Jour. Ind. Soc. Ag. Statistics
43 (1), 1991 : 40–48

FURTHER REMARKS ON SEQUENTIAL POINT ESTIMATION OF REGRESSION PARAMETERS

Anoop Chaturvedi
University of Allahabad, Allahabad
and
Ajit Chaturvedi
University of Jammu, Jammu(Tawi)
(Received : May, 1988)

Summary

The sequential procedure developed by Chaturvedi [2] for estimating the regression parameters in a linear model is further analyzed. Much simpler proof is provided for the asymptotic 'risk-efficiency' and second-order approximations are obtained for the expected sample size and 'regret' associated with the sequential procedure. The problem of bounded risk point estimation is also discussed.

Key words : Linear model, point estimation, stopping times, risk- efficiency, regret, second-order approximations.

Introduction

Consider the linear model

$$y_i = \underline{x}_i' \underline{\beta} + \varepsilon_i \ (i = 1, 2, \dots)$$

Where $\underline{x}_i$ is pxl vector of known constants, $\underline{\beta}$ is the pxl vector of unknown parameters, and ε_i's are disturbance terms, independent and normally distributed with mean 0 and variance σ^2. Having recorded $y_1, \dots, y_n$ on $\underline{x}_1, \dots, \underline{x}_n$, respectively, let $x_n = (\underline{x}_1 \dots \underline{x}_n)'$, $\underline{Y}_n = (y_1 \dots y_n)'$ Use the usual least squares estimator $\hat{\underline{\beta}}_n = (x'_n x_n)^{-1} x'_n \underline{y}_n$ to estimate $\underline{\beta}$.

Let the loss incurred in estimating $\underline{\beta}$ by $\hat{\underline{\beta}}_n$ be

$$L(\underline{\beta}, \hat{\underline{\beta}}_n) = A \left[\frac{1}{n} (\hat{\underline{\beta}}_n - \underline{\beta})' (x'_n x_n) (\hat{\underline{\beta}}_n - \underline{\beta}) \right]^{\alpha/2} + Cn^t, \tag{1.1}$$

Where A, α, C and t are known positive constants. The risk corresponding to the loss function (1.1) is

$$v_n(\sigma) = K(p, \alpha)\frac{\sigma^\alpha}{n^{\alpha/2}} + Cn^t \qquad (1.2)$$

where $K(p, \alpha) = A.2^{\alpha/2}\,\Gamma[(p+\alpha)/2]/\Gamma(p/2)$. For known σ, the sample size which minimizes $v_n(\sigma)$ is the smallest positive integer $n \geq n_o$, where

$$n_o = (\alpha/2Ct)\,K(p, \alpha)\,\sigma^\alpha]^{2/(2t+\alpha)} \qquad (1.3)$$

and setting $n=n_o$ in (1.2), the corresponding minimum risk is

$$v_{n_o}(\sigma) = C(1 + 2t/\alpha)\,n_o^t \qquad (1.4)$$

But, in the ignorance of σ, no fixed sample size procedure minimizes $v_n(\sigma)$ simultaneously for all values of σ. In such a situation, motivated by (1.3), adopt the following sequential procedure.

Let us define, for

$n \geq p + 1$, $\hat{\sigma}_n^2 = (n-p)^{-1}\, \underline{Y}'_n\,[I_n - X_n(X'_n X_n)^{-1} X'_n]\,\underline{Y}_n$ as the estimator for

σ^2, where I_k denotes a kxk identity matrix. Then, the stopping time $N=N(\sigma)$ is given by

$$N = \inf[\, n \geq m : n \geq \{(\alpha/2Ct)\,K(p, \alpha)\,\hat{\sigma}_n^\alpha\}^{2/(2T+\alpha)}\,] \qquad (1.5)$$

where $m(\geq p+1)$ being the initial sample size. When stop, estimate $\underline{\beta}$ by $\underline{\hat{\beta}}_N$

As in Starr [6] and Starr and Woodroofe [7], the 'risk- efficiency' and 'regret' of the sequential procedure (1.5) are defined, respectively, by

$$\eta(\sigma) = \bar{v}(\sigma)/v_{n_o}(\sigma)$$

and

$$\omega(\sigma) = \bar{v}(\sigma) - v_{n_o}(\sigma) \qquad (1.7)$$

where $\bar{v}(\sigma)$ is the risk associated with the sequential procedure, i.e.

