phi; x)}{\partial \lambda} = \alpha \sum_{i=1}^n \frac{1}{p_{GEN}(x_i; \phi)} \left(\frac{1}{\lambda} - x_i \right) f_2(x_i; \lambda) = 0$$

$$\frac{\partial \log L_{GEN}(\phi; x)}{\partial \mu} = \beta \sum_{i=1}^n \frac{1}{p_{GEN}(x_i; \phi)} \left(\frac{x_i - \mu}{\sigma^2} \right) f_3(x_i; \mu, \sigma) = 0$$

$$\frac{\partial \log L_{GEN}(\phi; x)}{\partial \sigma} = -\beta \sum_{i=1}^n \frac{1}{p_{GEN}(x_i; \phi)} \frac{1}{\sigma} \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_3(x_i; \mu, \sigma) = 0$$

$$\text{但し, } \Psi(\gamma) = \frac{\Gamma'(\gamma)}{\Gamma(\gamma)}$$

$$\Gamma'(\gamma) = \frac{d\Gamma(\gamma)}{d\gamma}$$

$$\Gamma(\gamma) = \int_0^{\infty} t^{\gamma-1} e^{-t} dt$$

を得る。指数分布と正規分布のモデルの場合と同様に分散共分散行列を求めるため各パラメータについて, 偏2階微分を求めると,

$$\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \alpha^2} = - \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} [-f_1(x_i; \xi, \gamma) + f_2(x_i; \lambda)]^2$$

$$\begin{aligned}
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \alpha \partial \beta} &= - \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [-f_1(x_i; \xi, \gamma) + f_2(x_i; \lambda)] \\
&\quad [-f_1(x_i; \xi, \gamma) + f_3(x_i; \mu, \sigma)] \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \alpha \partial \xi} &= - \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [(1-\beta)f_2(x_i; \lambda) + \beta f_3(x_i; \mu, \sigma)] \\
&\quad \left(\frac{\gamma}{\xi} - x_i\right) f_1(x_i; \xi, \gamma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \alpha \partial \gamma} &= \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [(1-\beta)f_2(x_i; \lambda) + \beta f_3(x_i; \mu, \sigma)] \\
&\quad [\Psi_{(\gamma)} - \log(\xi x_i)] f_1(x_i; \xi, \gamma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \alpha \partial \lambda} &= \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [(1-\beta)f_1(x_i; \xi, \gamma) + \beta f_3(x_i; \mu, \sigma)] \\
&\quad \left(\frac{1}{\lambda} - x_i\right) f_2(x_i; \lambda) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \alpha \partial \mu} &= -\beta \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [-f_1(x_i; \xi, \gamma) + f_2(x_i; \lambda)] \\
&\quad \left(\frac{x_i - \mu}{\sigma^2}\right) f_3(x_i; \lambda, \sigma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \alpha \partial \sigma} &= \beta \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [-f_1(x_i; \xi, \gamma) + f_2(x_i; \lambda)] \\
&\quad \frac{1}{\sigma} \left[1 - \left(\frac{x_i - \mu}{\sigma}\right)^2\right] f_3(x_i; \mu, \sigma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \beta^2} &= - \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [-f_1(x_i; \xi, \gamma) + f_3(x_i; \mu, \sigma)]^2 \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \beta \partial \xi} &= - \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [\alpha f_2(x_i; \lambda) + (1-\alpha)f_3(x_i; \mu, \sigma)] \\
&\quad \left(\frac{\gamma}{\xi} - x_i\right) f_1(x_i; \xi, \gamma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \beta \partial \gamma} &= \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [\Psi_{(\gamma)} - \log(\xi x_i)] [\alpha f_2(x_i; \lambda) \\
&\quad + (1-\alpha)f_3(x_i; \mu, \sigma)] f_1(x_i; \xi, \gamma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \beta \partial \lambda} &= \alpha \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [f_1(x_i; \xi, \gamma) - f_3(x_i; \mu, \sigma)] \\
&\quad \left(\frac{1}{\lambda} - x_i\right) f_2(x_i; \lambda, \xi, \gamma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \beta \partial \mu} &= \sum_{i=1}^n \frac{1}{[\hat{p}_{GEN}(x_i; \phi)]^2} [(1-\alpha)f_1(x_i; \xi, \gamma) + \alpha f_2(x_i; \lambda)]
\end{aligned}$$

