

この点に関しては、 $g(\lambda)$ をどう決めるかの問題とも関連し、今の段階では何とも言えない。今後に残された問題としておきたい。

ベイズ型コウホートモデル

統計数理研究所 中 村 隆

1. コウホート分析

社会の変化を分析するとき、時勢による推移だけでなく、加齢による変化や世代による相違を区別することが重要である。コウホート分析は、継続的調査で得られる何らかの数量特性を年齢層×調査時点別に集計したデータ表から、年齢と時代効果、そして表の斜め方向に現われるコウホート（同時出生集団）の効果を分離しようとする方法である。

ただし、コウホート分析には識別問題が存在し、三効果を一義に分解することは原理的に不可能であるとされていた（Fienberg and Mason, 1979）。これに対し、中村（1982）は、識別問題を克服するためにパラメータの漸進的変化の条件を取り込んだベイズ型コウホートモデルを提案し、種々のコウホートデータに適用している。

2. コウホート表

コウホート分析の分析対象であるデータ表をコウホート表と呼ぶ。コウホート表は、まず、集計されたデータの性質にしたがって、比率型と数量型に分類できる。比率型は表のセルに回答比率を配したもので、数量型は平均子ども数のような数量を配したものである。それぞれ異なる標本抽出モデルを仮定する。

さらに、コウホート表はその形式にしたがって、標準表・一般表・特殊表に分けることができる。標準表とは、年齢×時代の表で調査間隔と年齢区分幅が一致しているものであり、同一コウホートが表の斜めのセルを渡るので分析が比較的容易である。一般表とは、調査間隔と年齢区分幅が一致しない場合である。標準表と一般表が継続調査で得られる年齢×時代の形式であるのに対し、特殊表はそれ以外の形式の表であり、一時点の回顧調査で得られるコウホート×年齢の表などがある。

3. コウホートモデル

ある数量特性 x_{ij} を年齢×時代の形式に配するコウホート表データを考える。これに既知の関数 g を用いて変換を施した数量 y_{ij} を次のように分解する：

$$(1) \quad y_{ij} \equiv g(x_{ij}) \cong \mu + \mu_i^A + \mu_j^P + \mu_k^C,$$

$$(2) \quad \sum \mu_i^A = \sum \mu_j^P = \sum \mu_k^C = 0.$$

ここで、 μ_i^A 、 μ_j^P 、 μ_k^C はそれぞれ年齢・時代・コウホート効果のパラメータであり、 μ は総平均効果である。

(1)式は(2)式を考慮して次のように書くことができる：

$$(3) \quad \mathbf{y} = D\boldsymbol{\mu}.$$

ここに、 $\mu' = (\mu, \mu^*)$, $\mu^* = (\mu_1^A, \dots, \mu_{I-1}^A, \mu_1^P, \dots, \mu_{J-1}^P, \mu_1^C, \dots, \mu_{K-1}^C)$ であり、また $D = [1D^*]$ は、コウホート表のセルとパラメータとを結びつける適当なデザイン行列である。標準・一般・特殊というコウホート表の違いは、このデザイン行列を工夫することによって対処する。

