（i） 周辺化（margining）—適当な統計量への周辺モデルを作る；

（ii） 条件付（conditioning）—適当な統計量に関する条件付モデルを作る．

これらの方法でサブモデルを構成したとき，我々が知りたいのは，そのモデルが，上記の目的にどの程度かなったものであるのか—即ち，そのサブモデルが，もとのモデルの有する“当該パラメータに関する情報量”をどの程度保有しているか—である．

そのためには，上記の情報量の概念を明確に定義する必要がある．そうすれば，セパレート推測に現われる定性的諸概念を，情報量概念を通して把握することができ，セパレート推測の統計構造がより明らかになる．

本稿では，その定義について述べるだけとする．

2. “情報量”の定義

$\{P_\theta;\ \theta\in R^n\}$ を統計モデルとする．推測の対象は，パラメータ関数 $\gamma(\theta)=(\gamma_1(\theta),\ \cdots,\ \gamma_k(\theta))'$ である．そのとき，この統計モデルの有する“$\gamma(\theta)$ に関する情報量” $I_f(\theta;\ \gamma)$ を以下のように定義する．まず必要な記号を列挙する．

f_θ: モデル $\{P_\theta:\ \theta\in R^n\}$ の尤度関数；

$$\Gamma(\theta)=[\partial\gamma_i/\partial\theta_j:\ i=1,\ \cdots,\ k;\ j=1,\ \cdots,\ n],\quad \text{ただし，}\ \mathrm{rank}\ \Gamma(\theta)=k;$$

$$I_f(\theta)=[E_\theta(f'_{\theta i}\,f'_{\theta j}/f_\theta^2):\ i,\ j=1,\ \cdots,\ n],\quad \text{ここで，}\ f'_{\theta i}=\partial f_\theta/\partial\theta_i;$$

$$M_f(\theta)=\begin{bmatrix} I_f(\theta) & \Gamma(\theta)' \\ \Gamma(\theta) & 0 \end{bmatrix}.$$

定義． $I_f(\theta;\ \gamma)$ を次で定義する：

$$M_f(\theta)^{\sim}=\begin{bmatrix} C_1 & C_2' \\ C_2 & -I_f(\theta;\ \gamma) \end{bmatrix},$$

ここで，$M_f(\theta)^{\sim}$ は，$M_f(\theta)$ の対称かつ反射的な一般逆行列を表わす．

統計量 T が与えられたとき，その周辺モデルの尤度関数を g_θ, 条件付モデルの尤度関数を h_θ とする．これらサブモデルの有する $\gamma(\theta)$ に関する情報量 $I_g(\theta;\ \gamma)$, $I_h(\theta;\ \gamma)$ も上と同様に定義すればよい．

一般に，これら3つの情報量の間には，

$$I_f(\theta;\ \gamma)\geqq I_g(\theta;\ \gamma)+I_h(\theta;\ \gamma)$$

なる関係がある．等号が成り立つ場合，$I_f(\theta;\ \gamma)=I_g(\theta;\ \gamma)$ (or $=I_h(\theta;\ \gamma)$) の場合などが，セパレート推測に現われる重要な諸概念と極めて密接な関係がある．それらについては，別の機会に述べることにする．

経験分布のマルチンゲール項による適合度検定

安　芸　重　雄

Khmaladze [3] は複合仮説適合度検定問題において，その極限分布が分布型，真の母数及び推定量に依存しないような統計量を一般的に構成し得ることを示唆している．彼は estimated empirical process の漸近的なマルチンゲール項がブラウン運動に弱収束することを示したが，具体的な統計量は構成していない．ここではマルチンゲール項の関数となるような統計量の性質を調べるため単純仮説適合度検定問題について経験分布のマルチンゲール項の関数となる統計量を構成してその性質について報告する．

$X_1,\ X_2,\ \cdots,\ X_n$ は iid で $[0,\ 1]$ 上の一様分布に従うとする．F_n を $X_1,\ \cdots,\ X_n$ から得られる経験分布

とすると，empirical process のマルチンゲール項は

$$W_n(t)=\sqrt{n}\left(F_n(t)-\int_0^t \frac{1-F_n(s)}{1-s}\,ds\right)$$

となり仮説の下で標準ブラウン運動 $W(t)$ に弱収束する．
以下で具体的に統計量を構成する．

I. 線形汎関数

$h'\in L_2[0,1]$ となる h に対して統計量 T_n^h を $T_n^h=\int_0^1 h(t)dW_n(t)$ と定義すると T_n^h は漸近的に平均 0，分散 $\int_0^1 h^2(t)dt$ の正規分布に従う．

II. Neyman の smooth test

W_n の汎関数として Neyman の smooth test T_n を導くことができる．ここで

$$T_n=\frac{1}{n}\sum_{j=1}^{k}\left(\sum_{i=1}^{n}\pi_j(X_i)\right)^2,$$

$\pi_1,\cdots,\pi_k$ は $[0,1]$ 上の正規直交多項式である．

命題． $T_n=\sum_{j=1}^{k}\left(\int_0^1\frac{1}{1-t}\left(\int_0^t\pi_j'(s)(1-s)ds\right)dW_n(t)\right)^2$ と書ける．（証明は［1］参照）

III. Kolmogorov-Smirnov 型

$T_n=\sup_{0\leqq t\leqq 1}|W_n(t)|$ と定義すると T_n は $\sup_{0\leqq t\leqq 1}|W(t)|$ に弱収束するが，この分布は Butler［2］の統計量の漸近分布と一致する．また exact 分布も求まる．（［1］参照）

IV. Cramér-von Mises 型

$T_n=\int_0^1 W_n^2(t)dt$ と定義すると，T_n は $\int_0^1 W(t)^2dt$ に弱収束するが，この分布は Rothman-Woodroofe［5］の統計量の漸近分布と一致する．また，weight をつけた統計量 $\hat{T}_n=\int_0^1\phi(t)W_n^2(t)dt$ を考えると，$\hat{T}_n$ の漸近分布は MacNeil［4］の統計量の漸近分布と一致する．

これらの他にも χ^2 型等の統計量が考えられる．また以上の結果のうち漸近分布に関するものは estimated case にも適用できる．

参考文献

［1］ 安芸（1984）「ノンパラメトリック・ロバスト推定論」研究会予稿
［2］ Butler（1969） *Ann. Math. Statist.*
［3］ Khmaladze（1981） *Theory Prob. Appl.*
［4］ MacNeil（1974） *Ann. Statist.*
［5］ Rothman and Woodroofe（1972） *Ann. Math. Statist.*