数理解析研究所講究録
第 728 巻 1990 年 70-81

超母集団モデルにおける最適線形予測について

筑波大・理工　　河合　伸一 (Shinichi Kawai)

1．はじめに

標本調査法における推測理論には，固定母集団アプローチと超母集団アプローチがある．固定母集団アプローチでは，母集団の各ユニットの変量は固定されているが未知であるとして理論を展開する．

超母集団アプローチでは，母集団の各ユニットの変量は，確率構造が規定された確率変数であるとして理論を展開する．したがって，各ユニットの変量の値は，確率変数の実現値として扱われる．

標本調査法に，統計的推測理論でよく用いられる種々の基準を最初に適用したのは Godambe (1955) である．そこでは，ある一般的なクラスの下では一様最小分散不偏推定量が存在しないことが示されている．

固定母集団アプローチは，古くからある方法であるが，最近の多くの成果は超母集団アプローチから得られている．ここでは超母集団アプローチを用いる．そして，有限母集団の各ユニットの変量の総和を予測する問題を，ある種の不偏性を仮定した線形予測量のクラスの中で考え，適当な基準のもとで最適な予測量を見つける．

2．記号と定義

この節において， Cassel, Särndal, and Wretman (1977) にしたがって，記号と定義を含めて後のための準備を行う．

有限母集団の大きさを N とする．標本調査法では，どの標本をとるかを自分で選ぶことができる．それは，母集団を構成する各ユニットを識別するためのラベルがついているからである．このラベルが観測できるというのが標本調査法における推測の大きな特徴である．ここでは，ラベルを $1,\ldots,N$ の数字で表し，ラベル k のついたユニットをユニット k と呼ぶことにする．

ユニット k が抽出されて値 y_k を観測したとき，ラベルと値の対 (k, y_k) をラベル付き観測値という．

有限母集団を，ラベルの集合

$$\mathcal{U} = \{1, \ldots, k, \ldots, N\}$$

で表すことにする．また，

$$\mathbf{y} = (\, y_1, \ldots, y_N \,)$$

とする．

標本抽出において，ユニットを一つ一つ順番に抽出して得られる順序付き標本と，これから順序とユニットの重複を除いた順序無し標本とを区別する．今，順序付き標本を

$$\mathbf{s} = (\, k_1, \ldots, k_{n(\mathbf{s})} \,)$$

順序無し標本を

$$s = \{\, k_1, \ldots, k_{\nu(s)} \,\}$$

で定義する．ここで，$n(\mathbf{s})$ は順序付き標本 $\mathbf{s}$ の大きさ，$\nu(s)$ は順序無し標本 s の大きさを表す．また，順序付き標本の集合を $\mathcal{S}^*$，順序無し標本の集合を $\mathcal{S}$ で表す．

ラベル付き観測値 (k, y_k) を用いて，順序付きデータを

$$\mathbf{d} = (\, (k, y_k) : k \in \mathbf{s} \,)$$

順序無しデータを

$$d = \{\, (k, y_k) : k \in s \,\}$$

で定義する．

標本の抽出方法は，$\mathcal{S}^*$ あるいは $\mathcal{S}$ 上の確率分布 $p(\cdot)$ で与えられる．この確率分布 p のことをデザインと呼ぶことにする．標本調査法では，このデザイン p を調査者自ら与えることができる．

$$p(\mathrm{s}) \geq 0 \qquad \text{for all } \mathrm{s} \in \mathcal{S}^*$$

$$\sum_{\mathrm{s}\in\mathcal{S}^*} p(\mathrm{s}) = 1$$

を満足するような $\mathcal{S}^*$ 上の関数 $p(\mathrm{s})$ を順序付きデザインという．また，

$$p(s) \geq 0 \qquad \text{for all } s \in \mathcal{S}$$

$$\sum_{s\in\mathcal{S}} p(s) = 1$$

を満足するような $\mathcal{S}$ 上の関数 $p(s)$ を順序無しデザインという．

デザイン p は $p(\cdot)$ が $\mathbf{y}$ に依存しないとき無情報デザイン (noinformative design) と呼ぶ．ここでは無情報デザインのみを扱う．

