

順序統計量を利用した最適配置について

東京理科大 理工 宮 川 強

東京理科大 理工 小 谷 孝一

§ 1. はじめに

位置母数 (location parameter) μ , 尺度母数 (scale parameter) σ を持つ連続型分布よりの n 個の任意標本 $X_1, X_2, \dots, X_n$ を値の小さい方から, 大きい方へならべた所謂順序統計量 $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ のうちの k 個 (k は $1 \leq k \leq n$ なる自然数) の $X_{(n_1)} < X_{(n_2)} < \dots < X_{(n_k)}$ ($1 \leq n_1 < n_2 < \dots < n_k \leq n$) を用いて適当な統計量を作り, 未知母数 μ, σ を推定する問題は従来より考えられて来ている。このうちで, 特に母分布が所謂正規分布の場合が中心問題である。初め F. Mosteller (1946) により提唱され, その後 J. Ogawa (1951) 及び其の他の人々により次第に解決されて来ている。以下, 母分布を正規分布と考えることにし, k 個の標本 $X_{(n_1)}, X_{(n_2)}, \dots, X_{(n_k)}$ の適当な線形式で μ, σ を推定する場合に『如何なる $n_1, n_2, \dots, n_k$ を選んで, 如何なる線形式を用いれば, 分散が最小となるか。(即ち, 相対効率が最小となるか。』という問題を考えることにする。

(1) n が有限の場合

これに対する解答は, $k=2, 4$ の場合は, $n \leq 15$ に対して, H. L. Harter (1959), T. Ishikawa (1971) 等により示されており, $k=2, 4, 6, 8, 10$ の場合は $n \leq 30$ に対して, K. Miyakawa & K. Kotani (1972) によって示されている。(具体的には, 文献 6 を参照されたい。)

(2) n が十分大なる場合 (n が無限大の場合)

σ が既知で, μ 未知の場合は, J. Ogawa (1951) により, μ の線形推定量が, ほとんど完全に近い形で, $k \leq 10$ まで与えられる。しかし μ 既知で σ 未知の場合は, σ の線形推定量は, $k \leq 6$ までしか与えられるといえず, μ, σ 共に未知な場合の μ, σ の推定は, ほとんど未解決といってよい。

本稿の目的は, 以下の (a), (b) であり, 主として (a) を取扱う。

(a) n が十分大で, σ 既知, μ 未知の場合の μ の推定

J. Ogawa の場合は $k \leq 10$ まで, 解決されているが, 実は, $\eta_1, \eta_2, \dots, \eta_k$ を決定する決定方程式が, 数値的に解くにしても, かなり大変な作業であり, $k > 10$ に対しては, 不可能に近いので, ゆずがに線形推定量の相対効率が落ちても, 決定方程式が簡単になり解きやすくなる様な線形推定量であれば, 実用上は有用であらうとの見地から与えられる様な線形推定量を考えることにある。

(b) n が十分大で, μ 既知, σ 未知の場合の σ の推定

この場合は、 σ を線形推定量で推定するかわりに、 σ^2 を 2 次の推定量で推定することとを考える。この場合は、 $n_1, n_2, \dots, n_k$ を決定する決定方程式の複雑さの点では、本質的には、J. Ogawa のより改良されてない。又我々の予想では、相対効率が J. Ogawa のよりよくなるつもりであつたが $k \leq 6$ までの試算の段階では、まったく同じとなる。(§3 での (iii) の場合)

§2. n が十分大なるとき、 μ 未知、 σ 既知の場合の線形推定量について。

$0 < \lambda_1 < \lambda_2 < \dots < \lambda_k < 1$ なる $\lambda_1, \lambda_2, \dots, \lambda_k$ を考える。

