UIT Report Vol. 11, No. 2, pp. 175~178, 1980 연구논문집 제11권 제2호 pp. 175~178, 1980

A Study on the Linear Estimation of Scale and Location Parameters by Order Statistics

Yeo, Sung-Chil
Dept. of Industrial Engineering
(Received June 10, 1980)

〈Abstract〉

In many cases, unbiased estimators which are linear combinations of order statistics are best linear unbiased estimators. But some estimators which are linear combinations of order statistics may not be the best unbiased estimators.

Although all estimators of linear combinations of order statistics are less efficient than the best unbiased estimator, there are many advantages in the use of such linear estimators.

A better linear estimator, based on the order statistics, may be obtained by the generalized least-squares estimation procedure.

順序統計量에 의한 規模 및 位置母數의 線型推定에 관한 硏究

余 成 七
產業工學科
(1980. 6. 10 접수)

〈要 約〉

順序統計量의 一次結合으로 이루어진 不偏推定量이 最良線型不偏推定量이 되는 경우는 많이있다. 그러나 順序統計量에 의한 어떤 推定量은 最良不偏推定量이 아닐수가 있다. 비록 順序統計量의 一次結合으로 된 모든 推定量들이 最良不偏推定量보다 有効性이 떨어지지만 그러한 推定量을 사용할때 많은 잇김이 있는 경우가 있다. 本 論文은 順序統計量에 의한 보다 나은 線型推定量이 一般化한 最小自乘推定法으로 얻어질 수 있음을 제시하였다.

I. Introduction

In many cases, a linear combination of order statistics may yield the best linear unbiased estimator (BLUE). For example, the sample mean is a particular linear combination of order statistics. In the case of a rectangular distribution $R(\theta_1, \theta_2)$, $-\infty < \theta_1 < \theta_2 < \infty$, the BLUE's of θ_1 and θ_2 are linear combinations of the exteme statistics $X_{(1)}$ and $X_{(n)}$. While linear combinations of order statistics are not in many other cases best unbiased estimators. This is the case, for instance, when we are estimating the standard deviation σ of a normal distribution $N(\mu, \sigma^2)$. The best unbiased estimator(BUE) of σ, based on a sample $X_1, \cdots, X_n (n \geq 2)$, is $\hat{\sigma}_r = C_n \cdot [\sum_{i=1}^{n} (X_i - \bar{X})^2]^{1/2}$, where $C_n = \Gamma[(n-1)/2] \sqrt{2}\,\Gamma(n/2)$.

In the normal case, an unbiased estimator of σ can be obtained faster as a simple function of the range, which is $\tilde{\sigma}_n = (X_{(n)} - X_{(1)})/d_n$, where $d_n (n \geq 2)$ is the expectation of the range of a random sample from the standard normal distribution $N(0, 1)$. This estimator is in common use especially in the field of quality control.

The estimator $\tilde{\sigma}_n$ is less efficient (large variance) than $\hat{\sigma}_n$, and its efficiency declines rapidly as the sample size n increases. This is due to the fact that $\tilde{\sigma}_n$ is not a function of the sufficient statistics($\Sigma_{i=1}^n X_i$, $\Sigma_{i=1}^n X_i^2$), and the amount of information in the sample that it discards is considerable. A better linear estimator of σ, based on the order statistics, may be obtained by a method which will be discussed presently.

Although all estimators of σ which are linear combinations of order statistics are less efficient than the best unbiased estimator $\hat{\sigma}_n$, there are many advantages in the use of such linear estimators, especially in large sample situations. A good discussion of the usefulness of such less efficient estimators is provided by Mosteller [6].

In this paper we give the results of the generalized least-squares estimation procedure for estimating scale and location parameters by linear combinations of order statistics, and present a simple example of such a linear estimation for the case of an exponential distribution.

II. The Generalized Least-Squares Estimation by Order Statistics

Let $\mathscr{F}$ be a class distributions depending on scale and location parameters only. That is, every distribution function in $\mathscr{F}$ is of the form $F[(x-\mu)/\sigma]$, $-\infty<\mu<\infty$, $0<\sigma<\infty$. The quantity μ designates the location parameter and σ the scale parameter. The distribution function of the standardized random variable $U=(X-\mu)/\sigma$ is $F(\mu)$, which is of a known functional form. Let $X_{(1)}\leq X_{(2)}\leq\cdots\leq X_{(n)}$ be the order statistics and $U_{(r)}=(X_{(r)}-\mu)/\sigma$. Basic quantities for the estimation of μ and σ are the parameters

