（３）情報量と統計的推測理論

久保木久孝（電通大・電子情報）：Construction of inferential distributions for predictive fit ………… 93
石黒真木夫（統計数理研）：情報量規準とブートストラップ ………… 95
坂口実（名古屋商科大）：Entropy analysis of brand purchase behavior－Sudden change of market structure ………… 97
富澤貞男（東京理科大・理工）：２次元分割表における非一様連関性を測るShannonエントロピー型尺度 ………… 99
川端勉（電通大・電子情報）：自己相似測度とレートひずみ次元 ………… 101
山田作太郎（東京水産大）：実験の比較の理論の紹介 ………… 103
草間時武（早稲田大・理工）：統計的実験の比較（Selection Modelの場合）… 105
鈴木武（早稲田大・理工）：実験の比較についてー制限された決定問題における比較 ………… 107
吉田裕亮（統計数理研）：非可換解析と相対エントロピー ………… 109
堀部安一（静岡大・工）：木と情報とフィボナッチ数：Balance properties of optimal binary trees ………… 111
有田清三郎（川崎医大・数学）：MCQテストにおけるファジィ情報と得点分布 ………… 112
江口真透（島根大・理）：Projection method for statistical inference in nonlinear model ………… 114
鎌倉稔成（中央大・理工）：ひずみを持つ分布の平均のロバスト推定について ………… 116
柳本武美（統計数理研）：推定方程式の理論からみたモーメント法 ………… 118

Construction of inferential distributions for predictive fit

電通大電子情報　久保木 久孝

1. はじめに

独立な n 個の観測 $s_n = (x_1, \ldots, x_n)$ を使い，未来の観測値 y について言及したい．観測値 x_i と y は独立であるが，それらの同時分布 $g(x_1)\cdots g(x_n)f(y)$ は母数モデル $\{g(x_1|\theta)\cdots g(x_n|\theta)f(y|\theta) : \theta \in \Theta\}$ の元であるとする．そのとき，通常 y の予測分布 $f(y|s_n)$ は次の方法で与えられる：観測 s_n に基づく分布 $p(\theta|s_n)$，または事前分布 $p(\theta)$ と s_n に基づく事後分布 $p(\theta|s_n)$ を使い，

$$f(y|s_n) = \int f(y|\theta)p(\theta|s_n)\,d\theta.$$

このような目的で使う θの条件付き分布を Akaike (1978, *Biometrika*, 53–59) は推測分布 (inferential distribution) と呼んでいる．予測の悪さの評価には（負の）期待エントロピー

$$E_{s_n|\theta}[I\{f(\cdot|\theta), f(\cdot|s_n)\}] = E_{s_n|\theta}\left[\int f(y|\theta)\log\left\{\frac{f(y|\theta)}{f(y|s_n)}\right\}dy\right]$$

が用いられる．

さて，θ に事前分布を想定したくない又は想定することができないとき，客観的な推測分布 $p(\theta|s_n)$ をいかに構成するかという問題を考える．最近ふたつの方法—ブートストラップ法 (Harris, 1989, *Biometrika*, 675–684) およびコンディショニング法 (El-Sayyad et al., 1989, *Biometrika*, 343–348) —が提案されているが，それらにはまだ議論の余地がある．本稿の目的は，推測分布の新しい構成法を示すことである．その客観性の基となるのが Lauritzen-Hinkley の予測尤度の概念である．特に s_n のしたがうモデルが指数型であるとき，この方法は Jeffreys の prior を利用する方法と密接に関連することも示される．

2. 客観的な構成法

仮想的な追加実験を行い観測値 $t_m = (x_{n+1}, \ldots, x_{n+m})$ が得られたとする．これに基づく θ の最尤推定を $\hat{\vartheta}_m$ とする．実際の観測値 s_n を利用して $\hat{\vartheta}_m$ の動きを予測することを考える．もしこのための分布 $p^{L_m}(\hat{\vartheta}_m|s_n)$ が構成できれば，θ にその推定値 $\hat{\vartheta}_m$ を代入して作った y の予測分布 $f(y|\hat{\vartheta}_m)$ をそれで平均することで予測分布 $f^{L_m}(y|s_n)$ が構成できる．すなわち

$$f^{L_m}(y|s_n) = \int f(y|\theta)p^{L_m}(\theta|s_n)\,d\theta.$$

いま，u_n, v_m および w_{n+m} をそれぞれ s_n, t_m および (s_n, t_m) の最小十分な縮約であるとする．このとき，Hinkley (1979, *Ann. Statist*, 718–728) は u_n を与えたと

きの v_m の予測尤度 (Lauritzen-Hinkley の予測尤度) を

$$\mathrm{lik}^*(v_m|u_n) = h(u_n|w_{n+m})$$

と定式化した．ここで $h(u_n|w_{n+m})$ は w_{n+m} を与えたときの u_n の条件付き分布である．当然 $\hat{\vartheta}_m = \hat{\vartheta}(v_m)$ と表現できるから，u_n を与えたときの $\hat{\vartheta}_m$ の予測尤度 $l^*(\hat{\vartheta}_m|u_n)$ を次の関数の密度関数（存在を仮定して）で定義することにする：

$$L^*(d\hat{\vartheta}_m|u_n) = \int_{\{v_m : \hat{\vartheta}(v_m) \in d\hat{\vartheta}_m\}} \mathrm{lik}\left(\hat{\vartheta}(v_m)|v_m\right) \mathrm{lik}^*(v_m|u_n)\, dv_m.$$

ここで $\mathrm{lik}(\theta|v_m)$ は θ の尤度である．いま，予測尤度 $l^*(\hat{\vartheta}_m|u_n)$ を将来得られるであろう $\hat{\vartheta}_m$ の値にオーダーを付ける関数と考え，推測分布 $p^{L_m}(\theta|s_n)$ を

$$p^{L_m}(\theta|s_n) \propto l^*(\theta|u_n)$$

で構成することを提案する．

3. 指数型モデルにおける結果

いま，$g(x|\theta)$ は指数型であるとする：

$$g(x|\theta) = a(\theta)b(x)\exp\{\theta' T(x)\}, \qquad \theta \in \Theta \subset R^k.$$

このとき，$u_n = T(x_1) + \ldots + T(x_n)$ の密度は

$$g^{n*}(u_n|\theta) = a(\theta)^n b^{r*}(u_n)\exp(\theta' u_n)$$

と表現される．ここで，T の分散共分散行列を Σ_θ とすると，m が十分大きいとき

$$\mathrm{lik}(\hat{\vartheta}_m|v_m)\,\mathrm{lik}^*(v_m|u_n) \simeq (2\pi m)^{-\frac{k}{2}}|\Sigma_{\hat{\vartheta}_m}|^{-\frac{1}{2}} g^{n*}(u_n|\hat{\vartheta}_m)$$

という近似を得る．これを利用すると，$l^*(\hat{\vartheta}_m|u_n)$ の定義から次の形の推測分布 $p^L(\theta|s_n)$ を得る：

$$p^L(\theta|s_n) \propto l^*(\theta|u_n) \propto g^{n*}(u_n|\theta)|i(\theta)|^{\frac{1}{2}}.$$

ここで，$i(\theta)$ は $T(x)$ の Fisher 情報量行列である．すなわち，$p^L(\theta|s_n)$ は形式的には θ に Jeffreys の prior を仮定したときの事後分布と一致する．

情報量規準とブートストラップ

石黒真木夫
統計数理研究所

1990年11月15日

ABSTRACT

赤池 (Akaike ,1973) はデータに当てはめたモデルの良さを評価する情報量規準の推定量として AIC を提案した。AIC の摘要範囲は非常に広く、最尤法によるパラメータの推定が有効な場合にはいつでも使えると言ってよい。ここで「有効」と言っているのは、最尤推定量の誤差評価が有効という意味であり、必ずしも最尤推定量の誤差が小さくなくてもいいことが AIC の便利な点である。