$n_1, n_2, \dots, n_k$ は $n_i = [n\lambda_i] + 1$ ($i=1 \sim k$) で定める。

今、標準正規分布の p.d.f を $g(x)$ とするとき、

$$\lambda_i = \int_{-\infty}^{u_i} g(x) dx \quad (i=1 \sim k)$$

で u_i を定義し、 $f_i \equiv g(u_i)$ ($i=1 \sim k$) とする。

今、配置の対称性 (即ち $u_i = -u_{k-i+1}$ ($i=1 \sim k$)) を仮定し、次の形な線形推定量 ($\hat{\mu}$) を考える。

$$\hat{\mu} = \frac{1}{\sum_{i=1}^k f_i} \sum_{i=1}^k f_i X_{(n_i)}$$

この場合は

$$V(\hat{\mu}) = \frac{\sigma^2}{n \left(\sum_{i=1}^k f_i \right)^2} \left\{ \sum_{i=1}^k \lambda_i (1 - \lambda_i) + 2 \sum_{i < j}^{1 \sim k} \lambda_i (1 - \lambda_j) \right\}$$

相対効率 $\eta(\mu)$ は

$$\eta(\mu) = \frac{\sigma^2}{n \cdot V(\hat{\mu})} \quad \text{である。}$$

こゝでの λ_i (λ_i が定まれば, u_i, f_i も定まる。) ($i=1 \sim R$) をこの $\eta(\mu)$ を最大にする標に, 即ち $V(\hat{\mu})$ を最小にする標に決定する。

具体的に決定するのは, 次の決定方程式による。

$k=2L$ (L は $L \geq 1$ なる整数) のとき

$$\begin{cases} u_i = (2L+1-2i)u_L, & (i=1 \sim L-1) \\ \sum_{j=1}^L f_j + 2u_L \sum_{j=1}^L (2L+1-2j)\lambda_j = 0 & (L=1 \text{ のときはこの式の4倍}) \end{cases}$$

$k=2L+1$ (L は $L \geq 1$ なる整数) のとき

$$\begin{cases} u_i = (L+1-i)u_L, & (i=1 \sim L-1) \\ \left(2 \sum_{j=1}^L f_j + \frac{1}{\sqrt{2\pi}} \right) + 2u_L \left\{ \sum_{j=1}^L 2(L+1-j)\lambda_j + \frac{1}{8} \right\} = 0 & (L=1 \text{ のときはこの式の4倍}) \end{cases}$$

$k \leq 30$ まで具体的に決定してあるが, 例示として,

$k=6, 8, 10, 20$ の場合の線形推定量 $\hat{\mu}$ を具体的に書き下し, 又 J. Ogawa の場合の相対効率と比較する。以下標本を k 個用いたときの $\hat{\mu}$ を $\hat{\mu}_k$ と書くことにする。

$k=6$ のとき

$$\begin{aligned} \hat{\mu}_6 = & 0.0725 \{ X_{([0.0518n]+1)} + X_{([0.9482n]+1)} \} + 0.1691 \{ X_{([0.1644n]+1)} + X_{([0.8356n]+1)} \} \\ & + 0.2583 \{ X_{([0.3724n]+1)} + X_{([0.6276n]+1)} \} \end{aligned}$$

$$\gamma(\mu) = 0.9531 \quad (\text{Ogawa の } t_{\frac{8}{11}}^{\frac{8}{11}} \text{ 合 } 0.9559)$$

$k=8$ のとき

$$\begin{aligned} \hat{\mu}_8 = & 0.0383 \{X_{([0.0308n]+1)} + X_{([0.9692n]+1)}\} + 0.0901 \{X_{([0.0910n]+1)} + X_{([0.9090n]+1)}\} \\ & + 0.1594 \{X_{([0.2116n]+1)} + X_{([0.7884n]+1)}\} + 0.2120 \{X_{([0.3947n]+1)} + X_{([0.6053n]+1)}\} \end{aligned}$$

$$\gamma(\mu) = 0.9693 \quad (\text{Ogawa の } t_{\frac{8}{11}}^{\frac{8}{11}} \text{ 合 } 0.9722)$$

$k=10$ のとき

$$\begin{aligned} \hat{\mu}_{10} = & 0.0228 \{X_{([0.0203n]+1)} + X_{([0.9797n]+1)}\} + 0.0522 \{X_{([0.0557n]+1)} + X_{([0.9443n]+1)}\} \\ & + 0.0971 \{X_{([0.1278n]+1)} + X_{([0.8722n]+1)}\} + 0.1469 \{X_{([0.2476n]+1)} + X_{([0.7524n]+1)}\} \\ & + 0.1806 \{X_{([0.4100n]+1)} + X_{([0.5900n]+1)}\} \end{aligned}$$

$$\gamma(\mu) = 0.9781 \quad (\text{Ogawa の } t_{\frac{10}{11}}^{\frac{10}{11}} \text{ 合 } 0.9808)$$