$$\alpha_r=E(U_{(r)}),\quad r=1,2,\cdots,n,$$

$$v_{rs}=\mathrm{Cov}(U_{(r)},U_{(s)}),\quad r,s=1,2,\cdots,n.$$

Let $\boldsymbol{\alpha}=(\alpha_1,\cdots\alpha_n)$, $V=\|v_{rs}\,;\,r,s=1,2\cdots,n\|$ be the covariance matrix and $X'=(X_{(1)},\cdots,X_{(n)})$, where $v_{rs}=v_{sr}$ by symmetry of V. Then we have the linear model

$$\begin{aligned} X&=\mathbf{1}_n\mu+\boldsymbol{\alpha}\sigma+e\\ &=(\mathbf{1}_n\ \boldsymbol{\alpha})\begin{pmatrix}\mu\\ \sigma\end{pmatrix}+e \qquad (2.1)\end{aligned}$$

where $\mathbf{1}_n'=(1,1,\cdots,1)$, $e'=(\varepsilon_1\cdots,\varepsilon_n)$, $E(e)=\mathbf{0}$ and $E(ee')=V\sigma^2$. Here we assume that V is positive definite (it is always non-negative definite). Then there exists a nonsingular $n\times n$ matrix P such that $PP'=V$. If we pre-multiply both sides of the model (2.1) by a matrix P^{-1}, then the model (2.1) is tranformed into the model such that

$$\begin{aligned} X^*&=\mathbf{1}_n^*\sigma+\boldsymbol{\alpha}^*\sigma+e^{*}{}^{(1)}\\ &=(\mathbf{1}_n^*\ \boldsymbol{\alpha}^*)\begin{pmatrix}\mu\\ \sigma\end{pmatrix}+e^* \qquad (2.2)\end{aligned}$$

, where $E(e^*)=0$ and $E(e^*e^{*\prime})=\sigma^2 I$

Hence the model(2.2) satisfies the usual assumptions of an ordinary least-squares model. Now if we apply the least-squares estimation procedures to the transformed model (2.2), then we have the least-squares estimator of (μ,σ) such that

$$\begin{pmatrix}\mu\\ \sigma\end{pmatrix}=\begin{bmatrix}\mathbf{1}_n'V^{-1}\mathbf{1}_n' & \mathbf{1}_n'V^{-1}\boldsymbol{\alpha}\\ \boldsymbol{\alpha}'V^{-1}\mathbf{1}_n & \boldsymbol{\alpha}'V^{-1}\boldsymbol{\alpha}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{1}_n'V^{-1}X\\ \boldsymbol{\alpha}'V^{-1}X\end{bmatrix}$$

$$=\frac{1}{\xi}\begin{bmatrix}(\boldsymbol{\alpha}'V^{-1}\boldsymbol{\alpha})(\mathbf{1}_n'V^{-1}X)-(\boldsymbol{\alpha}'V^{-1}\mathbf{1}_n)(\boldsymbol{\alpha}'V^{-1}X)\\ (\mathbf{1}_n'V^{-1}\mathbf{1}_n)(\boldsymbol{\alpha}'V^{-1}\mathbf{1}_n)-(\mathbf{1}_n'V^{-1}\boldsymbol{\alpha})(\mathbf{1}_n'V^{-1}X)\end{bmatrix}$$

, where $\xi=(\mathbf{1}_n'V^{-1}\mathbf{1}_n)(\boldsymbol{\alpha}'V^{-1}\boldsymbol{\alpha})-(\mathbf{1}_n'V^{-1}\boldsymbol{\alpha})^2$.

If we let

$$D=\frac{1}{\xi}[V^{-1}(\mathbf{1}_n'\boldsymbol{\alpha}-\boldsymbol{\alpha}\mathbf{1}_n')V^{-1}]$$

, then the least-squares estimators of μ and σ are

$$\hat{\mu}=-\boldsymbol{\alpha}'DX$$

$$\hat{\sigma}=\mathbf{1}_n'DX$$

Since the covariance matrix Σ of the estimator $(\hat{\mu},\hat{\sigma})$ is

$$\begin{aligned}\Sigma&=\sigma^2\begin{bmatrix}\mathbf{1}_n'V^{-1}\mathbf{1}_n & \mathbf{1}_n'V^{-1}\boldsymbol{\alpha}\\ \mathbf{1}_n'V^{-1}\boldsymbol{\alpha} & \boldsymbol{\alpha}'V^{-1}\boldsymbol{\alpha}\end{bmatrix}^{-1}\\ &=\frac{\sigma^2}{\xi}\begin{bmatrix}\boldsymbol{\alpha}'V^{-1}\boldsymbol{\alpha} & -\mathbf{1}_n'V^{-1}\boldsymbol{\alpha}\\ -\mathbf{1}_n'V^{-1}\boldsymbol{\alpha} & \mathbf{1}_n'V^{-1}\mathbf{1}_n\end{bmatrix}\end{aligned}$$

we see that

$$\mathrm{Var}(\hat{\mu})=\frac{\sigma^2}{\xi}(\boldsymbol{\alpha}'V^{-1}\boldsymbol{\alpha}),$$

$$\mathrm{Var}(\hat{\sigma})=\frac{\sigma^2}{\xi}(\mathbf{1}_n'V^{-1}\mathbf{1}_n),$$