最尤推定が使われていない場合には AIC は使えない。AIC を導くにあたって対数尤度関数と情報量規準の関係、MLE の漸近正規性が使われているためである。

この論文では、推定値の誤差評価にブートストラップ法を用いることによって最尤法以外の方法でパラメータ推定がなされている場合にもその情報量の意味での評価ができることを示し、ＡＩＣが用いられていた分野はもちろん、従来ＡＢＩＣによる評価が用いられていた問題にも適用できることを示した。

1 WIC

Information Criterion, WIC is defined by

$$\begin{aligned} WIC = & - 2 \times \log f(x|\hat{\theta}) \\ & + 2 \times \text{“Bias correction”} \end{aligned} \tag{1}$$

where the “Bias correction” is defined by

$$\begin{aligned} & \text{“Bias correction”} \\ & = E_{x^*} \left\{ \log f(x^*|\hat{\theta}^*) - \log f(x|\hat{\theta}^*) \right\} \end{aligned}$$

where x^* is a simulated data obtained by the resampling technique. $\hat{\theta}^*$ is the estimate of parameter based on the data x^*. E_{x^*} denotes the expectation with respect to the variation of x^*, which is practically computed by the Monte Carlo method.

WIC can be computed even if the parameter estimation is not done by the maximum likelihood method. When the parameter estimation is done by the maximum likelihood method, and the number of data is large compared to the number of parameters to be estimated, the "Bias correction" coinsides with the number of parameters estimated.

2 数値例

*WIC*を分割表モデルの選択、ＡＲモデルの次数選択に適用した場合に*AIC*と同等な挙動を示すことを示し、次いで、ｐｅｎａｌｉｚｅｄ　ｌｅａｓｔ　ｓｑｕｒｅｓによる回帰曲線の推定と多項式回帰の評価、比較が情報量の意味で可能である事も示した。最後に 電波望遠鏡データ解析で用いられるＣＬＥＡＮとして知られているデータ解析法の制御にも応用できる事を示した。

参考文献

[1] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, *2nd Inter. Symp. on Information Theory* (Petyov, B.N. and Csai, F. eds.), Akademiai Kiado, Budapest, pp.267-281.

[2] Efron, B., (1979). Bootstrap Methods: Another Look at the Jackknife, *The Annals of Statistics*, 7, 1, pp.1-26.

[3] Efron, B., (1986). How biased is the apparent error rate of a prediction rule? *JASA*, 81, 394, pp.461-470.

[4] Sakamoto, Y., Ishiguro, M. & Kitagawa, G. (1986). *Akaike Information Criterion Statistics*, D.Reidel Publishing Company, Dordrecht/Tokyo.

[5] Shibata, R.(1989). Statistical Aspect of Model Selection, in *From Data to Model*, Ed. J. C. Willems, pp.215-240, Springer-Verlag.

[6] Wong, W. H.(1983) A note on the modified likelihood for density estimation *JASA*, 78, No.382, pp.461-463.

Entropy Analysis of Brand Purchase Behavior
―― Sudden Change of Market Structure

坂口　実　（名古屋商大）

§1. Entry of a New Brand（文献[1]）

Herniterの意味(1973)で均衡している 2-brand market の中に Br.3 が参入するとき新しい market share はどうなるか？　直観的予想は立て難い。

参入直後の市場構造を自然と思われる

［想定A_1］新Br.の参入に無関心な客層の中で既存brand の間の market share は参入の前後で不変である

のもとで考え

参入直後の全system entropy ⟶ max
subject to
(a) $\{p_{ij}\}$ についての規格条件,
(b) $\phi_i(\cdot)$ etcについての規格条件,
(c) $[A_1]$ のための条件

をやる。例えば $m=\langle 0.6, 0.4\rangle$ のときに、Br.3 が参入したら new share は $m_c' = \langle 0.331, 0.231, 0.438\rangle$ から $m_l' = \langle 0.493, 0.342, 0.165\rangle$ の間であると考えられる。

§2. Exit from the Existing Three-Brand Market（文献[2]）

Herniterの意味で均衡している 3-brand market から Br.3 が出てゆくとき、残された 2-brand market の share はどう変わるか？　直観的予想では share が $m=\langle m_1, m_2, m_3\rangle$ のとき Br.3 が退場した後には

$$m^{\circ}=\langle m_1, m_2\rangle + m_3\cdot\langle m_1/(m_1+m_2), m_2/(m_1+m_2)\rangle = \langle m_1/(m_1+m_2), m_2/(m_1+m_2)\rangle$$

であろうが、これは本当にそうだろうか？

退場直後の市場構造を、自然と思われる

［想定A_2］Br.3 に関心のあった客層の中で、その退場後の Br.1, 2 への移行率は $m_1:m_2$ に等しい

のもとで考え、前と同様の MEP analysis をやる。例えば $m=\langle 0.5, 0.3, 0.2\rangle$ のとき new share は $m'=\langle 0.630, 0.370\rangle$ になるが、直観では $m^{\circ}=\langle 0.625, 0.375\rangle$ である。

§3. Introduction of the 2nd choice（文献[3]）

Herniterの意味で均衡している 3-brand market において新たに 2nd choice を（重複なしに）選択する場合、ordered-choice-of-two はどうなるか？　直観的予想では share が $m=\langle m_1, m_2, m_3\rangle$ のときは $m^{\circ}_{ij}=m_i m_j/(1-m_i)\ (i=1,2,3; i\neq j)$ であろうが、本当にそうだろうか？

市場構造と、当然と思われる

[想定A_3] 2nd choice を選択する前後で、各 brand の 1stly chosen brand の share は $m_1:m_2:m_3$ である

のもとで前と同様の MEP analysis をやる。この結果、例えば $m=\langle 0.5, 0.3, 0.2\rangle$ のとき new share は

$$m' = \langle 0.268, 0.232, 0.180, 0.118, 0.114, 0.085\rangle,$$
$$\text{for } (1,2), (1,3), (2,1), (2,3), (3,1), (3,2) \text{ pair}$$

したがって、各 brand の 2nd choice としての share は $\langle 0.294, 0.353, 0.350\rangle$ である。これらは直観的予想では $m^\circ = \langle 0.300, 0.200, 0.214, 0.086, 0.125, 0.075\rangle$ と $\langle 0.339, 0.375, 0.286\rangle$ である。

§4. Final Remark.

§1 (§2) では新Br.の参入後 (3rd Br.の退場後) 新しい3-brand market (2-brand market) として均衡に至る。これらは一般に m' とは異なるものである。

[文献]

[0] J.D.Herniter, J. Marbeting Res. 10(1973), 361-375.

[1] M.Sakaguchi, Math. Japonica, 32(1987), 475-492.

[2] M.Sakaguchi, Math. Japonica, 33(1988), 943-955.

[3] M.Sakaguchi, Math. Japonica, 35(1990), 171-186(with A.Doi).

