

統計モデルとブートストラップ確率

(客員) 東京大学教養学部 岸野 洋久

いくつかの統計モデルが設定されたとき、どのモデルがデータに最もよく適合しているかを調べるモデル選択に加えて、実際にはそのモデルがどの程度他に比して勝っているか評価することが重要となることが多い。対比較による仮説検定から最終的な結論を導き出すのは、しばしば困難を来す。分子系統学において Felsenstein (1985) が、ブートストラップによりリサンプリングされたデータの中でそれぞれのモデルを支持する割合を算出し、これに基づいてモデルの信頼性を評価することを提示した。現在これは特にこの分野の解析では欠かせないアプローチとして普及している。

最近になって、この統計モデルの比較評価法の統計的特性がようやく調べられ始めた。Hillis and Bull (1993) はシミュレーションを通じて、選択されたモデルのブートストラップ確率が真のモデルを当てる確率を反映していないことなどを示し、これに基づく信頼性の評価法自体に疑問を投げかけた。これに対して、Zharkikh and Li (1992), Felsenstein and Kishino (1993) は簡単な例によりこうした現象を解析的に調べるとともに、ブートストラップ確率を第二種の過誤の安全サイドの評価として使うか、あるいはこれに基づいてモデルを集めた信頼集合を構成することを提案した。ブートストラップ確率の解釈、或いははっきりした信頼集合の構成などが今後の課題として残された。

参考文献

- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap, *Evolution*, **39**, 783-791.
- Felsenstein, J. and Kishino, H. (1993). Is there something wrong with the bootstrap on phylogenies? A reply to Hillis and Bull, *Systematic Biology* (in press).
- Hillis, D.M. and Bull, J.J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis, *Systematic Biology* (in press).
- Zharkikh, A. and Li, W.H. (1992). Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences, I: four taxa with a molecular clock, *Molecular Biology and Evolution*, **9**, 1119-1147.

領域統計研究系

平滑化法における推定量の性能

柳本 武美

1. 序

回帰モデル

$$y_i = g(i/n) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

において $\mu = (\mu_1, \dots, \mu_u)' = (g(1/n), \dots, g(1))$ と σ^2 を推定する問題を考える。ここでは関数 $g(\cdot)$ として母数モデルを仮定しないで、 $g(\cdot)$ が滑らかに変化するという仮定をおく。この問題は平滑化法あるいはノンパラメトリック回帰として理解され、多くの手法が提示されているが、それらの性能の比較研究は十分でない。

2. 比較の規準

推定量 $\hat{\mu}$ の良さを計る規準としては通常残差平方和の平均が用いられる。分散の推定量 $\hat{\sigma}^2$ の規準としてはバイアス, Kullback-Leibler loss など通常の規準が考えられる。従来は別々に研究されてきた。

推定量の良さを調べるためには他にも規準があることが望ましい。先ず $\|\mu\|^2$ の推定量の良さを調べることである。しかしこれは分散の推定量の評価をすれば良い。次に滑らかさ, 例えば $\sum(\mu_i - 2\mu_{i+1} + \mu_{i+2})^2$ の推定量を調べるのが考えられる。ところが既存の文献ではこの量の推定量は明示的には与えられていない。

3. 比較のための条件設定

ここでは平滑化法として経験 Bayes 法, Wahba による GCV を用いる方法を比較した。

条件設定での困難は妥当な関数 $g(\cdot)$ を選択することである。ここでは Korn et al. (1991) に見える関数を中心にした。さらに計画のための規準を豊富にすることが望ましい。このために平均ピタゴラス関係を適用する。良い推定量であればこの関係が近似的に成り立つと期待される。そうすると前に述べた新しい評価規準が示唆される。

4. 結 果

Simulation 研究で得られた知見は次の通りである。

- a) 既存の方法では平滑化不足のことが多い。この対策としては Buckley et al. (1988) の分散推定量を用いることによって少し改善される。
- b) GCV 法は経験 Bayes 法に比べて平滑化不足の傾向が強い。
- c) 分散の推定のためには経験 Bayes 法では条件付推測の処理をした方が良い。それでも分散は負の bias をもつ傾向がある。
- d) 平滑化法の性能はトレード・オフ母数と $g(\cdot)$ に依存する。Korn らの比較は σ^2 が小さいときに偏っている。

参 考 文 献

- Buckley, M.J., Eagleson, G.K. and Silverman, B.W. (1988). The estimation of residual variance in nonparametric regression, *Biometrika*, **75**, 189-199.
- Korn, R., Ansley, C.F. and Tharm, D. (1991). The performance of cross-validation and maximum likelihood estimators of spline smoothing parameters, *J. Amer. Statist. Assoc.*, **86**, 1042-1050.

じゃんけんモデルと非線型積分可能な力学系の確率モデル

伊 藤 栄 明

型 1, 型 2, 型 3 という 3 つの型の粒子があり, 1 は紙, 2 ははさみ, 3 は石, というように巡回的な強弱関係があるものとする。それぞれに n_1, n_2, n_3 個の粒子がはいっていたとし, 3 つの型の粒子の総数は n であり単位時間内にランダムに 2 つの粒子が衝突し弱いほうの粒子は強いほうの粒子に変化する。たとえば 1 の粒子と 2 の粒子は衝突により 2 の 2 粒子になるものとする。同じ型の 2 粒子の衝突では変化はおきないものとする。1, 2, 3 の比率はそれぞれ確率変数 $x_1(t), x_2(t), x_3(t)$ であらわされる。そのとき時刻 t での積の値を $x_1(t)x_2(t)x_3(t)$ とした条件付き期待値について次の関係が成り立つ (Itoh (1973, 1979b)),