(1) Here we have set $X^*=P^{-1}X$, $\mathbf{1}_n^*=P^{-1}\mathbf{1}_n$, $\boldsymbol{\alpha}^*=P^{-1}\boldsymbol{\alpha}$ and $e^*=P^{-1}e$

and $\mathrm{Cov}(\hat{\mu},\hat{\sigma}) = -\frac{\sigma^2}{\xi}(\mathbf{1}_n' V^{-1}\boldsymbol{\alpha})$.

According to the Gauss-Markov theorem, we conclude that the estimators are the best linear combinations of the order statistics.

Ⅲ. An Illustrative Example

Let $X_1, X_2, \cdots, X_n$ be independent identically distributed ($i.i.d.$) random variables having a 2-parameter exponential distribution, that is,

$$F\left(\frac{x-\mu}{\sigma}\right) = \begin{cases} 0, & \text{if } x<\mu, \\ 1-\exp\left\{-\frac{x-\mu}{\sigma}\right\} & \text{if } x \geq \mu, \end{cases}$$

where $-\infty<\mu<\infty$, $0<\sigma<\infty$; μ designates the location parameter and σ the scale parameter. Let $X_{(1)} \leq \cdots \leq X_{(n)}$ be the order statistics of the given sample and $U_{(r)} = (X_{(r)}-\mu)/\sigma$.

We prove first that

$$\alpha_r = E(U_{(r)}) = \sum_{i=1}^{r} (n-i+1)^{-1}, \quad r=1,2,\cdots,n \tag{3.1}$$

and then establish that, for all $1 \leq r < s \leq n$,

$$\mathrm{Var}(U_{(r)}) = \mathrm{Cov}(U_{(r)}, U_{(s)}) = \sum_{i=1}^{r}(n-i+1)^{-2}. \tag{3.2}$$

Consider the exponential distribution $F(x) = 1-e^{-x}$ for $x \geq 0$. Suppose that the random variable Z, having the distribution $F(x)$, respresents an observation on the lifetime of a certain apparatus. Let $Z_1, \cdots, Z_n$ be $i.i.d.$ random variables, representing the lifetime of n such apparatuses. Let $U_{(1)}, U_{(2)}, \cdots, U_{(n)}$ be the corresponding order statistics, representing the time of first failure, second failure, and so on. Since $Z_1, \cdots, Z_n$ are $i.i.d.$, having a distribution $F(x)$, the distribution function of the first failure time $U_{(1)}$ is $F(nx)$ and hence $E(U_{(1)}) = 1/n$ and $\mathrm{Var}(U_{(1)}) = 1/n^2$.

Since the stochastic process governing the failure time points is a Poisson process with a mean of one failure per time unit, the distribution of $U_{(2)} - U_{(1)}$ is again exponential with the expectation of

$$E(U_{(2)} - U_{(1)}) = (n-1)^{-1},$$

and

$$\mathrm{Var}(U_{(2)} - U_{(1)}) = (n-1)^{-2},$$

Furthermore, since $U_{(1)}$ and $U_{(2)} - U_{(1)}$ are independent we have

$$E(U_{(2)}) = n^{-1} + (n-1)^{-1},$$

and

$$\mathrm{Var}(U_{(2)}) = n^{-2} + (n-1)^{-2}.$$

In a similar manner (3.1) and the first part of (3.2) are proven for any $r=1,\cdots,n$. Finally, to prove that $\mathrm{Cov}(U_{(r)}, U_{(s)}) = \mathrm{Var}(U_{(r)})$ for all $r=s$, we write $U_{(s)} = U_{(r)} + (U_{(s)} - U_{(r)})$. Since the increment $U_{(s)} - U_{(r)}$ is independent of $U_{(r)}$ the required result is proven. Thus

$$\boldsymbol{\alpha} = \left(n^{-1}, n^{-1}+(n-1)^{-1}, \cdots, \sum_{i=1}^{r}(n-i+1)^{-1}, \cdots, \sum_{i=1}^{n}(n-i+1)^{-1}\right)$$

and

$$V = \begin{bmatrix} v_{11} & v_{11} & \cdots & v_{11} \\ \cdot & v_{22} & \cdots & v_{22} \\ & & \ddots & \\ \cdot & \cdot & & v_{nn} \end{bmatrix}$$

