5) Hasegawa, M. and Yano, T. (1984) *Proc. Japan Acad.*, 60B, 389-392.

ノンパラメトリックなハザード関数の推定

鎌 倉 稔 成

系列事象の統計的解析では，強度関数（intensity function），あるいはハザード関数が重要な役割を演ずる．ハザード関数は信頼性理論の分野では故障率関数と呼ばれており，パラメトリック，ノンパラメトリック，数々の推定法がある（Singpurwalla & Wong, 1893）．よく用いられるのはワイブル分布，ガンマ分布，レーリー分布などのパラメトリックなモデルを仮定してそのパラメータを推定するといったかたちでハザード関数を求める方法である．ハザード関数をノンパラメトリックに求めるのは，密度関数をノンパラメトリックに求めるのと同程度のむずかしさがある．通常は，累積ハザード関数を推定することが多い．累積ハザード関数の推定量としては，Nelson の推定量，Perterson の推定量がよく使われている．ハザード関数そのものの推定方法としては，actuarial estimates, kernel estimates, などがあり，漸近理論の研究もある．ここでは，ノンパラメトリックに滑らかなハザード関数の推定を行う問題を考える．滑らかなハザード関数を推定する方法としては，Klotz (1982) が滑らかな生存時間分布関数の推定に用いた B スプラインの利用が考えられる．この他，Akaike (1980) の Bayes モデルの利用も1つの方法といえる．本小稿では B スプラインによるハザード関数の推定を議論する．

スプライン関数は区分的な多項式であり，プロットされた点を通る滑らかな曲線を得るために用いられており，種々の応用が考えられている（市田・吉本，1979）．ここでは，基準化された B スプラインを用いる．Klotz (1982) では2階の B スプラインを用いており，滑らかな生存時間分布関数の推定を行うにはさほど問題はないが，ハザード関数の推定には2階の B スプラインでは十分とはいえない．尤度計算には B スプライン関数の積分が必要になるが，これは簡単には表現されない．筆者が数式処理プログラム REDUCE 3 を用いて $\int_0^x N_{k,r}(u)du$ の計算を行った結果を記述する．たとえば，2階の B スプラインでは次のようになる．

i) $x \leqq \tau_k$

0

ii) $\tau_k < x \leqq \tau_{k+1}$

$(\tau_k - x)^2/2(\tau_{k+1} - \tau_k)$

iii) $\tau_{k+1} < x \leqq \tau_{k+2}$

$-(\tau_{k+2}\tau_{k+1} + \tau_{k+2}\tau_k - 2\tau_{k+2}x - \tau_{k+1}\tau_k + x^2)/2(\tau_{k+2} - \tau_{k+1})$

iv) $\tau_{k+2} < x$

$(\tau_{k+2} - \tau_k)/2$

第3研究部

医学データ解析における不適切問題の解法

田 辺 國 士

肺内ガス交換機能解析に登場するある不適切問題の統計的解法を与えた．問題は被験者の呼気に含まれるガスの分圧を測定し，この情報から肺内での換気・血流比 r の分布を推定することである．分布を換気血流比 r の対数 $R = \log r$ の密度函数 $p(R)$ で表現するとき，定常ガス交換における拡散平衡に関する物質保存の原理から，観測値 d と密度函数 $p(R)$ の間に

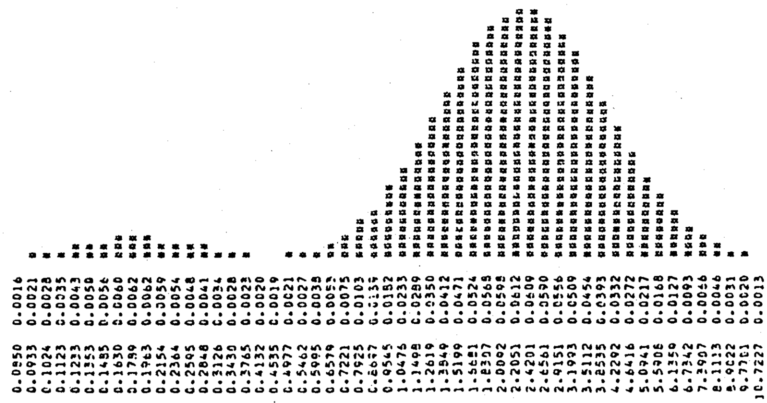
(1) $$d = K(p) + \varepsilon$$

なる関係式が導びかれる．ただし，K は線型写像，d は有限次ベクトル，ε は観測誤差ベクトルである．データ数（すなわち d の次元）は測定できるガスの種類の数であり，通常 3～6 である．このような数の少ないデータから p を推定するためには，p に関する事前情報が必要であり，統計的モデリングが要求される．従来は，$p(R)$ に関して 2 つの正規分布の混合分布などさまざまなパラメトリックモデル $p_\theta(R)$（θ のうち自由パラメタ数は n）を仮定し，最小二乗法などで，パラメター θ の値を推定していた．しかし，換気•血流比が混合正規分布でうまく表現できる医学的保証はなく，このモデルは多分に，データ数と未知パラメタ数の辻つま合せのために導入されたものである．その上，最小二乗法によるパラメタ推定において残差平方和が，多数の極値をもつという数値的困難をもかかえており，いずれかの極値を選んでも妥当な結果が得られないことが多い．実際，健康体の場合となんらかの疾患がある場合では $p(R)$ のプロファイルがかなり異なることが予想され，適当なパラメトリックモデルを構成することは困難である．

我々はこの困難を解決するため，$p(R)$ を R の離散点上の値 $(p_\lambda)\equiv\bar{p}$ で表現し (1) を離散近似したモデル

$$d=\bar{K}(\bar{p})+\varepsilon$$

を考え，Leonard の logistic density transform を用いて $\bar{p}$ を f でパラメトライズし，f に関する smoothness prior を導入し，ベイズの方法によって，$p(R)$ の推定値を得た．得られた結果は，医学的に見ても不自然な点はないことが認められた．（この研究は慶大医学部川城丈夫氏との共同研究中）



推定例

Hamming 距離によるランダムパッキング

伊 藤 栄 明

符号化理論における Golay code より 24 次元空間における球の密な規則的充塡がみちびかれることは知られている．ここでは Hamming 距離にもとづくランダムパッキングを考える．各座標が 0 と 1 からなる次元 d の点を考える．そのうちのひとつを先ずランダムにえらぶ．それを I_1 とする．$r(\geqq 1)$ 個の点がすでにえらばれたとすると I_{r+1} は I_1, I_2, $\cdots$, I_r のいずれからも k 以上はなれている点のなかから等確率でえらばれるものとし，これを可能なかぎりつづける．このようにして得られる点の個数 $X(d, k)$ についての計算機実験を行った．$k=2, 3$ について，充塡率 $X(d, k)/2^d$ の平均値の従う実験