数理解析研究所講究録
197巻 1974年 84-92

順序統計量を利用しての最適配置について

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, 尺度母度 (scale parameter) σ を持つ連続型分布よりの n 個の任意標本 $X_1, X_2, \cdots\cdots, X_n$ 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 $X_{(1)} < X_{(2)} < \cdots\cdots < X_{(n)}$ のうちの k 個 (k は $1 \leqq k \leqq n$ なる自然数) の $X_{(n_1)} < X_{(n_2)} < \cdots\cdots < X_{(n_k)}$ ($1 \leqq n_1 < n_2 < \cdots\cdots < n_k \leqq n$) 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きている。以下、母分布を正規分布と考えることにし、k 個の標本 $X_{(n_1)}, X_{(n_2)}, \cdots\cdots, X_{(n_k)}$ の適当な線形式で μ, σ を推定する場合に『如何なる $n_1, n_2, \cdots, n_k$ 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、$n \leqq 15$ 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 $n \leqq 30$ 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、$k \leqq 10$まで与えられている。しかしμ既知でσ未知の場合は、σの線形推定量は、$k \leqq 6$までしか与えられていず、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$k \leqq 10$まで解決されているが、実は、$n_1, n_2, \cdots, n_k$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、わずかに線形推定量の相対効率がおちても、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、実用上は有用であろうとの見地から§2で与える様な線形推定量を考えることにする。

(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を考える。この場合は，$n_1, n_2, \cdots n_k$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ったが $k \leqq 6$ 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。

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

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

今，標準正規分布の p.d.f を $g(x)$ とするとき，

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

で u_i を定義し，$f_i \equiv g(u_i) \quad (i=1\sim k)$ 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，次の様な線形推定量 ($\hat{\mu}$) 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

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

である。

こゝでの λ_i(λ_iが定まれば, u_i, f_iも定まる。)($i=1\sim k$)をこの$\eta(\mu)$を最大にする様に, 即ち$V(\hat{\mu})$を最小にする様に決定する。具体的に決定するのは, 次の決定方程式による。

$k=2\nu$ (νは$\nu\geqq 1$なる整数)のとき

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

$k=2\nu+1$ (νは$\nu\geqq 1$なる整数)のとき

$$\begin{cases} u_i=(\nu+1-i)u_\nu, \quad (i=1\sim \overline{\nu-1}) \\ \left(2\sum_{j=1}^{\nu} f_j+\frac{1}{\sqrt{2\pi}}\right)+2u_\nu\left\{\sum_{j=1}^{\nu}2(\nu+1-j)\lambda_j+\frac{1}{8}\right\}=0 \\ \qquad (\nu=1\text{のときはこの式のみ}) \end{cases}$$

$k\leqq 30$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。

<u>$k=6$ のとき</u>

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

$\eta(\mu) = 0.9531$ （Ogawaの場合 0.9559）

<u>$k=8$ のとき</u>

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

$\eta(\mu) = 0.9693$ （Ogawaの場合 0.9722）

<u>$k=10$ のとき</u>

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

$\eta(\mu) = 0.9781$ （Ogawaの場合 0.9808）

<u>$k=20$ のとき</u>

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

$\eta(\mu) = 0.9927.$

§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て。

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

$k=2k'$（k'は$k'\geq 1$なる整数）の場合を考える。

(i)　σの推定量として，§2と同様な発想で，類推として，線形推定量（$\hat{\sigma}$）を考える。

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

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，以下のものを考える。

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

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($\hat{\hat{\sigma}}^2$)を考える。

$$\hat{\hat{\sigma}}^2 = (\underset{\sim}{X}_{()} - \underset{\sim}{\mu})' B (\underset{\sim}{X}_{()} - \underset{\sim}{\mu})$$

但し, $\underset{(k\times k)}{B} = \| b_{ij} \|$: 対称かつ実対称かつ非負定符号行列.

$\hat{\hat{\sigma}}^2$が不偏性をもつために $\underset{\sim}{u}' B \underset{\sim}{u} = 1$ とする。

こゝで

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

とする。

この場合は

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

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

又 $\lambda_{ij} = \dfrac{\lambda_i (1-\lambda_j)}{n f_i f_j} \quad (i \leqq j)$

相対効率 $\eta_3(\sigma^2)$は

$$\eta_3(\sigma^2) = \frac{2}{n} \cdot \frac{\sigma^4}{V(\hat{\hat{\sigma}}^2)}$$

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,簡便推定量の様に思われるが,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，決定方程式の解きにくさも同様である。$\mu=0$ の場合をあとで例示する。

(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，決定方程式の解きにくさも同様である。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。

<u>$k=2$ のとき</u>

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

$$\eta_2(\sigma^2) = 0.6522 \quad (\text{Ogawa の場合 } 0.653)$$

<u>$k=4$ のとき</u>

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

$$\eta_2(\sigma^2) = 0.8240 \quad (\text{Ogawa の場合 } 0.824)$$

<u>$k=6$ のとき</u>

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

$$\eta_2(\sigma^2) = 0.8926 \quad (\text{Ogawa の場合 } 0.893)$$

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).