これより先は，順序無しデータからつくった統計量の十分性より，順序無し標本，データ，デザインのみを扱う．

$$p(s) > 0 \quad \Rightarrow \quad \nu(s) = n \qquad \text{for all } s \in \mathcal{S}$$

であるようなデザインを固定実質標本サイズデザイン (fixed effective size design) と呼び $FES(n)$ と書く．

ユニット k を含む標本の集合を

$$C_k = \{\, s : k \in s \,\}$$

で表す．$\mathcal{S}$ 上のデザイン p が与えられたとき，ユニット k を抽出する確率は

$$\Pi_k = \sum_{s\in C_k} p(s)$$

で与えられる． Π_k をユニット k の被抽出確率 (inclusion probability) と呼ぶ．

超母集団アプローチでは， $y_1,\ldots,y_N$ は確率変数の実現値として表される．有限母集団は超母集団からの N 個の標本と考え， $Y_1,\ldots,Y_N$ は超母集団において定められた分布 ξ に従うものとする．

データにおいて， y_k , s を確率変数で置き換えたものを

$$D = \{\ (k, y_k) : k \in S\ \}$$

$$\mathcal{D} = \{\ (k, Y_k) : k \in S\ \}$$

$$d = \{\ (k, Y_k) : k \in s\ \}$$

と定義する．

次に， $T = \sum_{k=1}^{N} Y_k$ の予測量として主なものを紹介しておく．

線形予測量

$$t(\ \mathcal{D}\ ) = w_{0S} + \sum_{k \in S} w_{kS} Y_k$$

斉次線形予測量 (線形予測量で $w_{kS} = 0$ のもの)

$$t(\ \mathcal{D}\ ) = \sum_{k \in S} w_{kS} Y_k$$

ここで重み w_{kS} は S に依存して良いことに注意する．

Horvitz-Thompson 予測量

$$t_{HT} = \sum_{k \in S} \frac{Y_k}{\Pi_k}$$

Generalized difference 予測量

$$t_{GD} = \sum_{k \in S} \frac{Y_k - e_k}{\Pi_k} + \sum_{k=1}^{N} e_k$$

ここで、 e_k $(k=1,\ldots,N)$ は任意の定数である.

デザイン p が与えられたときの，予測量 t のデザインに関する期待値，分散を，それぞれ，

$$E_p(\ t(\mathcal{D})\ ) = \sum_{s\in\mathcal{S}} p(\ s\ )t(\ d\ )$$

$$V_p(\ t(\mathcal{D})\ ) = \sum_{s\in\mathcal{S}} p(\ s\ )\{t(\ d\ ) - E_p(\ t\ )\}^2$$

で定義する．デザイン p が与えられたとき，予測量 t について

$$E_p(\ t(D)\ ) = \sum_{k=1}^{N} y_k \quad \text{for all} \quad \mathbf{y}\in\mathbf{R}^N$$

が成り立つとき，予測量 t は p–不偏であるという.

t の p–不偏性について，次の定理２．１がなりたつ.（たとえば Cassel, Särndal, and Wretman (1977) を見よ）

定理２．１． T の p–不偏予測量 t が存在するための必要十分条件は，$\Pi_k > 0$ $(k=1,\ldots,N)$ である.

証明. すべての k について $\Pi_k > 0$ ならば， Horvitz-Thompson 予測量が， p–不偏である．逆に，ある k_0 について $\Pi_{k_0} = 0$ ならば，どんな予測量も y_{k_0} に依存しないことになるから， T の p–不偏予測量とはなりえない.

$Y_1,\ldots,Y_N$ の同時確率分布を ξ とするとき， $Q = Q(Y_1,\ldots,Y_N)$ の期待値，分散を，それぞれ

$$\mathcal{E}(Q) = \int Q d\xi$$

$$\mathcal{V}(Q) = \int \{Q - \mathcal{E}(Q)\}^2 d\xi$$

$Q_1 = Q_1(Y_1,\ldots,Y_N)$, $Q_2 = Q_2(Y_1,\ldots,Y_N)$ $(Q_1 \neq Q_2)$ の共分散を

$$C(Q_1\,,\,Q_2) = \int \{Q_1 - \mathcal{E}(Q_1)\}\{Q_2 - \mathcal{E}(Q_2)\}d\xi$$

で定義する.