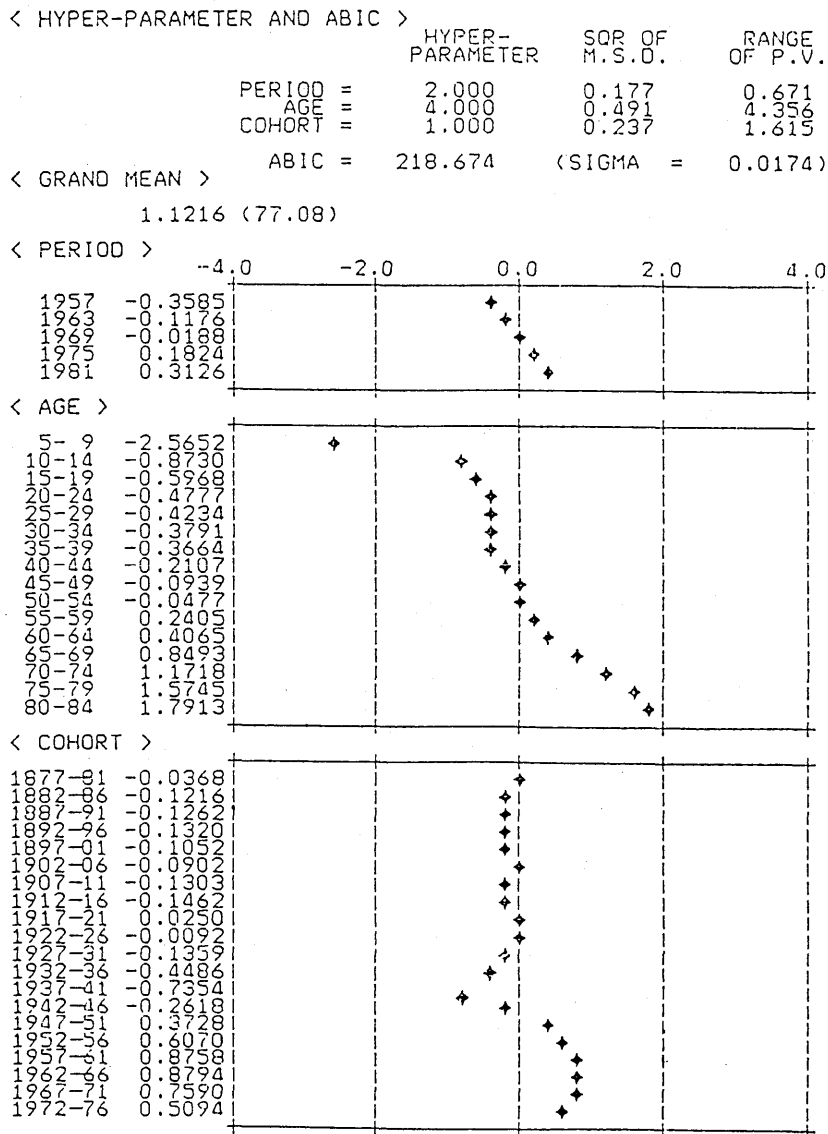


図1. ベイズ型モデルによる齲蝕率のコウホート分析結果（下顎第1大臼歯，男）

Fig.1. Estimated Effects for the Lower First Molar DMFT of Males, based on the Bayesian Cohort Model.

4. 標本抽出モデル

比率型の場合, x を比率 p , g をロジット変換 $\ln p/(1-p)$ とし, 標本抽出モデルに積の二項分布を仮定すれば, モデルの尤度 f は

$$(4) \quad f(\mathbf{n} | \mathbf{p}(\boldsymbol{\mu})) = \prod \binom{N_i}{n_i} p_i^{n_i} / (1-p_i)^{N_i-n_i}$$

となる. ここで, N_i は各セルの標本の大きさ, n_i は正反応度数である.

数量型の場合, y が正規分布に従うと仮定すると, モデルの尤度 f は

$$(5) \quad f(\mathbf{x} | \boldsymbol{\mu}, \sigma^2) = (2\pi\sigma^2)^{-L/2} |G| \cdot \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y}-D\boldsymbol{\mu})' W^{-1}(\mathbf{y}-D\boldsymbol{\mu})\right\}$$

となる. ここで, $G = \text{diag}\{\dots, dg(x_i)/dx, \dots\}$, $|W|=1$ である.

最尤法を用いてパラメータを推定するのがこれまでの通常のやり方だが, コウホート分析における識別問題のために解は不定となる. $\boldsymbol{\mu}$ を特定するには何らかの付加条件が必要である.

5. パラメータの漸進的変化の条件

不定な解の中から, 節約原理を満たすような解を求めるために, パラメータの漸進的変化の条件をモデルに取り込む. 具体的には, パラメータの1次階差の2乗和

$$(6) \quad \sum(\mu_i^A - \mu_{i+1}^A)^2 / \sigma_A^2 + \sum(\mu_j^P - \mu_{j+1}^P)^2 / \sigma_P^2 + \sum(\mu_k^C - \mu_{k+1}^C)^2 / \sigma_C^2$$

を小さくすることを考える. すなわち, これを正規分布の密度関数の形に書換えた

$$(7) \quad \pi(\boldsymbol{\mu} | \sigma_A^2, \sigma_P^2, \sigma_C^2, \sigma^2) = (2\pi)^{-M/2} |D_S' S^{-1} D_S|^{1/2} \exp\left\{-\frac{1}{2} \boldsymbol{\mu}'_* D_S' S^{-1} D_S \boldsymbol{\mu}_*\right\}$$

がなるべく大きくなるような解を求める. これを, 尤度 f を大きくすることとのバランスを考えて実現し, 結局,

$$(8) \quad f \cdot \pi$$

を最大化することによってパラメータ $\boldsymbol{\mu}$ を推定する. ここで, $\sigma^2 S = \text{diag}\{\sigma_A^2, \dots, \sigma_A^2, \sigma_P^2, \dots, \sigma_P^2, \sigma_C^2, \dots, \sigma_C^2\}$. また, D_S は1次階差を表現する行列である.