$$\bar{v}(\sigma) = K(p,\alpha)\sigma^\alpha E(N^{-\alpha/2}) + CE(N^t)$$

$$= (2Ct/\alpha) n_o^{(t+\alpha/2)} E(N^{-\alpha/2}) + CE(N^t) \tag{1.8}$$

Chaturvedi [2], using much complicated algebra and many probability inequalities, obtained a condition on the starting sample size m, which ensures the asymptotic (as $\sigma \to \infty$) 'risk-efficiency' of the sequential procedure (1.5) and proved that $\lim_{\sigma \to \infty} \eta(\sigma) = 1$ if $m > p + \alpha^2/(\alpha + 2t)$ He also obtained first-order asymptotics for the 'regret' and proved that, for $t = 1$, $\lim_{\sigma \to \infty} \omega(\sigma) = O(1)$ if and only if $m \geq p + \alpha$. For $\alpha=2$ and $t=1$, i.e., when the loss is quadratic plus linear cost of sampling, Chaturvedi [2] derived second-order approximations for the 'regret' and proved that $\lim_{\sigma \to \infty} \omega(\sigma) = 1 + O(1)$ for $m \geq p+1$.

The purpose of this note is to obtain a much simpler proof of the asymptotic 'risk-efficiency' and second-order approximations for the 'regret'. In section 2, a condition on the initial sample size is determined which could guarantee asymptotic 'risk- efficiency'. In section 3, improving the bounds for 'regret' obtained by Chaturvedi [2], second order approximations are achieved. Finally, in section 4, the problem of bounded risk point estimation of $\underline{\beta}$ is discussed.

2. *Asymptotic Risk-Efficiency*

We first establish three lemmas.

Lemma 1

For any $\lambda(>0)$ fixed and $m \geq p + 1$, $\lim_{\sigma \to \infty} E(N/n_o)^{\lambda} = 1$

Proof: If follows from the definition (1.5) of N that

$$\left\{(\alpha/2Ct) K(p,\alpha)\, \hat{\sigma}_N^{\alpha}\right\}^{2/(2t+\alpha)} \leq N \leq \left\{(\alpha/2Ct) K(p,\alpha)\, \hat{\sigma}_N^{\alpha}\right\}^{2/(2t+\alpha)} + (m-1)$$

or,

$$(\hat{\sigma}_N/\sigma)^{2\alpha/(2t+\alpha)} \leq (N/n_0) \leq (\hat{\sigma}_N/\sigma)^{2\alpha/(2t+\alpha)} + (m-1)/(n_o)$$

which on using the facts that $\{\lim_{\sigma \to \infty} N = \infty$, $\lim_{N \to \infty} = \sigma$ a.s. and

$\lim_{\sigma \to \infty} n_o = \infty$, leads us to the result that

$$\lim_{\sigma \to \infty} (N/n_o) = 1 \text{ a.s.} \tag{2.1}$$

Using the fact that [see, Judge, and Bock (1978, p.20, Theorem A.2.16) for proof] $(n-p)\hat{\sigma}_n^2 = \sum_{j=1}^{n-p} Z_j$, with $Z_j \sim \chi^2(1)$, we obtain from (1.5),

$$(N/n_o)^\lambda \leq [(\hat{\sigma}_n^2/\sigma^2)^{\alpha/(2t+\alpha)} + (m-1)/(n_o)]^\lambda$$

$$= [\{(n-p)^{-1} \sum_{j=1}^{n-p} Z_j\}^{\alpha/(2t+\alpha)} + (m-1)/(n_o)]^\lambda$$

It follows from Wiener ergodic theorem (see, Khan [5]) that

$$E\left[\sup_{n \geq p+1} \left\{(n-p)^{-1}\sum_{j=1}^{n-p} Z_j\right\}^{\alpha\lambda/(\alpha+2t)}\right] < \infty.$$

Thus, $(N/n_0)^\lambda$ is integrable and the lemma follows from (2.1) and dominated convergence theorem.