2次元分割表における非一様連関性を測る Shannonエントロピー型尺度

富澤貞男　東京理科大・理工

順序カテゴリーの$R\times C$分割表において，(i,j)セル確率をp_{ij}とする．行と列変数の独立性(I)モデルの拡張であるGoodman (1979)の一様連関(U)モデルは，

$$p_{ij}=\alpha_i\beta_j\theta^{ij}\quad(i=1,2,\cdots,R;\ j=1,2,\cdots,C),$$

のように与えられる．$\theta=1$とおいた特別な場合がIモデルである．今，$R\times C$表において隣接した行(iと$i+1$)と列(jと$j+1$)からなる2×2表に対する局所オッズ比をΘ_{ij}と略記する．即ち，$\Theta_{ij}=(p_{ij}p_{i+1,j+1})/(p_{i+1,j}p_{i,j+1})$．Uモデルは，

$$\Theta_{ij}=\theta\quad(i=1,2,\cdots,R-1;\ j=1,2,\cdots,C-1),$$

のようにも表わせる．ところで，Iモデルからの隔たりの程度を測るための尺度は，数多く提案されている．

本講演の目的は，Uモデルからの隔たりの程度を測るための尺度を提案することである．提案される尺度は，Iモデルが成り立たないときに利用され，いくつかの分割表においてUモデルからの隔たりの程度を比較するのに役立つ．今，次のような尺度を導入する．

$$\phi_U = 1 - H(\Theta^*) / [\log (R-1)(C-1)],$$

ただし，$H(\Theta^*) = -\sum_{i=1}^{R-1}\sum_{j=1}^{C-1} \Theta_{ij}^* \log \Theta_{ij}^*$，$\Theta_{ij}^* = \Theta_{ij}/\nu$，$\nu = \sum_{i=1}^{R-1}\sum_{j=1}^{C-1} \Theta_{ij}$.

ここに $H(\Theta^*)$ は $\{\Theta_{ij}^*\}$ に基づくShannonエントロピーである。

また，Kullback-Leibler情報量を用いて ϕ_U は次のようにも表わせる。

$$\phi_U = I(\Theta^*; \Theta^U) / [\log (R-1)(C-1)],$$

ただし，$I(\Theta^*; \Theta^U) = \sum_{i=1}^{R-1}\sum_{j=1}^{C-1} \Theta_{ij}^* \log(\Theta_{ij}^* / \Theta_{ij}^U)$，$\Theta_{ij}^U = 1/(R-1)(C-1)$.

今，$0 \le \phi_U < 1$ であり，Uモデルが成り立つための必要十分条件は $\phi_U = 0$ である。ここで，(i, j) セル観測度数を n_{ij} とし，$\hat{p}_{ij} = n_{ij}/n$，$n = \Sigma\Sigma n_{ij}$ とし，対応する ϕ_U の推定値を $\hat{\phi}_U$ とすると，delta method を使って，$\sqrt{n}(\hat{\phi}_U - \phi_U)$ は平均0，分散 σ^2 の正規分布に漸近的に従う。ここに，

$$\sigma^2 = \sum_{i=1}^{R}\sum_{j=1}^{C} (L_{ij}^2 / P_{ij}) / [\nu \log (R-1)(C-1)]^2,$$

ただし，

$$L_{ij} = (\Theta_{i-1,j-1} - \Theta_{i-1,j} - \Theta_{i,j-1} + \Theta_{ij})\delta/\nu + (\Theta_{i-1,j-1} \log \Theta_{i-1,j-1}$$
$$- \Theta_{i-1,j} \log \Theta_{i-1,j} - \Theta_{i,j-1} \log \Theta_{i,j-1} + \Theta_{ij} \log \Theta_{ij}), \qquad 0 \log 0 = 0,$$
$$\delta = -\sum_{s=1}^{R-1}\sum_{t=1}^{C-1} \Theta_{st} \log \Theta_{st}, \quad \Theta_{s0} = \Theta_{0t} = \Theta_{Rt} = \Theta_{sC} = 0 \left(\begin{matrix} s = 0, 1, \cdots, R \\ t = 0, 1, \cdots, C \end{matrix}\right).$$

これより，ϕ_U の信頼区間が作れ，それに基づいて，いくつかの分割表におけるUモデルからの隔たりの程度が比較できる。

参考文献：Tomizawa(1991). Statist. & Prob. Letters, to appear.
Goodman, L. A. (1979). J. Amer. Statist. Assoc., 74, 537-552.

自己相似測度とレートひずみ次元

電気通信大学　川端　勉

自己相似集合とはユークリッド空間 (R^k) 内の図形であって、自分の相似図形の直和で表せるものをいう。このような図形はハウスドルフ次元がしばしば非整数となることがあるので、フラクタル図形の典型と見なされている。この図形上に定義された測度のレートひずみ関数を考える。この関数は、情報源の入力系列をを符号化したときに、復号の後に生ずるひずみを、平均として一定レベル D 以下に押さえられるために必要な、一入力あたりの符号化ビットレートである。ひずみは、ユークリッド距離の冪乗 (r 乗) とする。レートひずみ関数は、情報源が測度 (μ) の独立同一分布に従っているときに次のような定式化がなされている：

$$R(D) = \inf_{\nu : E[d(X,Y)] \leq D} I(\mu, \nu) \, ,$$

ただし、(X, Y) は同時分布 $\mu\nu$ に従うものとする。直線や平面上に滑らかに定義された確率分布に対してレートひずみ関数を評価することは可能であるが、特に、ひずみが小さいときには、$R(D) \sim -\frac{\rho}{r} \log D$ の様に振る舞うことが証明できる。ここで、ρ は分布の台のハウスドルフ次元である。本研究では、この事実が自己相似集合を台とする確率分布にも適用できることを示した。

さて、$S = \cup_{i=1}^{m} W_i S$ (ただし W_i は S の相似変換) としよう。このとき $\forall x \in S$ はある部分集合 $W_i S$ に属するので、$\forall t > 0$ に対し $x \in W_{b_1} W_{b_2} \ldots W_{b_t} S$ なる系列 $\mathbf{b} = b_1 b_2 \ldots b_t \ldots$ が定まる。これにより、S 上の測度 μ は $M = \{1, \ldots, m\}$ 上の測度 μ^* に変換される。μ^* が定常エルゴート的ならば、

$$\rho(\mu) \stackrel{def}{=} \frac{H_\infty(\mu^*)}{E_{\mu^*}[-\log \alpha_{b_1}]}$$

という量を定義することができる。分子はエントロピーレート、分母は平均対数縮小率である。W_i のスケールを α_i とする。$\forall x \in S$ に対して、$inf\{n : \prod_{t=1}^{n} \alpha_t < \delta\}$ は $W_{b_1} W_{b_2} \ldots W_{b_n} S$ が S の δ 倍以下となる初めての n を表す。$D_r^*(1) \stackrel{def}{=} \min_y \max_{x \in S} d(x, y)$ $(< \infty)$ とし、最小値を与える y を y_{min} と書くことにする。$W_{b_n}^{-1} W_{b_{n-1}}^{-1} \ldots W_{b_1}^{-1} y_{min}$ を X_δ と表すと、これはほぼ一様な S の量子化スキームを与えている。X_δ のレートひずみ関数 $R_\delta(D)$ ともとの $R(D)$ の関係は次の定理により与えられる。

Theorem 1

$$R_\delta((D^{1/r}+\delta D_r^*(1)^{1/r})^r) \le R(D)$$
$$\le R_\delta((D^{1/r}-\delta D_r^*(1)^{1/r})^r) ,$$

ただし、$R_\delta(D) \stackrel{def}{=} \infty$ *for* $D<0$ と置いた。

X_δ がなす点の集合を Λ_δ と表す。このとき

$$\inf_{\delta>0} \min_{i\neq j} \left\{ \min_{x\in\Lambda_{\delta/\alpha_i},\, x'\in\Lambda_{\delta/\alpha_j}} d(W_i(x),W_j(x'))^{1/r} \right\} = \epsilon > 0 ,$$

が満たされるなら S はカントル的と呼ぶ。これに関して次の定理を得た。

Theorem 2 S をカントル的集合をすると、

$$\lim_{D\to 0}\left[-\frac{rR(D)}{\log D}\right] = \rho(\mu) .$$

さらに興味深い性質として、

Theorem 3

$$\sup_{\mu^*\, stationary} \left\{ \frac{H_\infty(\mu^*)}{E_{\mu^*}(-\log\alpha_{b_1})} \right\} = d^*,$$