6. ベイズ型モデル

超パラメータの σ_A^2 等が決まれば(7)式の最大化は容易である. ただし, σ_A^2 等の決定が問題である. これに対しては, Akaike (1980) の提案したベイズ型情報量規準 ABIC を最小にする方式をとる.

比率型コウホートモデルの場合,

$$(9) \quad \begin{aligned} \text{ABIC} = & 2\sum\{n_i \ln n_i/\hat{n}_i + (N_i - n_i) \ln(N_i - n_i)/(N_i - \hat{n}_i)\} \\ & + \hat{\boldsymbol{\mu}}'_* D_S' S^{-1} D_S \hat{\boldsymbol{\mu}}_* - \ln |D_S' S^{-1} D_S| \\ & + \ln |D_V' V^{-1} D_V + D_S' S^{-1} D_S| \end{aligned}$$

となる. ここで, $D_V = D_* - (1' V^{-1} 1)^{-1} 1' V^{-1} D_*$, $V = \text{diag}\{\dots, N_i/\hat{n}_i(N_i - \hat{n}_i), \dots\}$.

数量型コウホートモデルの場合,

$$(10) \quad \text{ABIC} = L \ln \hat{\sigma}^2 - 2 \ln |G| - \ln |D'_s S^{-1} D_s| + \ln |D'_w W^{-1} D_w + D'_s S^{-1} D_s|$$

となる。ここで、 $\hat{\sigma}^2 = \frac{1}{L} \{(\mathbf{y} - D\hat{\boldsymbol{\mu}})' W^{-1}(\mathbf{y} - D\hat{\boldsymbol{\mu}}) + \hat{\boldsymbol{\mu}}' D'_s S^{-1} D_s \hat{\boldsymbol{\mu}}\}$, $D_w = D_* - (\mathbf{1}' W^{-1} \mathbf{1})^{-1} \cdot \mathbf{1} \mathbf{1}' W^{-1} D_*$ である。

7. 適用例

歯科疾患実態調査は昭和32年から5回継続して調査されている。ただし、調査間隔は6年、年齢区分は5歳刻みであり、一般コウホート表のデータとなる。オーバーラップを許すコウホート区分を想定してデザイン行列を設定すればよい。齲蝕経験率であるので比率型のモデルを用いている。ここでは、男の下顎第1大臼歯についての分析結果を図1に示す。実際にはコウホート効果のパラメータはもっとたくさんあるが、図では見やすくするためにまびいてある。

結果をみると、まず、年齢効果が大きい。時代効果は漸増傾向にある。コウホート効果では、昭和10年前後生まれで齲蝕が相対的に少なく、戦後生まれで急速に増加している。しかし、昭和40年以降生まれの若いコウホートでの齲蝕の減少傾向が認められる。男女別に比べると20歳～50歳における年齢効果が異なっていることがみられた(那須他, 1984)。

参考文献

- Akaike, H. (1980). Likelihood and the Bayes Procedure, *Bayesian Statistics* (Eds. J.M. Bernardo, et al.). Valencia: University Press.
- Fienberg, S.E. and Mason, W.M. (1979). Identification and Estimation of Age-Period-Cohort Models in the Analysis of Discrete Archival Data, *Sociological Methodology, 1979*.
- Glenn, N.D. (1977). *Cohort Analysis*, Sage Publications. (藤田訳「コウホート分析法」, 朝倉書店).
- 中村 隆(1982). ベイズ型コウホート・モデル—標準コウホート表への適用—, 「統計数理研究所彙報」, 29巻2号.
- 那須郁夫, 森本 基, 中村 隆(1984). 下顎第1大臼歯齲蝕経験のコウホート分析—歯科疾患実態調査報告資料による—. 「口腔衛生学会雑誌」, 34巻3号.

セメントロータリキルンの統計的制御

システム総合開発(株) 八木原 彬 殷

1. 何が問題だったか

セメント焼成工程は原料の流れ方向の順に、① 予熱、仮焼反応を受け持つSPタワー、② 焼成反応を主体とするロータリキルン及び、③ クリンカ(セメント半製品)を空冷するクリンカ・クーラの3つの工程から成る一種の向流型熱交換機である。各工程が安定しているときは、各々の独立性が顕著で、各工程単独で制御していても何ら問題はないが、一旦系が乱れると相互干渉が発生し、図1に示したような20～25分周期のハンチング現象を起す。この現象は系内に存在する装置上のfeedback loopによるもので、完全なる遅れ振動として把えることが出来るが、図5に示したように従来の1入力1出力の定値制御ではこれを安定にするのは困難である。図1のようなハンチングが突如として起るため(何らかの原因はあるはずだが)、オペレー