$$\begin{aligned}
& \left(\frac{x_i - \mu}{\sigma^2} \right) f_3(x_i; \mu, \sigma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \beta \partial \sigma} &= - \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \frac{1}{\sigma} \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_3(x_i; \mu, \sigma) \\
& \quad [(1-\alpha)f_1(x_i; \xi, \gamma) + \alpha f_2(x_i; \lambda)] \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \xi^2} &= (1-\alpha-\beta) \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \left\{ -(1-\alpha-\beta) \frac{\gamma}{\xi^2} f_1(x_i; \xi, \gamma) \right. \\
& \quad \left. + \left[\frac{\gamma(\gamma-1)}{\xi^2} - \frac{2\gamma}{\xi} x_i + x_i^2 [\alpha f_2(x_i; \lambda) + \beta f_3(x_i; \mu, \sigma)] \right] f_1(x_i; \xi, \gamma) \right\} \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \xi \partial \gamma} &= (1-\alpha-\beta) \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \left\{ (1-\alpha-\beta) \frac{1}{\xi} f_1(x_i; \xi, \gamma) \right. \\
& \quad \left. + \left[\frac{1}{\xi} - \left(\frac{\gamma}{\xi} - x_i \right) (\Psi_{(\gamma)} - \log(\xi x_i)) \right] [\alpha f_2(x_i; \lambda) \right. \right. \\
& \quad \left. \left. + \beta f_3(x_i; \mu, \sigma)] \right\} f_1(x_i; \xi, \gamma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \xi \partial \lambda} &= -(1-\alpha-\beta) \alpha \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \left(\frac{\gamma}{\xi} - x_i \right) f_1(x_i; \xi, \gamma) \\
& \quad \left(\frac{1}{\lambda} - x_i \right) f_2(x_i; \lambda) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \xi \partial \mu} &= -(1-\alpha-\beta) \beta \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \left(\frac{\gamma}{\xi} - x_i \right) f_1(x_i; \xi, \gamma) \\
& \quad \left(\frac{x_i - \mu}{\sigma^2} \right) f_3(x_i; \mu, \sigma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \xi \partial \sigma} &= (1-\alpha-\beta) \beta \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \left(\frac{\gamma}{\xi} - x_i \right) f_1(x_i; \xi, \gamma) \\
& \quad \frac{1}{\sigma} \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_3(x_i; \mu, \sigma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \gamma^2} &= -(1-\alpha-\beta) \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \left\{ (1-\alpha-\beta) \Psi_{(\gamma)} f_1(x_i; \xi, \gamma) \right. \\
& \quad \left. - [(\log(\xi x_i))^2 - 2\Psi_{(\gamma)} \log(\xi x_i) - \Psi'_{(\gamma)} + (\Psi_{(\gamma)})^2] [\alpha f_2(x_i; \lambda) \right. \right. \\
& \quad \left. \left. + \beta f_3(x_i; \mu, \sigma)] \right\} f_1(x_i; \xi, \gamma) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \gamma \partial \lambda} &= -(1-\alpha-\beta) \alpha \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} [\log(\xi x_i) - \Psi_{(\gamma)}] \\
& \quad f_1(x_i; \xi, \gamma) \left(\frac{1}{\lambda} - x_i \right) f_2(x_i; \lambda) \\
\frac{\partial^2 \log L_{GEN}(\phi; x)}{\partial \gamma \partial \mu} &= -(1-\alpha-\beta) \beta \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} [\log(\xi x_i) - \Psi_{(\gamma)}]
\end{aligned}$$

$$f_1(x_i; \xi, \gamma) \left(\frac{x_i - \mu}{\sigma} \right) f_3(x_i; \mu, \sigma)$$

$$\frac{\partial^2 \log L_{GEN}(\phi; \mathbf{x})}{\partial \gamma \partial \sigma} = (1 - \alpha - \beta) \beta \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} [\log(\xi x_i) - \Psi(\gamma)]$$

$$f_1(x_i; \xi, \gamma) \frac{1}{\sigma} \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] f_3(x_i; \mu, \sigma)$$

$$\frac{\partial^2 \log L_{GEN}(\phi; \mathbf{x})}{\partial \lambda^2} = -\alpha \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \left\{ \left(\frac{2}{\lambda} x_i - x_i^2 \right) [(1 - \alpha - \beta) f_1(x_i; \xi, \gamma) \right.$$

$$\left. + \beta f_3(x_i; \mu, \sigma)] + \alpha \frac{1}{\lambda^2} f_2(x_i; \lambda) \right\} f_2(x_i; \lambda)$$