ただし、d^* は S の *Hausdorff-Besicovitch* 次元すなわち、$\sum_{i=1}^m \alpha_i^{d^*} = 1$ なる唯一の根である。

を得た。

実験の比較の理論の紹介

東京水産大　山田作太郎

1. 定義

推測したいパラメーターは $\theta \in \Theta$ とする。この時２つの観測

$$X = (\mathcal{X}, \{P_\theta ; \theta \in \Theta\}), \quad Y = (\mathcal{Y}, \{Q_\theta ; \theta \in \Theta\})$$

のどちらからも推測できる（θ に関する情報を得ることができる）とする。どちらを観測したら得か？　どちらが θ の推測により多くの情報を含んでいるか？　Y の方が X よりよいとき、Y は X に対して十分であるという。次の定理の条件(1)又は(2)を形式的な定義として採用することが多い。

定理1. 次の２つの条件は同値である。

(1) ある既知の分布に従う確率変数 Z と、(Y, Z) の関数 h が存在して、もし Y が Q_θ に従えば、$h(Y, Z)$ は P_θ に従う。（又は、$P_\theta(A) = \int M(A, y)\, dQ_\theta$, $\forall \theta \in \Theta$ なる M がある。）

(2) パラメーター空間を Θ とする任意の決定問題と、観測 X に基づく任意の決定関数 δ に対し、Y に基づく決定関数 δ' がとれて、危険関数について

$$r(\delta', \theta) \leqq r(\delta, \theta) \quad \text{for all } \theta \in \Theta$$

が成立する。

(1)又は(2)が成立するとき、$Y \geqq X$ としるす。」

他に「各観測ごとに、その実験が含む"情報量"を計算し、それで比較する」という考え方もある。

定理2. Y が X に対して十分とする。この時、X で estimable な任意のパラメーター関数は Y でも estimable である。

（証明）$g(\theta)$ はパラメーター関数で、$t(x)$ が存在して

$$\int t(x)\, dP_\theta(x) = g(\theta), \quad \forall \theta \in \Theta \quad \text{とする。}$$

M は Y から X へのマルコフ核で $P_\theta = M \cdot Q_\theta$ $\forall \theta \in \Theta$ が成立するとする。この時 $g(\theta) = \int \left[\int t(x)\, M(dx, y) \right] dQ_\theta(y)$ より、$y \longrightarrow \int t(x)\, M(dx, y)$ は g の不偏推定量である。」

2. 例

(1) $X_1,\dots,X_n$ は独立に $N(\theta,\sigma_1^2)$ に、$Y_1,\dots,Y_n$ は独立に $N(\theta,\sigma_2^2)$ に従うとする。σ_1^2, σ_2^2 は既知で $\sigma_1^2>\sigma_2^2$ とする。θ は未知とする。この時 $Y=(Y_1,\dots,Y_n)$ は $X=(X_1,\dots,X_n)$ に対して十分である。$Z_1,\dots,Z_n$ は $Y_1,\dots,Y_n$ とそれぞれ独立に $N(0,\sigma_1^2-\sigma_2^2)$ に従うとすると、Y_i+Z_i は $N(\theta,\sigma_1^2)$ に従うからである。一方 X は Y に対して十分でない。θ のUMVUEの分散は、X については σ_1^2/n、Y については σ_2^2/n で、$\sigma_1^2/n>\sigma_2^2/n$ であるから(定理1(2))。

(2) $X_1,\dots,X_n$ は独立に $N(\theta,\sigma^2)$ に、$Y_1,\dots,Y_n$ は独立に $N(\theta,\rho^2\sigma^2)$ に従うとする。ρ は既知で $0<\rho<1$ とする。Case 1. $\theta=\theta_0$ が既知とする(一般性を失わず $\theta_0=0$ とする)。$Y_i/\rho\sim N(0,\sigma^2)\sim X_i$、かつ、$\rho X_i\sim N(0,\rho^2\sigma^2)\sim Y_i$ より X と Y は同値である。Case 2. θ,σ が共に未知とする。次のHansen-Torgersenの結果より、X と Y は比較不可能である($X\gtrsim Y$, $Y\gtrsim X$ とも成立しない)。しかしFisher情報行列の意味では Y は X に対して十分である。

「Hansen-Torgersenの定理:線形模型 $\mathcal{E}_A,\mathcal{E}_B$ を考える。

$\mathcal{E}_A$: $X=A'\beta+e$, $A'(n_A\times k)$ は既知。$e\sim N(0,\sigma^2 I)$

$\mathcal{E}_B$: $Y=B'\beta+e$, $B'(n_B\times k)$ は既知。$e\sim N(0,\sigma^2 I)$

σ^2 が既知のとき、$\Theta=(-\infty,\infty)^k$、未知のとき $\Theta=(-\infty,\infty)^k\times(0,\infty)$

このとき (ア) σ^2 が既知のとき、$\mathcal{E}_A\gtrsim\mathcal{E}_B\Leftrightarrow AA'\geq BB'$

(イ) σ^2 が未知のとき、$\mathcal{E}_A\gtrsim\mathcal{E}_B\Leftrightarrow AA'\geq BB'$、$n_A\geq n_B+\mathrm{rank}(AA'-BB')$ 」

(3) $\theta\in(0,1)\equiv\Theta$ とする。

$$p(x|\theta)={}_nC_x\,\theta^x(1-\theta)^{n-x},\quad X=(\{0,1,\dots,n\},\{p(x|\theta);\theta\in\Theta\})$$
$$g(y|\theta)=e^{-n\theta}\cdot(n\theta)^y\cdot(y!)^{-1},\quad Y=(\{0,1,\dots\},\{g(y|\theta);\theta\in\Theta\})$$
$$r(z|\theta)=(1-\theta)^z\cdot\theta,\quad Z=(\{0,1,\dots\},\{r(z|\theta)|\theta\in\Theta\})$$

と定義する。この時定理2を用いて X,Y,Z は比較不可能であることが示される。

3. 問題点

すべての決定問題に対して改良されねばならないという条件は強すぎる。最近Lehmannは決定問題のクラスを制限した実験の比較を考察している。

統計的実験の比較（Selection Modelの場合）

早大理工　草間時武

Xが密度関数$g(x,\theta)$をもつが、$x\in S$であるxしか、我々には観測できないならば、我々は

$$f(y,\theta)=\frac{g(y,\theta)}{P_\theta(X\in S)}\quad y\in S,\qquad =0\ (その他)$$

なる密度関数をもつ実験を対象にしていることになる．原実験（original model）と、この実験（selection model）を比較するBayarriとDeGrootの仕事を紹介する．Xにもとずく実験$\mathcal{E}_X$と，Yにもとずく実験$\mathcal{E}_Y$が同一の母数空間をもつとき、$\mathcal{E}_X$が$\mathcal{E}_Y$より，Blackwellの意味でmore informativeなら$\mathcal{E}_X\succsim\mathcal{E}_Y$とかくことにする．また$\mathcal{E}_X$が$\mathcal{E}_Y$よりpairwisely more informativeであるとき$\mathcal{E}_X\succsim_2\mathcal{E}_Y$とかく．Fisher情報量を比較する場合は$\mathcal{E}_X\succsim_F\mathcal{E}_Y$とかく．

$$\mathcal{E}_X\succsim\mathcal{E}_Y\ \Rightarrow\ \mathcal{E}_X\succsim_2\mathcal{E}_Y\ \Rightarrow\ \mathcal{E}_X\succsim_F\mathcal{E}_Y$$