3．問題の設定

$Y_1,\ldots,Y_N$ は,

$$\mathcal{E}(Y_k) = \mu_k \qquad (k = 1,\ldots,N)$$

$$\mathcal{V}(Y_k) = \sigma_k^2 \qquad (k = 1,\ldots,N)$$

$$C(Y_k, Y_l) = \rho\sigma_k\sigma_l \qquad (k \neq l = 1,\ldots,N)$$

ここで, μ_k, $\sigma_k(>0)$, $(k = 1,\ldots,N)$, $\rho(-1/(N-1) \leq \rho \leq 1)$ は未知.

を満たすような同時確率分布 ξ に従うとする. いまこれをモデル M と呼ぶことにする. そして, モデル M, デザイン p のもとで標本 s をとり, $T = \sum_{k=1}^{N} Y_k$ を予測する問題を考える.

予測量のクラスとしては, 斉次線形 p–不偏予測量のクラス $\mathcal{L}_{0u}$ と, 線形 p–不偏予測量のクラス $\mathcal{L}_u$ を考える. そして,

$$\mathcal{E}E_p(t-T)^2$$

を最小にするような, デザインと予測量の組 (p,t) を $\mathcal{L}_{0u}$, $\mathcal{L}_u$の中からそれぞれ選ぶ.

4．斉次線形 p–不偏予測量のクラスでの最適性

$t \in \mathcal{L}_{0u}$ のとき, $t = \sum_{k\in S} w_{kS}Y_k$ の形で書き表される. また, t が p–不偏であるための必要十分条件は,

$$\sum_{s\in C_k} w_{ks}p(\,s\,) = 1 \quad \text{for} \quad k = 1,\ldots,N$$

である.

$\mathcal{L}_{0u}$ での最適性について，次の定理４．１が成り立つ.

定理 ４．１． (Arnab (1986)) モデル M のもとで，p を $\Pi_k > 0$ $(k = 1,\ldots,N)$ であるような任意のデザイン，$t \in \mathcal{L}_{0u}$ とするとき，$\rho \geq 0$ ならば,

$$\mathcal{E}E_p(t-T)^2 \geq (1-\rho)\sum_{k=1}^{N}\sigma_k^2\left(\frac{1}{\Pi_k}-1\right)$$

である．等号成立は,

$$\text{(i)}\ w_{ks} = \frac{1}{\Pi_k} \quad \text{for all} \quad k,\ s \quad \text{with} \quad p(\ s\ ) > 0$$

$$\text{(ii)}\ \sum_{k\in s} w_{ks}\sigma_k = \sum_{k=1}^{N}\sigma_k \quad \text{for all} \quad s \quad \text{with} \quad p(\ s\ ) > 0$$

$$\text{(iii)}\ \sum_{k\in s} w_{ks}\mu_k = \sum_{k=1}^{N}\mu_k \quad \text{for all} \quad s \quad \text{with} \quad p(\ s\ ) > 0$$

を同時に満たすときに限る.

(i)-(iii) を満たすとき，最適予測量は

$$t_{HT} = \sum_{k\in S}\frac{Y_k}{\Pi_k}$$

となる.

５．線形 p–不偏予測量のクラスでの最適性

$t \in \mathcal{L}_u$ のとき，$t = w_{0S} + \sum_{k\in S} w_{kS}Y_k$ の形で書き表される．また，t が p–不偏であるための必要十分条件は,

$$\sum_{s\in\mathcal{S}} w_{0s}p(\ s\ ) = 0$$

かつ,

$$\sum_{s\in C_k} w_{ks}p(\ s\ )=1 \quad \text{for} \quad k=1,\ldots,N$$

である.

$\mathcal{L}_u$ での最適性について、次の定理５．１が成り立つ．これは定理４．１の拡張になっている.