$$\frac{\partial^2 \log L_{GEN}(\phi; \mathbf{x})}{\partial \lambda \partial \mu} = -\alpha \beta \sum_{i=1}^n \left(\frac{1}{[p_{GEN}(x_i; \phi)]^2} \left(\frac{1}{\lambda} - x_i \right) f_2(x_i; \lambda) \left(\frac{x_i - \mu}{\sigma} \right) \right.$$

$$\left. f_3(x_i; \mu, \sigma) \right)$$

$$\frac{\partial^2 \log L_{GEN}(\phi; \mathbf{x})}{\partial \lambda \partial \sigma} = \alpha \beta \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \left(\frac{1}{\lambda} - x_i \right) f_2(x_i; \lambda) \frac{1}{\sigma} \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right]$$

$$f_3(x_i; \mu, \sigma)$$

$$\frac{\partial^2 \log L_{GEN}(\phi; \mathbf{x})}{\partial \mu^2} = -\beta \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \frac{1}{\sigma^2} \left\{ \left[1 - \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] [(1 - \alpha - \beta) \right.$$

$$\left. f_1(x_i; \xi, \gamma) + \alpha f_2(x_i; \lambda)] + \beta f_3(x_i; \mu, \sigma) \right\} f_3(x_i; \mu, \sigma)$$

$$\frac{\partial^2 \log L_{GEN}(\phi; \mathbf{x})}{\partial \mu \partial \sigma} = -\beta \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \frac{1}{\sigma^2} \left\{ \left[3 \frac{x_i - \mu}{\sigma} - \left(\frac{x_i - \mu}{\sigma} \right)^3 \right] \right.$$

$$\left. [(1 - \alpha - \beta) f_1(x_i; \xi, \gamma) + \alpha f_2(x_i; \lambda)] + 2\beta \frac{x_i - \mu}{\sigma} \right.$$

$$\left. f_3(x_i; \mu, \sigma) \right\} f_3(x_i; \mu, \sigma)$$

$$\frac{\partial^2 \log L_{GEN}(\phi; \mathbf{x})}{\partial \sigma^2} = \beta \sum_{i=1}^n \frac{1}{[p_{GEN}(x_i; \phi)]^2} \frac{1}{\sigma^2} \left\{ \left[2 - 5 \left(\frac{x_i - \mu}{\sigma} \right)^2 + \left(\frac{x_i - \mu}{\sigma} \right)^4 \right] \right.$$

$$\left. [(1 - \alpha - \beta) f_1(x_i; \xi, \gamma) + \alpha f_2(x_i; \lambda)] + \beta \left[1 - 3 \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] \right.$$

$$\left. f_3(x_i; \mu, \sigma) \right\} f_3(x_i; \mu, \sigma)$$

但し, $p_{GEN}(x_i; \phi) = (1 - \alpha - \beta) f_1(x_i; \xi, \gamma) + \alpha f_2(x_i; \lambda) + \beta f_3(x_i; \mu, \sigma)$

$$\Psi'_{(\gamma)} = \frac{d\Psi_{(\gamma)}}{d\gamma}$$

$$\Psi_{(\gamma)} = \frac{d \log \Gamma_{(\gamma)}}{d\gamma}$$

$$\Gamma(r) = \int_0^{\infty} t^{r-1} e^{-t} dt$$

を得る。このモデルについても SAS の NLIN プロシージャを用いガウス・ニュートン法によって推定を行った。推定されたパラメータによって得られた混合密度による出現確率を求め、相対度数とともに図 8 (a), (b), (c) に示す。図から指数分布と正規分布による混合密度のモデルよりもよく適合していることがわかる。さらに適合度をくらべるため AIC を求めた。AIC は、

$$-2\log(\text{最大尤度}) + 2p$$

ここで、 p はパラメータの数

で定義される統計量でモデル間の適合性検定に有効なもので、より小さな値をとるモデルが適合性が良いと判定される。それぞれのモデルに推定されたパラメータを用いて求めた AIC を表 3 に示す。