がなりたつ．

まずXが$N(0,\theta)$にしたがう場合を考え、$(-\infty,-\tau_i]\cup[\tau_i,\infty)$に制限された観測を$Y_i$とし、$Y_i$にもとずく実験を$\mathcal{E}_i$とする．$0\leqq\tau_1<\tau_2<\infty$とすると$\mathcal{E}_2\succsim_2\mathcal{E}_1$である．したがって$\tau_1=0$とすれば、original model $\mathcal{E}$に対して$\mathcal{E}_i\succsim_2\mathcal{E}$である．これらの結果は$[\tau_i,\infty)$に制限しても同じである．$X$が$N(\theta,1)$にしたがう場合は結果は逆である．すなわち$-\infty<\tau_1<\tau_2<\infty$，$[\tau_i,\infty)$に制限した実験を$\mathcal{E}_i$とすると$\mathcal{E}_1\succsim_2\mathcal{E}_2$である．

これらの問題で$\succsim_2$を$\succsim$におきかえられるかどうかは未解決である．

次に2項分布とPoisson分布の実験を考えよう．これらの分布にもとずく実験を$\mathcal{E}_X$，0を除外した実験（zero truncated）を $\mathcal{E}_Y$ とすると $\mathcal{E}_X \gtrsim_2 \mathcal{E}_Y$ である。しかし $\mathcal{E}_X \gtrsim \mathcal{E}_Y$ はなりたたない。$\mathcal{E}_X$ や $\mathcal{E}_Y$ が dominated であっても、sufficiency の場合と違い、pairwisely more informative から more informative は云えないのである。しかし、一般に $\mathcal{E}_1 \gtrsim_2 \mathcal{E}_2$, $\mathcal{E}_2 \gtrsim_2 \mathcal{E}_1$ のとき $\mathcal{E}_1 \approx_2 \mathcal{E}_2$ とかき、$\mathcal{E}_1 \gtrsim \mathcal{E}_2$, $\mathcal{E}_2 \gtrsim \mathcal{E}_1$ のとき、$\mathcal{E}_1 \approx \mathcal{E}_2$ とかくと、$\mathcal{E}_1, \mathcal{E}_2$ が dominated のとき $\mathcal{E}_1 \approx_2 \mathcal{E}_2 \Longrightarrow \mathcal{E}_1 \approx \mathcal{E}_2$ がなりたつ。

Weighted function $w(y)$ をえらび $f(y,\theta) = w(y)g(y,\theta)/E_\theta[w(X)]$ を密度関数とする実験は selection model の拡張である。$w(y) = 1_S(y)$ とすれば、今まで述べてきた場合となる。いくつかの結果をあげると、

$\mathcal{E}_X$ を母数 θ の指数分布の実験、$\mathcal{E}_Y$ を weighted function $w(x) = e^{-ax}$ $(x>0)$ である実験とすると $\mathcal{E}_X \gtrsim \mathcal{E}_Y$.

$\mathcal{E}_X$ を $N(\theta, 1)$ の実験、$\mathcal{E}_Y$ を $w(x) = e^{ax}$ $(-\infty < a < \infty)$ による実験とすると $\mathcal{E}_X \simeq \mathcal{E}_Y$ である。

$\mathcal{E}_X$ を Poisson 分布の実験、$w(x) = a^x$ $(a>0,\ a \neq 1)$ とすると $\mathcal{E}_X \gtrsim \mathcal{E}_Y$ $(a<1)$, $\mathcal{E}_Y \gtrsim \mathcal{E}_X$ $(a>1)$ である。

もっとも、しばしば用いられる weighted function は、size-biased weighted function $w(x) = x$ である。この場合次のような結果がある。

$\mathcal{E}_X$ を $B(n,\theta)$ の実験（θ が未知母数）とし、$\mathcal{E}_Y$ を $w(x) = x$ とした実験（$x = 0, 1, \cdots, n$）とすると $\mathcal{E}_X \gtrsim \mathcal{E}_Y$.

$\mathcal{E}_X$ を $P_o(\theta)$ の実験、$\mathcal{E}_Y$ を $w(x) = x$ $(x = 0, 1, 2, \cdots)$ とする実験とすると $\mathcal{E}_X \simeq \mathcal{E}_Y$.

$\mathcal{E}_X$ を負の二項分布 $g(x,\theta) = \binom{r+x-1}{x}\theta^r(1-\theta)^x$ $(x = 0, 1, 2, \cdots)$ の実験（θ が未知母数）、$\mathcal{E}_Y$ を $w(x) = x$ $(x = 0, 1, 2, \cdots)$ である実験とすると $\mathcal{E}_Y \gtrsim \mathcal{E}_X$.

(1) M.J. Bayarri and M.H. DeGroot: Information in Selection Models. In Probability and Bayesian Statistics (R. Viertl, ed.), 39-51 New York: Plenum Press. (1987)

(2) M.J. Bayarri and M.H. DeGroot: Comparison of Experiments with Weighted Distributions, Statistical Data Analysis and Inference (Y. Dodge, ed) Elsevier Science Publishers B.V. (1989)

「実験の比較について—制限された決定問題における比較」

早稲田大学理工学部
鈴木 武

$\mathcal{E}=(\mathcal{X},\mathcal{A},\{P_\theta;\theta\in\Theta\})$と$\mathcal{F}=(\mathcal{Y},\mathcal{B},\{Q_\theta;\theta\in\Theta\})$を2つの与えられた統計的実験とする。Blackwell (1951, 1953) あるいは Le Cam (1964; 1986) によれば次の条件がみたされるとき，$\mathcal{E}$は$\mathcal{F}$より"more informative"($\mathcal{E}\geqq\mathcal{F}$)であるといわれる：任意の決定空間$(T,\mathcal{T})$，任意の有界損失関数$L_\theta(t)$，$\mathcal{F}$にもとづく任意の危険関数$\delta(\theta)$に対して，$\mathcal{E}$にもとづく危険関数$\gamma(\theta)$が存在して，$\gamma(\theta)\leqq\delta(\theta)$, $^\forall\theta\in\Theta$がなりたつ。

この定義における条件はきわめて強く，与えられた問題によっては常識に反する結果がおこることが Lehmann (1988) により指摘された。そこで，この不自然さを解消する一つの試みとして，単調尤度比をもつ分布族の比較においては，対象とする決定問題を，いわゆる単

調決定問題に制限したうえで危険関数の値を比較することがLehmann(1988)により提案された。この新しい意味で実験$\mathcal{E}$が$\mathcal{F}$より"more informative"のとき，Lehmann(1988)は，$\mathcal{E}$は$\mathcal{F}$より"more effective"であるとよんだ。この定義にしたがえば上記のような事態はおこらない。実際，$\mathcal{E}$,$\mathcal{F}$をそれぞれ単調尤度比をもつ実験で，各θにつきQ_θはnon-atomicとしたとき，次の結果がえられる。

定理(Lehmann(1988)) $\mathcal{F}$が$\mathcal{E}$より"more effective"であるための必要十分条件は，各xを固定したとき，$G_\theta^{-1}[F_\theta(x)]$が$\theta$の関数として非減少関数となることである。ここに$G_\theta$, F_θはそれぞれQ_θ, P_θの分布関数をあらわす。

さらに"more effective"なる概念を一般化して，$\varepsilon>0$を与えられた定数としたとき，それぞれの実験に対応する危険関数の値をεの範囲内で一方が他方を改良できるという意味で，いわば"ε-more effective"なる概念も考えられ，これについての議論は今後の課題である。

非可換解析と相対エントロピー

統計数理研究所　　吉　田　裕　亮

有限型フォン・ノイマン環と条件付き期待値

この講演では、主としてフォン・ノイマン環における相対エントロピーの話を行なった。一般のフォン・ノイマン環は量子化された測度空間に対応し、特に有限型と呼ばれるフォン・ノイマン環は量子化された確率空間に対応している。

まず，はじめに有限型フォン・ノイマン環とは何であるか，確率空間の非可換への拡張といわれるのは何故かを概説した.