定理５．１． モデル M のもとで，p を $\Pi_k > 0 \quad (k=1,\ldots,N)$ であるような任意のデザイン，$t\in\mathcal{L}_u$ とするとき，$\rho\geq 0$ ならば,

$$\mathcal{E}E_p(t-T)^2 \geq (1-\rho)\sum_{k=1}^{N}\sigma_k^2\left(\frac{1}{\Pi_k}-1\right)$$

である。等号成立は,

$$\text{(i)}\ w_{ks}=\frac{1}{\Pi_k} \quad \text{for all} \quad k,\ s \quad \text{with} \quad p(\ s\ )>0$$

$$\text{(ii)}\ \sum_{k\in s} w_{ks}\sigma_k=\sum_{k=1}^{N}\sigma_k \quad \text{for all} \quad s \quad \text{with} \quad p(\ s\ )>0$$

$$\text{(iii)}\ w_{0s}+\sum_{k\in s} w_{ks}\mu_k=\sum_{k=1}^{N}\mu_k \quad \text{for all} \quad s \quad \text{with} \quad p(\ s\ )>0$$

を同時に満たすときに限る.

(i)-(iii) を満たすとき，最適予測量は

$$t_{GD}=\sum_{k\in S}\frac{Y_k-\mu_k}{\Pi_k}+\sum_{k=1}^{N}\mu_k$$

となる.

証明.

$$\mathcal{E}E_p(\,t-T\,)^2=\mathcal{E}E_p(\,t^2\,)-\mathcal{E}(\,T^2\,)$$

$$
\begin{aligned}
&= \mathcal{E}\left\{\sum_{s\in\mathcal{S}} p(s)\left(w_{0s}+\sum_{k\in s} w_{ks}\,Y_k\right)^2\right\} - \mathcal{E}(T^2) \\
&= \mathcal{E}\left[\sum_{s\in\mathcal{S}} p(s)\left\{w_{0s}^2 + 2\,w_{0s}\left(\sum_{k\in s} w_{ks}\,Y_k\right)\right\}\right] \\
&\quad + \mathcal{E}\left\{\sum_{s\in\mathcal{S}} p(s)\left(\sum_{k\in s} w_{ks}\,Y_k\right)^2\right\} - \mathcal{E}(T^2) \\
&= \sum_{s\in\mathcal{S}} p(s)\left\{w_{0s}^2 + 2\,w_{0s}\left(\sum_{k\in s} w_{ks}\,\mu_k\right)\right\} \\
&\quad + (1-\rho)\sum_{k=1}^{N}\sigma_k^2 \sum_{s\in C_k} p(s)\,w_{ks}^2 \\
&\quad + \rho\sum_{s\in\mathcal{S}} p(s)\left(\sum_{k\in s} w_{ks}\,\sigma_k\right)^2 \\
&\quad + \sum_{s\in\mathcal{S}} p(s)\left(\sum_{k\in s} w_{ks}\,\mu_k\right)^2 - \mathcal{E}(T^2) \\
&= (1-\rho)\sum_{k=1}^{N}\sigma_k^2 \sum_{s\in C_k} p(s)\,w_{ks}^2 \\
&\quad + \rho\sum_{s\in\mathcal{S}} p(s)\left(\sum_{k\in s} w_{ks}\,\sigma_k\right)^2 \\
&\quad + \sum_{s\in\mathcal{S}} p(s)\left(w_{0s}+\sum_{k\in s} w_{ks}\,\mu_k\right)^2 - \mathcal{E}(T^2).
\end{aligned}
$$

Schwarz の不等式を用いて,

$$
\sum_{s\in\mathcal{S}} w_{ks}^2\,p(s) \geq \frac{\left\{\sum_{s\in C_k} w_{ks}\,p(s)\right\}^2}{\sum_{s\in C_k} p(s)}.
$$

$$= \frac{1}{\Pi_k}, \qquad (k = 1, \ldots, N),$$

等号成立は,

$$(5.1) \qquad w_{ks} = \frac{1}{\Pi_k} \text{ for all } k,\ s \text{ with } p(s) > 0,$$

に限る. また,

$$\sum_{s \in \mathcal{S}} p(s) \left(\sum_{k \in s} w_{ks}\, \sigma_k \right)^2 \geq \left\{ \sum_{s \in \mathcal{S}} p(s) \sum_{k \in s} w_{ks}\, \sigma_k \right\}^2$$
$$= \left\{ \sum_{k=1}^{N} \sigma_k \sum_{s \in C_k} w_{ks}\, p(s) \right\}^2$$
$$= \left(\sum_{k=1}^{N} \sigma_k \right)^2,$$

