る作業になり，一方限定されたモデル間の比較ではしばしば常識に反する結果を与える．

宮野は，リッジ回帰を適用した解決法を提案している．パラメータの0次階差（ノルム）を小さくするという点で漸進的変化の条件と類似しているが，リッジ係数を決定する基準については未解決である．この基準にABICを用いたとしても3効果に対する重みが同一なことから，3効果が同程度に現れるような解に限定される．

石井は，不定解の中からパラメータの1次階差が最小になる解を求めようとしている．このやり方も，3効果に対する重みを決定する基準をもっていない．

丹後は，メッシュコウホートという概念を持ち出し，データのセルを分割することによって識別問題が解決できるとしている．しかし，この方法は，暗黙裡に3効果のパラメータに1次階差の制約を導入したものと同等である．また，その重みが固定されており，識別問題を解決しているとはいえない．

附属統計技術員養成所

分割表に対する混合モデルと潜在構造分析

鈴 木 義 一 郎

一般に，2×2分割表は，3つのパラメータを用いて

$pq+d$	$p(1-q)-d$	p
$(1-p)\,q-d$	$(1-p)\,(1-q)+d$	$1-p$
q	$1-q$	1

のように表わされることが分かる．これを $M\{p, q\,;\,d\}$ と表わす．この d というパラメータは，対角要素の積の差であり，同時に独立モデル $M\{p, q\,;\,0\}$ からの偏差を表し

$$-b \leqq d \leqq a$$

の範囲で変動する．ここで

$$a=\min\{p(1-q), (1-p)q\}$$
$$b=\min\{pq, (1-p)(1-q)\}$$

下の表に，パラメータ数が1つ（網目の部分）または2つの特殊モデルで，$d \geqq 0$ のケースを示す．

次に，

$$M\{p, q\,;\,d\}, M\{\tilde{p}, \tilde{q}\,;\,\tilde{d}\}$$

という2つのモデルが，それぞれ α, $1-\alpha$ という比率で混合されたモデルは

$$M\{\bar{p}, \bar{q}\,;\,\bar{d}+e\}$$

	Independent Model	Dependent Model	The Most Dependent Model
	$I(p,q)=M\{p,q\,;\,0\}$	$M\{p, q\,;\,d\}$ three parameter model	$U^+(p, q)=M\{p, q\,;\,p(1-q)\}$ $L^+(p, q)=M\{p, q\,;\,(1-p)q\}$
対 称 $(q=p)$	$I_0(p)=M\{p, p\}$	$S^+(p, d)=M\{p, p\,;\,d\}$ $(0 \leqq d \leqq p(1-p))$	$D^+(p)=M\{p, p\,;\,p(1-p)\}$ positive diagonal model
歪対称 $(q=1-p)$	$I_0^*(p)=M\{p, 1-p\}$	$T\,(p, d)=M\{p, 1-p\,;\,d\}$ $(0 \leqq d \leqq p^2 \wedge (1-p^2))$	$U^+(p, 1-p)\ (p \leqq 1/2)$ $L^+(p, 1-p)\ (p \geqq 1/2)$

のように表わされる．ここで

$$\bar{p}=\alpha p+(1-\alpha)\tilde{p}$$
$$\bar{q}=\alpha q+(1-\alpha)\tilde{q}$$
$$\bar{d}=\alpha d+(1-\alpha)\tilde{d}$$
$$e=\alpha(1-\alpha)(p-\tilde{p})(q-\tilde{q})$$

例えば，大竜巻のような天災が起きるか否かの予想と，実際に起きたか否かにより結果を分類した表に，完璧な予測に対応する対角モデル $D^{+}(p)$ と，当てずっぽな予測に対応する独立モデル $I(p, q)$ との混合モデルを対峙させると，α はその予想のパターンの“的中率”を示す．

また，ある期間に結婚した夫婦の前歴で，夫が初婚か再婚か，妻が初婚か再婚かで分類した表には，三角モデル $U^{+}(p, q)$ と独立モデル $I(p, q)$ 又は三角モデル $S\left(\frac{p+q}{2}, p-\left(\frac{p+q}{2}\right)^{2}\right)$ との混合モデルを対応させると，α は初婚の夫が再婚の妻との結婚を忌避する尺度を示すものと考えられる．

この他，各種の混合モデルを対峙させて，分割表に潜在している「構造」に探りを入れることができる．

ベイズ型密度関数推定モデルについて

坂　元　慶　行

滑かな密度関数をもつ限り適用可能な密度関数の推定法を［1］で提案した．この方法は，赤池の GALTHY の考え方に基礎を置き，具体的には次の手順によっている．

① データ x を，適当な分布関数 $Q(x)$ を用いて，$y=Q(x)$ に変換
② 区間［0, 1］上のヒストグラムのベイズ推定値 $\hat{h}(y)$ を求める
③ $\hat{h}(y)$ を2次スプラインで平滑化し，$\tilde{h}(y)$ を求める
④ x の密度関数の推定値として $f(x)=\tilde{h}(Q(x))q(x)$ をとる

この方法においては，② で，尤度 $L(\boldsymbol{h})$ の $\boldsymbol{h}$ に対して事前分布 $\pi(\boldsymbol{h} \mid u^{2}, h_{-1}, h_{0})$ を仮定し，ハイパー・パラメータを，

$$\mathrm{ABIC}=(-2)\log\int L(\boldsymbol{h})\pi(\boldsymbol{h} \mid u^{2}, h_{-1}, h_{0})d\boldsymbol{h}+2\ (\text{ハイパー・パラメータ数})$$

によって決定する．この方法は，従来順位統計量によって処理されることの多かった“(2標本)ノンパラメトリック検定”に対する新たな情報量統計学的アプローチの可能性をひらくと考えられる．すなわち，

$$g(y \mid u^{2}, h_{-1}, h_{0})=\int L(\boldsymbol{h})\pi(\boldsymbol{h} \mid u^{2}, h_{-1}, h_{0})d\boldsymbol{h}$$

が確率分布をなす場合，分布型未知の2つの母集団分布 g_1, g_2 から m, n 個の標本 $\{y_i^{(1)}\}$, $\{y_i^{(2)}\}$ がとられたとき，2つのモデル，

$$\text{MODEL}\,(0):\ y_i^{(j)}\sim g(y \mid u^{2}, h_{-1}, h_{0}) \qquad j=1, 2$$
$$\text{MODEL}\,(1):\ y_i^{(j)}\sim g(y \mid u_{(j)}^{2}, h_{-1}^{(j)}, h_{0}^{(j)}), \qquad j=1, 2$$

を想定し，ABIC の最小値を与えるモデルを採択することによって実現できると考えられる．

いくつかの実験結果から，正答率には遜色ないことがわかった．しかし，この方法の特長は分布型そのものの良好な推定が可能な点にある．

［1］ Ishiguro, M. and Sakamoto, Y. (1984). A Bayesian Approach to the Probability Density Estimation, *Ann. Inst. Statist. Math.*, 36, 3, 523-538.
［2］ 坂元慶行 (1985). カテゴリカルデータのモデル分析，共立出版．(印刷中)