ヒルベルト空間からそれ自身への線形写像を作用素という．ヒルベルト空間 $\mathfrak{H}$ 上の有界線形作用素の全体を $B(\mathfrak{H})$ で表す．$B(\mathfrak{H})$ の部分 $*-$ 代数のうち，弱位相で閉じていて，恒等作用素を含むものをフォン・ノイマン環という.

また可換フォン・ノイマン環の場合には，ある局所コンパクト Hausdorff 空間 Γ とそれ上の正値ラドン測度 μ が存在し，測度空間 (Γ,μ) 上の関数環 $L^\infty(\Gamma,\mu)$ と同型になる．これより，可換フォン・ノイマン環の研究は測度空間の研究にそれぞれ対応することがわかる.

フォン・ノイマン環 $\mathcal{M}$ が有限型のフォン・ノイマン環であるとは，$\mathcal{M}$ 上に線形汎関数 τ が存在し，

$$\tau(h) \geq 0 \quad (h \in \mathcal{M}^+), \quad \tau(1) = 1$$
$$\tau(x^*x) = \tau(xx^*)$$

を満たすときをいう．また，このような線形汎関数 τ をトレースという．有限型と呼ばれる所以は恒等作用素 1 のトレースの値が有限値（通常は 1 に規格化しておく）となるところにある.

確率空間を $(\Omega,\mathfrak{B},\mu)$ とし，その上の $L^\infty(\Omega,\mathfrak{B},\mu)$ を考える．これには各点毎の和，積，スカラー倍及び複素共役をとる $*-$ 演算で $*-$ 代数の構造が入り，この関数環はヒルベルト空間 $L^2(\Omega,\mathfrak{B},\mu)$ 上の可換なフォン・ノイマン環と見れる．さらに

$$\tau(f) = \int_\Omega f(\omega)d\mu(\omega)$$

とおくと τ は有限な忠実，正規トレースとなる．Ω 上の μ による積分は τ による値の評価と一致する.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

$\mathcal{M}$ を有限型フォン・ノイマン環とし，その上の忠実，正規トレースを τ とする．ただし，τ は $\tau(1)=1$ と規格化しておく．$\mathcal{N}$ を $\mathcal{M}$ の部分フォン・ノイマン環とする．このとき，任意の $x\in\mathcal{M}$ と任意の $y\in\mathcal{N}$ に対して，条件

$$\tau(E(x)y)=\tau(xy)$$

を満たす $\mathcal{M}$ から $\mathcal{N}$ への線形写像 E がトレース τ に依存して，一意に定まる．この線形写像 E を $\mathcal{M}$ から $\mathcal{N}$ への τ− 不変な条件付き期待値といい，$E_{\mathcal{N}}^{\mathcal{M}}$ と表す．

相対エントロピー

次に相対エントロピーの定義を紹介し、今までに得られたそれらの持つ性質や計算に有効な公式の紹介を行なった．$\mathcal{M}$ を有限型フォン・ノイマン環で，その忠実，正規，規格化されたトレースを τ とする．また，$\mathcal{N}_1,\mathcal{N}_2$ を $\mathcal{M}$ の部分フォン・ノイマン環とする．

$E_{\mathcal{N}_i}$ でそれぞれ $\mathcal{M}$ から $\mathcal{N}_i$ $(i=1,2)$ への τ-不変な条件付き期待値を表すものとする．また，$S(\mathcal{M})$ でもって $\mathcal{M}$ の単位の有限分割全体の集合を表すものとする．

このとき $\mathcal{N}_1$ の $\mathcal{N}_2$ に関する相対エントロピー $H(\mathcal{N}_1|\mathcal{N}_2)$ は

$$H(\mathcal{N}_1|\mathcal{N}_2)=\sup_{\Delta\in S(\mathcal{M})}\sum_i(\tau\eta E_{\mathcal{N}_2}(x_i)-\tau\eta E_{\mathcal{N}_1}(x_i))$$

と定義される．

ここで η は $\eta(t)=-t\log t\ (t>0), \eta(0)=0$ で定義される $[0,\infty]$ 上の連続関数である．ところで，$\mathcal{M}$ が可換の場合には，$\mathcal{N}_1,\mathcal{N}_2$ は，ある $\mathcal{M}$ の単位の分割 P_1,P_2 から生成される $\mathcal{M}$ の部分環であって，相対エントロピー $H(\mathcal{N}_1|\mathcal{N}_2)$ は，古典的な相対エントロピー $h(P_1|P_2)$ と一致する．このような意味で，いま定義された相対エントロピーは確率論でいうところの相対エントロピーの拡張になっている．

また、この相対エントロピーは 2 つの環の距離

$$\delta(\mathcal{N}_1,\mathcal{N}_2)=\sup\{\|x-E_{\mathcal{N}_2}(x)\|_2;x\in\mathcal{N}_1,\|x\|\leq 1\}$$

と両立する量でもある．

講演では相対エントロピーの評価を行なうためには条件付き期待値の分解が重要であることも報告した．

木と情報とフィボナッチ数

: Balance Properties of Optimal Binary Trees

静岡大学工学部　堀部安一

We consider a binary tree with leaves weighted by probabilities. An optimal tree (e.g., the well-known Huffman tree) is a tree that has minimum average path length. Also interesting is a binary tree which is constructed, in a top-down way, by successively dichotomizing the sequence of the leaf-weights as equally as possible. This weight-balanced tree is known to be nearly optimal. The inequality relation: $w_{k-1} \geqslant w_k + w_{k+1}$ is shown to hold commonly in both types of trees, where w_k is the weight of an arbitrary internal node at level k, and w_{k-1} and w_{k+1} are its father's and son's weights. Based on this relation, we show some properties of those trees, all related to the concept of weight-balancing and all deeply connected with the Fibonacci numbers and the Golden section.

Let $a:(1-a)$ be called α-balanced if $a \in [1-\alpha, \alpha]$ $(0<a<1,\ 1/2 \leqslant \alpha \leqslant 1)$. We first show that, if $(w_k/w_{k-1}):(1-w_k/w_{k-1})$ is out of ψ-balance such that $w_k/w_{k-1} > \psi$, then $(w_{k+1}/w_k):(1-w_{k+1}/w_k)$ is ψ-balanced (i.e., the ψ-balance is "immediately recovered" at the next level node), here $= (\sqrt{5}-1)/2$. The number ψ is then seen to be the critical value for such an immediate restoration of the balance.

We next show $1/3 \leqslant w_k/w_{k-1} \leqslant 2/3$ (best possible bounds), where the left inequality holds in the case that node $\bar{w}_k$ is also internal. From the right inequality $w_k/w_{k-1} \leqslant F_3/F_4$, the bound $w_k \leqslant 2/F_{k+3}$ (also best possible) can be proved, using the Fibonacci numbers F_n $(F_0=0,\ F_1=1)$. It follows then from this that the geometric mean of the probabilities w_i/w_{i-1} along the path $w_0, w_1, \ldots, w_k$ form the root $w_0 = 1$ to an internal node w_k is upper bounded by $((\sqrt{5}+2)/4)^{1/k}\psi$ for $k \geqslant 2$.