, where $v_{rs} = \Sigma_{i=1}^{r}(n-i+1)^{-2}$ for all $r \leq s$, $r=1,2,\cdots,n$, and $v_{rs} = v_{sr}$ Since the inverse of V is (2) V^{-1}

$$= \begin{bmatrix} (n-1)^2+n^2 & -(n-1)^2 & & \bigcirc \\ -(n-1)^2 & (n-2)^2+(n-1)^2 & -(n-2)^2 & \\ & -(n-2)^2 & \ddots & -1 \\ \bigcirc & & \ddots & 1 \\ & & -1 & \end{bmatrix}$$

we have that $\mathbf{1}_n' V^{-1} = (n^2, \mathbf{0}'_{n-1})$ and $\boldsymbol{\alpha}' V^{-1} = \mathbf{1}_n'$. From this fact we see that

$$\xi = (\mathbf{1}_n' V^{-1}\mathbf{1}_n)(\boldsymbol{\alpha}' V^{-1}\boldsymbol{\alpha}) - (\mathbf{1}_n' V^{-1}\boldsymbol{\alpha})^2$$
$$= n^2(n-1) \text{ and}$$

$$D = \frac{1}{\xi}\left[(V^{-1}\mathbf{1}_n)(\boldsymbol{\alpha}' V^{-1}) - (V^{-1}\boldsymbol{\alpha})(\mathbf{1}_n' V^{-1})\right]$$

$$= \frac{1}{n^2(n-1)} \begin{bmatrix} 0 & n^2 \cdots n^2 \\ -n^2 & \\ \vdots & \bigcirc \\ -n^2 & \end{bmatrix}$$

Thus we obtain that

$$\hat{\mu} = -\boldsymbol{\alpha}' D X = \frac{nX_{(1)} - \overline{X}}{n-1}$$

$$\hat{\sigma} = \mathbf{1}_n' D X = \frac{n\overline{X} - X_{(1)}}{n-1}$$

The covariance matrix Σ of the estimators $\hat{\mu}$

(2) See Graybill [3] pp. 168—169.

and $\hat{\sigma}$ is

$$\Sigma=\frac{\sigma^2}{\xi}\begin{bmatrix}\boldsymbol{\alpha}'V^{-1}\boldsymbol{\alpha} & -\mathbf{1}_n'V^{-1}\boldsymbol{\alpha}\\ -\mathbf{1}_n'V^{-1}\boldsymbol{\alpha} & \mathbf{1}_n'V^{-1}\mathbf{1}_n\end{bmatrix}$$

$$=\sigma^2\begin{bmatrix}\dfrac{1}{n(n-1)} & \dfrac{-1}{n(n-1)}\\ \dfrac{-1}{n(n-1)} & \dfrac{1}{n-1}\end{bmatrix}$$

Hence we have that

$$\mathrm{Var}(\hat{\mu})=\frac{\sigma^2}{n(n-1)},\quad \mathrm{Var}(\hat{\sigma})=\frac{\sigma^2}{n-1},$$

and $\mathrm{Cov}(\hat{\mu},\hat{\sigma})=-\dfrac{\sigma^2}{n(n-1)}$.

In the above example we see that $(X_{(1)},\overline{X})$ is a minimal sufficient statistic for the 2-parameter exponential family.(3 Hence by the Rao-Blackwell theorem, $\hat{\mu}$ and $\hat{\sigma}$ are not only BLUE's but also BUE's.

Ⅳ. Conclusion

We have applied the least-squares theory for estimating scale and location parameters by linear combinations of order statistics. The advantage of estimators based on order statistics is especially great in situations where trimming or censoring of observations in the extremes is part of the experimental model (e.g., in life testing experiments). The efficiency of such estimation procedures, when the distributions are normal, exponential, gamma, rectangular, etc., has been studied and discussed by Greenberg and Sarhan[4].

References

1. Çinlar, E., Introduction to Stochastic Process, Prentice Hall, Inc., Englewood Cliffs, N.J., 1975.
2. Graybill, F.A., An Introduction to Linear Statistical Models, Vol.1, McGraw-Hill: New York, 1961.
3. Graybill, F.A., Introduction to Matrices with Applications in Statistics, Wadsworth: Belmont, California, 1969.
4. Greenberg, B.G., Sarhan, A.E., Matrix inversion, its interest and application in analysis of data, *Jour. Amer. Statist.* 51 : 755—766(1959).
5. Mood, A.M., Graybill, F.A., Boes, D.C., Introduction to the Theory of Statistics, Third Edition, McGraw-Hill: New York, 1974.
6. Mosteller, F., On some useful "inefficient" statistics, *Ann. Math. Statist.*, 17 : 377—408 (1946).
7. Seber, G.A.F., Linear Regression Analysis, Wiley: New York, 1977.
8. Zacks, S., The Theory of Statistical Inference, Wiley: New York, 1971.

(3) See Zacks[8] pp.31—32.