$k=20$ のとき

$$\begin{aligned} \hat{\mu}_{20} = & 0.0042 \{X_{([0.0055n]+1)} + X_{([0.9945n]+1)}\} + 0.0080 \{X_{([0.0114n]+1)} + X_{([0.9886n]+1)}\} \\ & + 0.0143 \{X_{([0.0223n]+1)} + X_{([0.9777n]+1)}\} + 0.0236 \{X_{([0.0409n]+1)} + X_{([0.9591n]+1)}\} \\ & + 0.0364 \{X_{([0.0705n]+1)} + X_{([0.9295n]+1)}\} + 0.0520 \{X_{([0.1142n]+1)} + X_{([0.8858n]+1)}\} \\ & + 0.0693 \{X_{([0.1744n]+1)} + X_{([0.8256n]+1)}\} + 0.0859 \{X_{([0.2517n]+1)} + X_{([0.7483n]+1)}\} \\ & + 0.0992 \{X_{([0.3440n]+1)} + X_{([0.6560n]+1)}\} + 0.1065 \{X_{([0.4467n]+1)} + X_{([0.5533n]+1)}\} \end{aligned}$$

$$\gamma(\mu) = 0.9927.$$

§ 3. n が十分大なるとき, μ 既知, σ 未知の場合の推定量につ

い 2.

$\lambda_i, u_i, n_i \ (i=1 \sim k)$ 等の記号は § 2 と同じ。

$k=2L$ (L は $L \geq 1$ の整数) の場合を考える。

(i) σ の推定量として, §2 と同様を理想正規分布として, 線形推定量 ($\hat{\sigma}$) を考える。

$$\hat{\sigma} = \frac{1}{\left(-2 \sum_{i=1}^k u_i f_i\right)} \sum_{i=1}^k f_i (X_{(n_{k-i+1})} - X_{(n_i)})$$

この場合は

$$V(\hat{\sigma}) = \frac{\sigma^2}{2n} \cdot \frac{1}{\left(\sum_{i=1}^k u_i f_i\right)^2} \left\{ \sum_{i=1}^k \lambda_i (1-2\lambda_i) + 2 \sum_{i < j}^{1 \sim k} \lambda_i (1-2\lambda_j) \right\}$$

相対効率 $\eta_1(\sigma)$ は

$$\eta_1(\sigma) = \frac{\sigma^2}{2n \cdot V(\hat{\sigma})}$$

である。

(ii) σ^2 の推定量 ($\hat{\sigma}^2$) として, 以下のものである。

$$\hat{\sigma}^2 = \frac{1}{\left(\sum_{i=1}^k f_i u_i^2\right)} \sum_{i=1}^k f_i (X_{(n_i)} - \mu)^2$$

この場合は

$$V(\hat{\sigma}^2) = \frac{4\sigma^4}{n \left(\sum_{i=1}^k f_i u_i^2\right)} \left\{ \sum_{i=1}^k u_i^2 \lambda_i (1-\lambda_i) + 2 \sum_{i < j} u_i u_j \lambda_i (1-\lambda_j) \right\}$$

相対効率 $\eta_2(\sigma^2)$ は

$$\eta_2(\sigma^2) = \frac{2}{n} \cdot \frac{\sigma^4}{V(\hat{\sigma}^2)}$$

である。

(iii) (ii) の場合を若干一般化して、以下の推定量 ($\hat{\sigma}^2$) を考える。

$$\hat{\sigma}^2 = (X_0 - \mu)' B (X_0 - \mu)$$

但し、 $B = \| b_{ij} \|_{(k \times k)}$: 対称かつ正対称かつ非負定符号行列。

$\hat{\sigma}^2$ の不偏性のためには $u' B u = 1$ とする。

こゝで

$$X_0 = \begin{bmatrix} X_{(n_1)} \\ X_{(n_2)} \\ \vdots \\ X_{(n_k)} \end{bmatrix}_{(k \times 1)}, \quad \mu = \mu \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(k \times 1)}, \quad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}$$