科研費シンポジウム「情報量と統計的推測理論」

ＭＣＱテストにおけるファジィ情報と得点分布

川崎医科大学　数学教室　有田清三郎

多肢選択テスト（Multiple-Choice Questions，以下ＭＣＱと略す）は医師国家試験をはじめ、歯科医師、看護婦、臨床検査技師、栄養士、リハビリテーション技師（ＯＴ，ＰＴ）等の国家試験、公務員採用資格試験、共通１次テストなど、広い分野にわたって実施されている。

ＭＣＱでは５肢択一式、２連式、３連式、複合連式、Ｋタイプなどの種々の形式があるが、５肢択一式では５つの選択肢が与えられ、その中に１個の正選択肢が隠されている。従って、受験者は何も知らなくても偶然に（確率１／５で）正解を言い当てることができる。これが従来から指摘されてきた「あて推量」（厳重には正答率の１／５のあて推量）である。ＭＣＱではあて推量で得点を増加することができる。（"The random guessing model", Callendra）

これに対して、我々はこのあて推量とは別の不完全な知識によるＭＣＱ特有の得点に影響を与える因子を見つけた。すなわち、

（１）与えられた選択肢のうち、特定の選択肢または複数個の選択肢を知っていれば、確実に正答ができる。（正答ターミナル）

（２）与えられた選択肢の中には、特定の１肢または複数個の選択肢を誤って理解していれば必ず誤答する。（誤答ターミナル）

（３）解答を答えるに不完全な知識では、無知識による完全なあて推量による解答のほかに正答率１／２，１／３等のあて推量による解答方式がある。

本研究では、まず５肢択一式、２連式、Ｋタイプ等の種々の形式について正答ターミナル、ランダム・ゲッシング領域、及びＭＣＱの禁止則を数学モデルによって明らかにした。また、それらの推移確率によって得点分布を求め、不完全情報による正の効果、負の効果を示した。

次に、ＭＣＱテストのうち、最もシンプルな形式の２肢択一型について、完全な情報と不完全なあいまい情報による知識と得点の数理モデルをつくり、これに基づいて受験者の得点分布を求めた。

次に、３肢択一型についての数理モデルを作り、２肢択一型との比較検討を

行い一般のｍ肢択一型への拡張理論を構築した。

（１）受験者のあいまいな知識を数量的に表現し、正しい知識、誤った知識、不明等の知識空間の中で受験者の知識を点として位置づけた。

（２）従来のＭＣＱモデルでは、完全な知識による正解と正答率１／ｍのランダム・ゲッシングだけが考察され、また誤った知識に基づく考察はなされなかった。今回の研究では完全な知識だけでなく、部分的な知識や不完全なあいまいな知識に基づく数理モデルを構成し、正答ターミナル（確実に正答する）の領域、誤答ターミナル（確実に誤答する）の領域、及び正答率が１／２，１／３等と種々の値をとるランダム・ゲッシング領域を数量的に明らかにした。

（３）受験者の知識はただ単に正答ターミナル領域、誤答ターミナル領域、ランダム・ゲッシング領域にあるかという知識空間の位置を示すのみならず、正答ターミナル領域あるいは誤答ターミナル領域への推移確率を定義できる。これを基に、あいまいな知識の他の知識状態への推移を確率モデルを使って数量化し、種々の各知識に対する得点分布を明らかにした。

（４）この得点分布によって、正確にわかっている選択肢の数（受験者の正知識）が得点と比例せず、また誤って理解した選択肢の数（受験者の誤知識）が得点に対応しないことを示した。

文献

(1) Arita, S., Saito, T. and Nasu, I. (1986): Score distribution of examinees with partial knowledge by the test of multiple-choice questions. Proc. of the Second Japan-China Symposium on Statistics, 5-8.

(2) Arita, S. and Saito, T. (1988): Fuzzy inference in the Multiple-Choice Question Test -Study on the principles of Fuzzy Rules-. 4th Fuzzy System Symposium, 303-307.

(3) Arita, S. and Saito, T. (1988): Fuzzy inference for finding the correct code in the multiple-choice question test. Proc. of International Workshop on Fuzzy System Applications, 131-132.

Projection method for statistical inference
in nonlinear model

S. Eguchi （江口真透）
Department of Mathematics, Shimane University
（島根大，理，数学）

The main purpose is to introduce and give a further development of the projection method for a nonlinear model on the basis of classical regression theory. Let θ be a p-parameter vector in the space Θ designating the population distribution with density $p(\mathbf{x}|\theta)$ and let $\bar{\mathbf{y}} = \sum_{i=1}^{N} \mathbf{y}(\mathbf{x}_i)/N$ be a higher dimensional statistic vector with mean vector $\eta(\theta)$ based on a random sample $\mathbf{x}_1, \cdots, \mathbf{x}_N$. The key idea is to find a one-to-one transformation $\mathbf{t}$ of $\mathbf{y}$ such that the image $\mathbf{t}(\mathcal{M})$ is a flat subset in the transformed space with $\mathcal{M} = \{ \eta(\theta) \mid \theta \in \Theta \}$, or equivalently of form $\mathbf{t}(\eta(\theta)) = X\theta + \mathbf{b}$, where X is a constant matrix of size n×p. Here the original parameter vector θ may be reparametrized. Next we define a generalized square distance

$$S(\theta) = N\{\mathbf{t}(\bar{\mathbf{y}}) - (X\theta + \mathbf{b})\}^T \hat{V}^{-1}\{\mathbf{t}(\bar{\mathbf{y}}) - (X\theta + \mathbf{b})\}$$

where $\hat{V}$ is a consistent estimate of the variance matrix of $\mathbf{t}(\bar{\mathbf{y}})$,

$$\hat{V} = \left(\frac{\partial \eta}{\partial \mathbf{t}} \mathrm{Var}(\eta) \left(\frac{\partial \eta}{\partial \mathbf{t}} \right)^T\right)_{\eta=\bar{\mathbf{y}}}.$$

Immediately $S(\theta)$ has a minimum $S(\hat{\theta})$ with $\hat{\theta} = (X^T\hat{V}^{-1}X)^{-1}X^T\hat{V}^{-1}\mathbf{t}(\bar{\mathbf{y}})$, so that

$$S(\theta) = S(\hat{\theta}) + N(\hat{\theta} - \theta)^T X^T \hat{V}\, X(\hat{\theta} - \theta),$$

which are asymptotically independent and distributed as χ^2_{n-p} and χ^2_p with the dimension n of $\mathbf{t}$. Consequently we propose the statistic $S(\hat{\theta})$ as a test for a hypothesis H : $\eta \in \mathcal{M}$ for the mean vector η of $\bar{\mathbf{y}}$. Simultaneously $\hat{\theta}$ works as an estimator of θ under the null hypothesis H.

We apply the projection method to a problem of testing Hardy-Weinberg equilibrium (HWE). In a HWE population with alleles $A_1, \cdots, A_m$, the respective frequencies $p_1, \cdots, p_m$ are summed as the homozygotes and heterozygotes

$$\sum_i p_i^2 + 2\sum_{i<j} p_i p_j$$

in the unit.

The HWE testing is fundamental and important in population genetics and statistical analysis. We are concerned with further research for HWE testing. The main purpose is to develop an option of the projection method for detecting an individual genotype or some group which deviates from HWE. Hernandez and Weir have given the disequilibrium coefficient approach to the problem where the coefficient is defined as the difference between the frequency of each heterozygote and the Hardy-Weinberg expected frequency. We detect a genotype away from HWE by succesive application of the projection method to the problem of testing whether the coefficient corresponding to the target genotype is zero or not.

We next analyze the real data in the HLA system by the use of the proposed method. The HLA system is characterized in polymorphism and incomplete identification. Thus the system is considered as a generalized ABO like system with identified codominant alleles and a recessive allele, see Gart and Nam (1984). In this context the recessive allele is considered as a set of unidentified alleles at the stage when sampling. The projection method for the overall test of deviations from Hardy-Weinberg equilibrium (HWE) at a locus in the HLA system is discussed in Eguchi and Matsuura (1990).

