連続成功連の数の分布

平　野　勝　臣

本年度の研究

(1)　これまでの離散確率分布についての研究を幾つかまとめた (Aki and Hirano (1993a, 1993b), Hirano and Aki (1993)).

(2)　システムの信頼性，とくに Consecutive-k-out-of-n : F systems について調べた. これは来年度に継続される.

ここでは Hirano and Aki (1993) の要旨を述べる.

要　　旨

$X_0, X_1, \ldots, X_n$ を初期分布と推移確率が指定された，0 か 1 のいずれかの値をとる time-homogeneous なマルコフ連鎖とする. 値 1 をとることを成功と言うことにする.

確率変数 $X_1, \ldots, X_n$ の系列において，条件 $X_0=0$ が与えられたときと，$X_0=1$ が与えられたときの，長さ k 以上の連続成功連の起こる回数の確率関数とその漸化式，確率生成母関数とその漸化式を与えた. このことからベルヌーイ試行の場合について，長さ k 以上の連続成功連の起こる回数の平均や分散など，また確率を求める便利な式も与えた.

つぎに連続した成功をオーバーラップして数えるとき，上と同様な結果を導いた. 即ち，条件 $X_0=0$ が与えられたときと，$X_0=1$ が与えられたときの，長さ k の連続成功連の起こる回数の確率関数とその漸化式，確率生成母関数とその漸化式を与えた.

参　考　文　献

Aki, S. and Hirano, K. (1993a). Discrete distributions related to succession events in a two-state Markov chain, *Statistical Sciences and Data Analysis*; *Proceedings of the Third Pacific Area Statistical Conference* (ed. K. Matusita et al.), 467-474, VSP, Utrecht.

Aki, S. and Hirano, K. (1993b). On waiting time distributions of succession events, Research Memo., No. 473, The Institute of Statistical Mathematics, Tokyo.

Hirano, K. and Aki, S. (1993). On number of occurrences of success runs of specified length in a two-state Markov chain, *Statistica Sinica*, **3**, 313-320.

ブートストラップ法と予測誤差推定

小　西　貞　則

ブートストラップ法を予測誤差の推定問題に適用することによって，従来解析的アプローチが難しかった様々な問題に対して，有効な解を与えることが可能となりつつある. このような例として，情報量規準 AIC (Akaike (1973)) に於ける対数尤度のバイアス補正，判別分析に於ける誤判別率の推定問題を取り上げ，特に，バイアス推定の変動の一減少法について検討した.

未知の確率分布 $G(x)(g(x))$ からの大きさ n の無作為標本を $\boldsymbol{\chi}_n$ とする. 観測されたデータを通して得られる情報をもとに確率分布モデルを構成し，これを $f(x\mid\boldsymbol{\theta})$ とおく. 標本 $\boldsymbol{\chi}_n$ に基づいて推定された一つの予測確率分布 $\hat{f}(z\mid\boldsymbol{\chi}_n)$ と，この標本を生成した確率分布との距離を Kullback-Leibler 情報量で測るとき，予測確率分布の対数尤度のバイアス

$$E_G[T(G;\ \boldsymbol{\chi}_n)]=E_G\left[\int g(z)\log \hat{f}(z\mid \boldsymbol{\chi}_n)dz-\frac{1}{n}\sum_{i=1}^{n}\log \hat{f}(X_i\mid \boldsymbol{\chi}_n)\right]$$

の推定とその補正が本質的となる．従って，一つの量 $T(G;\ \boldsymbol{\chi}_n)$ の変動を可能な限り制御し，精度良く期待値を推定する方法が求められる．

一般に，推定量の変動はその影響関数を通して大体の様相を把握することができる．いま，ある正則条件のもとで $T(G;\ \boldsymbol{\chi}_n)$ は，その影響関数 $IF(x;\ G)$ を用いて

$$T(G;\ \boldsymbol{\chi}_n)=\frac{1}{n}\sum_{i=1}^{n}IF(X_i;\ G)+o_p(n^{-1/2}) \tag{1}$$

と表されたとする．ブートストラップ法を適用して期待値を推定するとき，影響関数の期待値は 0 であることを利用すると上式は

$$\begin{aligned}T(G;\ \boldsymbol{\chi}_n)-\frac{1}{n}\sum_{i=1}^{n}IF(X_i;\ G)&=\int g(z)\log \hat{f}(z\mid \boldsymbol{\chi}_n)dz-\int g(z)\log f(z\mid \boldsymbol{\theta})dz\\&\quad+\frac{1}{n}\sum_{i=1}^{n}\log f(X_i\mid \boldsymbol{\theta})-\frac{1}{n}\sum_{i=1}^{n}\log \hat{f}(X_i\mid \boldsymbol{\chi}_n)\end{aligned} \tag{2}$$

の期待値の推定にブートストラップ法を適用することと同じことになる．(1) 式と (2) 式の違いは，(1) 式の漸近分散のオーダーが $O(1/n)$ であるのに対して，(2) 式の漸近分散のオーダーは $O(1/n^2)$ となるところにある．

情報量規準 EIC の構成に於ける (2) 式の利用は，北川 (1991) によって指摘され数値的にその良さが確かめられているが，理論的にもその有効性が示された．以上の考え方は，一般に予測誤差の推定問題において，バイアス推定の変動をおさえる一つの方法として用いることができ，ここでは特に判別分析に於ける誤判別率推定の問題に対して検討を行った．

参 考 文 献

Akaike, H. (1973). Information theory and an extention of the maximum likelihood principle, *Proceedings of 2nd International Symposium on Information Theory* (eds. B.N. Petrov and F. Csaki), 267-281, Akademiai Kiado, Budapest.

北川源四郎 (1991). 対数尤度のブートストラップについて,「時系列に関する推測の理論と応用」, 共同研究レポート, No. 31, 175-179, 統計数理研究所.

第 60 回日本統計学会 (1992). 共通テーマ：ベイズ統計学の理論と応用 (II), 日本統計学会講演予稿集, 255-271.

Semimartingale Regression Model のノンパラメトリック推定

吉 田 朋 広

次の分解を持つ semimartingale regression model を考える．

$$\begin{aligned}&X_i(t)=X_i(0)+B_i(t)+M_i(t), \qquad t\in[0,\ 1], \qquad i=1, 2, \dots, n,\\&B_i(t)=\int_0^t Y_i(s)b(s,\ Z_i(s))ds,\\&M_i(t)=M_i^c(t)+M_i^d(t),\\&\langle M_i^c\rangle(t)=\int_0^t Y_i(s)c^2(s,\ Z_i(s))ds,\\&M_i^d(t)=\int_0^t\int_{R^*}x[\mu_i(ds,\ dx)-\nu_i(ds,\ dx)],\\&\nu_i(ds,\ dx)=Y_i(s)n(s,\ x,\ Z_i(s))\nu_0(dx)ds.\end{aligned} \tag{1}$$