とする。

この場合は

$$V(\hat{\sigma}^2) = 4\sigma^4 u' B' V B u \quad (u' B u = 1)$$

但し、 $V = \| \lambda_{ij} \|_{(k \times k)}$: 対称である。

$$\text{又 } \lambda_{ij} = \frac{\lambda_i(1-\lambda_j)}{n s_i s_j} \quad (i \leq j)$$

相対効率 $\gamma_3(\sigma^2)$ は

$$\gamma_3(\sigma^2) = \frac{2}{n} \cdot \frac{\sigma^4}{V(\hat{\sigma}^2)}$$

である。

以上 (i), (ii), (iii) について

(i) は J. Ogawa の推定量よりは、直観的には、簡便推定量の称に思われるが、実は決定方程式が、かなり複雑となり、初期の目的を達成したとは思われない。

(ii) については、 $k=2, 4, 6$ を計算してみると、結果的には

J. Ogawa の相対効率と同じになり、決定方程式の解まにくさとも同称である。 $\mu=0$ の場合 E あとで例示する。

(iii) は (ii) を一般化した形での 2 次形式推定量である。結果的には、 $k=2, 4, 6$ のときは、J. Ogawa の相対効率と同じになり、決定方程式の解まにくさとも同称である。しかし、相対効率を J. Ogawa のより、あげ"得る"一応の方法と思われる。この場合推定量を具体的に定める場合、この 2 次形式推定量は一意に定まらず、種々の場合がある。

以下 $\hat{\sigma}^2$ を $k=2, 4, 6$ の場合に具体的に書き下す。以下標本 k 個を用いる場合の $\hat{\sigma}^2$ を $\hat{\sigma}_k^2$ と書くことにする。

$k=2$ のとき

$$\hat{\sigma}_2^2 = 0.2276 \{ X_{([0.0791n]+1)}^2 + X_{([0.9209n]+1)}^2 \}$$

$$\gamma_2(\sigma^2) = 0.6522 \quad (\text{Ogawa の場合 } 0.653)$$

$k=4$ のとき

$$\hat{\sigma}_4^2 = 0.0552 \{ X_{([0.0226n]+1)}^2 + X_{([0.9774n]+1)}^2 \} + 0.2115 \{ X_{([0.2115n]+1)}^2 + X_{([0.7885n]+1)}^2 \}$$

$$\gamma_2(\sigma^2) = 0.8240 \quad (\text{Ogawa の場合 } 0.824)$$

$k=6$ のとき

$$\hat{\sigma}_6^2 = 0.0202 \{ X_{([0.0099n]+1)}^2 + X_{([0.9901n]+1)}^2 \} + 0.0803 \{ X_{([0.0513n]+1)}^2 + X_{([0.9487n]+1)}^2 \} \\ + 0.1924 \{ X_{([0.1692n]+1)}^2 + X_{([0.8308n]+1)}^2 \}$$

$$\gamma_2(\sigma^2) = 0.8926 \quad (\text{Ogawa の場合 } 0.893)$$

参考文献

- [1] F. Mosteller, "On Some Useful Inefficient Statistics,"
Ann. Math. Statist., Vol 17. (1946)
- [2] J. Ogawa, "Contribution to the Theory of Systematic
Statistics I," Osaka Math. Jour., Vol 3 (1951)
- [3] W. J. Dixon, "Estimation of the Mean and Standard
Deviation of the Normal Population,"
Ann. Math. Statist., Vol 28 (1957)
- [4] A. E. Sarhan and B. G. Greenberg, "Contribution to Order
Statistics," Wiley, New York, (1962)
- [5] T. Ishikawa, "系統統計量について,"
相模工大紀要. Vol 5. (1971)
- [6] K. Miyakawa and K. Kotani, "On the Estimation of the
Location and Scale Parameters based on Order
Statistics in the Case of Normal Population,"
東京理科大. 理学専攻科(数学専攻)紀要. Vol 1. (1972).