ひずみを持つ分布の平均のロバスト推定について

中央大学　鎌倉稔成

1. はじめに

ロバストということばはいろいろな意味で使われるが、最も広く使われているロバストはHuber流の ε-contaminationによるものだろう。しかしながら、最も一般的な定義としは、Hampel et al.(1986)の

"In a broad informal sense, robust statistics is a body of knowledge, partly formalized into 'theories of robustness,' relating to deviations from idealized assumptions in statistics."

である。実際にはロバスト推定に対する理論の多くはHuber流の定義によるもので、仮定する分布に異った分布がわずか混入した場合になるべくその影響を受けないように推定方式を設計するのがそのねらいである。このとき、仮定したモデルを$\{F_\theta\}$とすれば、データが未知の構造を持ったモデル、

$$(1-\varepsilon)F_\theta+\varepsilon H$$

から得られたものと考える。HはF$_0$とは異る分布であり、未知である。つまり、Hにあまり影響を受けずに θ を推測しようというものである。

これに対して仮定したモデル$\{F_\theta\}$とデータを与える真のモデル$\{G_\theta\}$がまったく違うモデルの場合も考えられる。たとえば、仮定した分布が正規分布としたとき、真のモデルが裾の重い歪みのある分布の場合がこれにあたる。

2. 歪みのある分布のロバスト推定

まず、Huber流のロバスト推定について考える。ロバスト推定に対する研究はほとんどが分布に対称性を仮定している。Hussein(1989)は対数正規分布、ワイブル分布について最適なB-robust推定量を構成するためのアルゴリズムを与えている。これを簡単にみることにする。

$X_1,X_2,\cdots,X_n$ は互いに独立で共通のc.d.f.F_θを持つものと仮定する。ここで関心のあるパラメータ θ の統計量として

$$T_n(X_1,\cdots,X_n)=T(F_n)$$

を考える。Tはフィッシャーの一致性を持つものとする。汎関数TがB-robustであるとは、影響関数IF(.;T,F)があるノルムについて有界関数である場合をいう。ただし、影響関数IFは

$$IF(x;T,F)=\lim_{\varepsilon\to 0}\frac{T[(1-\varepsilon)F+\varepsilon\delta_x]-T(F)}{\varepsilon}$$

で与えられる。ただし、δ_xはxのみでmassを持つ退化した分布関数である。ここ

にvon Misesの展開に必要な正則条件のもとで、

$$T(F_n) - T(F) = \int IF(x;T,F)\,dF_n(x) + o_p(n^{-\frac{1}{2}})$$
$$\int IF(x;T,F)\,dF(x) = 0$$
$$\sqrt{n}[T(F_n) - T(F)] \xrightarrow{L} N(0, AV(T,F))$$
$$AV(T,F) = \int IF(x;T,F)(IF(x;T,F))'\,dF(x)$$

である。Hampel(1968)はscore function,

$$\tilde{\eta}(x,\theta) = [s(x,\theta_*) - a]_{-b}^{b}$$

を持つM-estimatorが θ のoptimal B-robust estimator であることを示した。Peracchi(1987)はこの結果を一般化し、p次元のM-estimatorに対する定理を与えた。さらに、この解を求めるためのアルゴリズムを与えたのがHussein(1989)である。

3. 信頼区間とロバスト推定

一般にロバスト推定量から区間推定を構成するのはむずかしい。統計量の正確な分布は複雑すぎて漸近論に頼らざるを得ない。したがって、小標本の性質をみるためにはシミュレーションやブート・ストラップ等の方法を使う必要がある。ここでは、Staudte and Sheather(1990)の例2のデータについて信頼区間を比較してみる。

	μ	95%信頼区間	
Student's t	31.79	30.88	32.71
Sign	31.95	31.01	32.45
Wilcoxon	31.85	31.20	32.50
2*5%trmean	31.86	31.24	32.51
2*10%trmean	31.85	31.21	32.49
Huber $\psi_{1.28}$	31.86	31.27	32.45

Kamakura and Yanagimoto(1989)にワイブル分布の平均の推測に、

$$T(\mu) = 2n\left(\frac{\bar{x}^2}{s^2} - \frac{1}{n}\right)\left(\frac{\bar{x}}{\mu} - \log\frac{\bar{x}}{\mu} - 1\right)$$

を提案し、$T(\mu) = t_{1-\alpha/4}^2(n-1)$ の解として信頼区間を構成できるとした。この統計量は仮定する分布がワイブル分布によらなず、真の平均が μ のとき、自由度１の χ^2 分布に従うという意味でロバストであるということが示されている。また、信頼区間については、(30.75,32.89)とt分布の結果とほとんど同じである。sを平均絶対偏差で置き換えればロバストな結果が得られる。

推定方程式の理論からみたモーメント法

統計数理研究所　　柳本　武美

1．序

モーメント法は従来 quick and dirty な方法としてその計算の簡便性が強調されてきた。しかし計算の負担が大幅に軽減された現在ではその長所は重要ではない。一方モーメント法は直観的に分かり易い。またその式が簡単であることから、尤度法に比べてロバストであることが期待される。

2．推定方程式の理論

尤度推定関数 $L(\underline{x};\theta)$ は次の正則条件を満たすことが多い。

a) $E(L(\underline{x};\theta)) = 0$　　　　（不偏性）

b) $E(L'(\underline{x};\theta)) = -E(L^2(\underline{x};\theta))$　　（情報不偏性）

c) $L(\underline{x};\theta) = 0$ は θ に関して一意の解をもつ。$\hat{\theta}$ を根としたとき、$L(\underline{x};\theta)$ は $\theta < \hat{\theta}$ で正、$\theta > \hat{\theta}$ で負。

Godambe（１９６０）は不偏な推定関数 $g(\underline{x};\theta)$ について一つの criterion

$$M(g(\underline{x};\theta)) = V(g(\underline{x};\theta))/E(g'(\underline{x};\theta))^2$$

を導入した。そして尤度推定関数が criterion $M(\cdot)$ を最小にするという意味で最適であることを示した。

3．研究の方法

モーメント法を拡張するために次の方法で研究を行なった。

a）　関心のある母数を簡単な低次のモーメントからなる方程式の解として与える。

b）　低次のモーメントに不偏な標本モーメントを代入して推定関数を得る。

c）　推定関数の良さを Godambe の規準で評価する。

更に次の方針も採用している。

d）　すべての母数を推定する。

e）　対応する検定をも導出する。

4．研究の成果

イ）　Mantel-Haenszel 推定量の最適性を与えた。　また Mantel検定との関連を与えた。

イ'）　オッズ比が canonical link として現われることを指摘して、負の2項分布での同様の推定量、検定量が構成できることを示した。

ロ）　2つの指数分布の平均の比　$\theta = \lambda_1/\lambda_2$　の推測では

$$(m-1/m)(\bar{x}/\bar{y}) - \theta = 0$$

よりも

$$\bar{x} - \theta\bar{y} = 0$$

が良い。

ハ）　変動係数　$\theta = \sigma^2/\mu^2$　の推定では

$$\Sigma(x_i-\bar{x})^2/(n-1) - \theta\{\bar{x}^2-\Sigma(x_i-\bar{x})^2/n(n-1)\} = 0$$

を根として推定する。

等が得られる。

著者（ら）の文献

Yanagimoto, T. (1990a). Combining moment estimates of a parameter common through strata (with discussion). J. Statist. Plan. Inf. 25, 187-198.

Yanagimoto, T. and Yamamoto, E. (1990). The role of unbiasedness in estimating equations. Submitted to Estimating Function (ed. by V.P. Godambe) Oxford University Press.

Yanagimoto, T. (1990b). The Mantel-Haenszel statistics for the extended odds ratio in the negative binomial distribution, Manuscript.

Yanagimoto, T. and Yamamoto, E. Moment methods for the linear canonical link regression model. (In preparation).