等号成立は,

$$(5.2) \qquad \sum_{k \in s} w_{ks} \sigma_k = \sum_{k=1}^{N} \sigma_k \text{ for all } s \text{ with } p(s) > 0,$$

に限る. さらに,

$$\sum_{s \in \mathcal{S}} p(s) \left(w_{0s} + \sum_{k \in s} w_{ks}\, \mu_k \right)^2 \geq \left\{ \sum_{s \in \mathcal{S}} p(s) \left(w_{0s} + \sum_{k \in s} w_{ks}\, \mu_k \right) \right\}^2$$
$$= \left\{ \sum_{k=1}^{N} \mu_k \sum_{s \in C_k} w_{ks}\, p(s) \right\}^2$$
$$= \left(\sum_{k=1}^{N} \mu_k \right)^2,$$

等号成立は,

$$(5.3)\qquad w_{0s}+\sum_{k\in s}w_{ks}\mu_k=\sum_{k=1}^{N}\mu_k \quad \text{for all } s \text{ with } p(s)>0,$$

に限る. したがって,

$$\begin{aligned}\mathcal{E}E_p(t-T)^2 &\geq (1-\rho)\sum_{k=1}^{N}\frac{\sigma_k^2}{\Pi_k}+\rho\left(\sum_{k=1}^{N}\sigma_k\right)^2+\left(\sum_{k=1}^{N}\mu_k\right)^2\\ &\quad-\left\{\sum_{k=1}^{N}\sigma_k^2+\left(\sum_{k=1}^{N}\mu_k\right)^2+\rho\sum_{k=1}^{N}\sum_{\substack{l=1\\ l\neq k}}^{N}\sigma_k\sigma_l\right\}\\ &=(1-\rho)\sum_{k=1}^{N}\sigma_k^2\left(\frac{1}{\Pi_k}-1\right).\end{aligned}$$

等号成立は, (5.1) − (5.3) を同時に満たすときに限る.

モデル M の特別な場合として, $(Y_k-b_k)/a_k\ (k=1,\ldots,N)$ が, 共通の平均 μ, 分散 σ^2, 共分散 $\rho\sigma^2$ を持つような分布に従うモデルを考え, これをモデル G_T とする. ここで, $a_k\,(0<a_k<N/n), b_k\ (k=1,\ldots,N)$ は既知, $\sum_{k=1}^{N}a_k=N$ とし, μ, $\sigma(>0)$, $\rho(-1/(N-1)\leq\rho\leq 1)$ は未知とする.

モデル G_T のもとで, 定理5.1での条件 (ii) を満たすような $FES(n)$ デザインは,

$$\Pi_k=fa_k \qquad \left(f=\frac{n}{N}\right) \qquad (k=1,\ldots,N)$$

となるようなものに限る. 今これを p_0 とする. このとき, t_{GD} は

$$t_{GD_0}=\sum_{k\in S}\frac{Y_k-b_k}{fa_k}+\sum_{k=1}^{N}b_k$$

となる.

定理5．1の系として，次のことがいえる．

系. モデル G_T のもとで，予測量のクラスを $\mathcal{L}_u$，デザイン p を任意の $FES(n)$ デザインとするとき，$\rho \geq 0$ ならば，

$$\mathcal{E}E_p(t-T)^2 \geq \mathcal{E}E_{p_0}(t_{GD_0}-T)^2$$

である．

注意. Cassel, Särndal, Wretman (1976, 1977) は，$\rho < 0$ のときも上の系がなりたつことを示した．$\rho \geq 0$ のとき，定理5．1は Cassel, Särndal, Wretman (1976, 1977) の結果を含む．

参考文献

Arnab R. (1986). Optimal Prediction for a Finite Population Total with Connected Designs and Related Model-Based Results. *Metrika* **33** 79-84.

Cassel C.M., Särndal C.E., Wretman J.H. (1976). Some results on generalized difference estimation and generalized regression estimation for finite populations. *Biometrika* **63** 615-620.

Cassel C.M., Särndal C.E., Wretman J.H. (1977). *Foundations of Inference in Survey Sampling.* New York: Wiley.

Godambe, V. P. (1955). A unified theory of sampling from finite populations. *J. R. Statist. Soc. B* **17** 269-278.