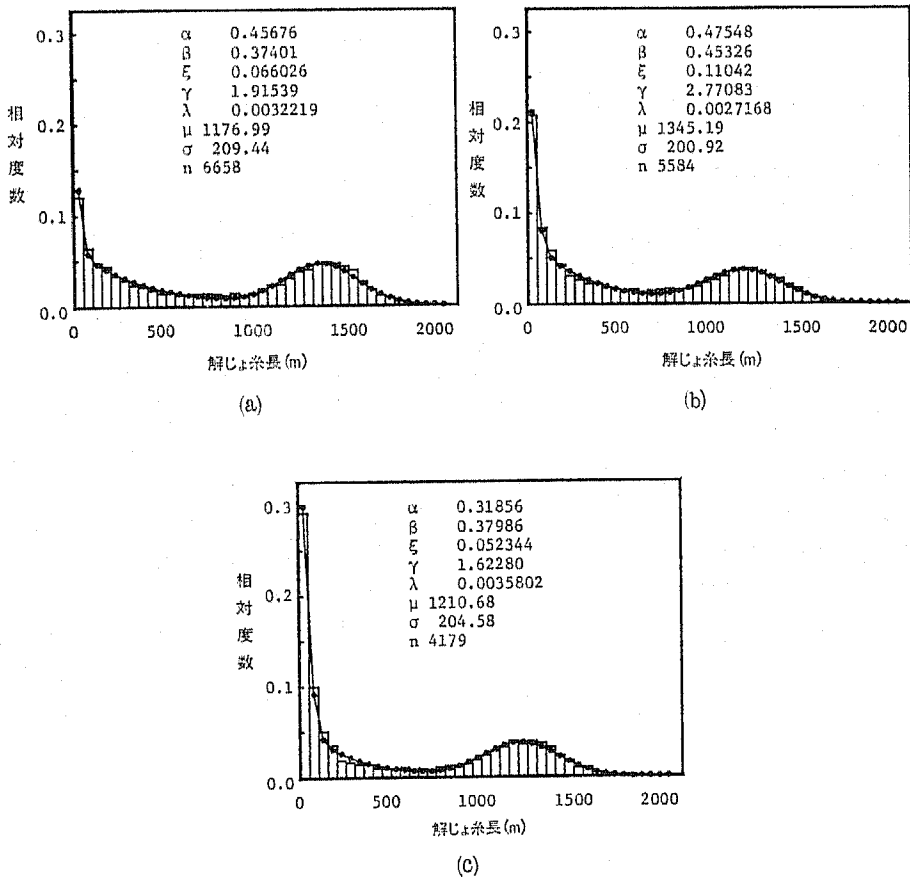


図 8 ガンマ分布、指数分布と正規分布による解じょ糸長分布へのあてはめ

表3 モデル, 蚕期別の AIC

蚕	期	モ デ ル	
		EN	GEN
春	(a)	81537.56	81313.63
初 秋	(b)	94234.81	93696.13
晩 秋	(c)	57233.52	56711.29

文 献

- BMDP statistical software (1981)
 DAY, N.E. : Biometrika, 56, 463-474 (1969)
 DEMPSTER, A.P., LAIRD, N.M. & RUBIN, D.B. : J.R. statist. soc., B, 39, 1-38(1977)
 FISHER, R.A. : Phil. Trans. Roy. Soc., A, 222, 309-368 (1921)
 FRYER, J.G. & ROBERTSON, C.A. : Biometrika, 59, 639-648 (1972)
 HASSELBIAD, V. : Technometrics, 8, 431-444 (1966)
 MCLACHLAN, G.J. & BASFORD, K.E. : Mixture models, Inference and applications to clustering, (1988) Marcel Dekker
 三浦幹彦, 西岡孝彦, 嶋崎昭典 : 応用統計学, 17, 29-42 (1988)
 PEARSON, K. : Phil. Trans. Roy. Soc., A, 185, 71-110 (1894)
 SAS User's Guide, statistics (1982)
 TAN, W.Y. & CHANG, W.C. : J. Amer. Statist. Assoc., 67, 702-708 (1972)
 TITTERNGTON, D.M., SMITH, A.F.M. & MAKOV, U.E. : Statistical analysis of finite mixture distributions, (1985) Wiley

Summary

Estimation of the parameters of finite mixture distribution and its applications

Takahiko NISHIOKA and Mikihiko MIURA

Finite mixture distributions are used as an statistical model of the data in the wide range of areas and many studies have been done about estimating the parameters of the mixture of normal distributions.

In this paper we discussed about the maximum likelihood estimation of the parameters of other continuous component densities, particularly both gamma mixtures and the mixture of an exponential and a normal, by EM algorithm and statistical packages such as SAS or BMDP. Precision of the estimation was also discussed. The results were applied for the statistical analysis of cocoon filament data.