Lemma 2 (Hayre [3])

Let $Z_1, Z_2, \ldots$ be i.i.d. Chi-squared with one degree of freedom, and let

$$\eta_k = \inf_{n \geq k} \left\{n^{-1}(Z_1 + Z_2 + \ldots + Z_n)\right\}$$

Then, for $1 > 0$, $E(\eta_k^{-1}) < \infty$ if and only if $k > 21$.

Lemma 3

For $\lambda\ (>0)$ fixed, $\lim_{\sigma \to \infty} E(n_0/N)^\lambda = 1$ if $m > p + 2\alpha\lambda/(\alpha+2t)$

Proof: It follows from the definition (1.5) of N that

$$(n_0/N)^\lambda \leq (\sigma/\hat{\sigma}_n)^{2\alpha\lambda/(\alpha+2t)}$$

$$= [(n-p)^{-1} \sum_{j=1}^{n-p} Z_j]^{-\alpha\lambda/(\alpha+2t)}$$

Since $n \geq m$ applying Lemma 2, we conclude that $E(n_0/N)^\lambda < \infty$ if $m > p+2\alpha\lambda/(\alpha+2t)$. Hence, $(n_0/N)^\lambda$ is integrabale for all $m > p+2\alpha\lambda/(\alpha+2t)$ and the lemma follows from (2.1) and dominated convergence theorem.

The main result of this section is stated in the following theorem, which provides a condition on starting sample size ensuring asymptotic "risk-efficiency".

Theorem 1

$$\lim_{\sigma \to \infty} \eta(\sigma) = 1, \quad \text{if} \quad m > p + \alpha^2/(\alpha+2t)$$

Proof : Making substitutions from (1.4) and (1.8) in (1.6), we obtain after some algebra,

$$\eta(\sigma) = (1+2t/\alpha)^{-1} [E(N/n_0)^t + (2t/\alpha) E(n_0/N)^{\alpha/2}]$$

The proof is now an immediate consequence of Lemmas 1 and 3.

Remark 1: It is concluded that the condition obtained on m in Theorem 1 is consistent with that obtained in Theorem 1 of Chaturvedi [2]. However, here we have not studied the behaviour of $\eta(\sigma)$ for the cases when (i) $m = p+\alpha^2/(\alpha+2t)$ and (ii) $m < p+\alpha^2/(\alpha+2t)$, but one may not be interested in the situations when the asymptotic risk-efficiency is not achieved.

3. *Second-Order Approximations for the Regret*

The following theorem provides-second-order approximations for the expected sample size associated with the sequential procedure (1.5).

Theorem 2 :

For t=1 and $m > p + 2\alpha/(\alpha+2)$, as $\sigma \to \infty$,

$$E(N) = n_0 + \frac{\alpha}{(\alpha+2)} \{\nu - 2(\alpha+1)/(\alpha+2)\} - p + O(1)$$

Proof: For t=1, the stopping rule (1.5) may be re-written as

$$N = \inf [n \geq m : \sum_{j=1}^{n-p} Z_j \leq n_0^{-(1+2\alpha^{-1})} n^{(1+2\alpha^{-1})} (n-p)] \quad (3.1)$$

Let us define a new stopping time t_σ by

$$t_\sigma = \inf[n \geq m-p : \sum_{j=1}^{n} Z_j \leq n_0^{-(1+2\alpha^{-1})} n^{2(1+\alpha^{-1})} (1+pn^{-1})^{(1+2\alpha^{-1})}] \quad (3.2)$$

Following the proof of Lemma 1 in Swanepoel and van Wyk [8], it can be shown that the steeping rules (3.1) and (3.2) follow the same probability distribution. From (3.2) and equation (1.1) of Woodroofe [9], we obtain in the notations of Woodroofe [9],

$$m = m-p, \; S_n + \sum_{j=1}^{n} Z_j, \; c = n_0^{-(1+2\alpha^{-1})}, \; \alpha = 2(1+\alpha^{-1})$$

$$L(n) = (1+pn^{-1})^{(1+2\alpha^{-1})}, \ \beta = (1+2\alpha^{-1})^{-1}, \ L_0 = p(1+2\alpha^{-1}), \ a = \tfrac{1}{2}.$$

Now, from Theorem 2.4 of Woodroofe [9], we obtain for all $m > p+2\alpha/(\alpha+2)$, as $\sigma \to \infty$,

$$E(t_\sigma) = n_0 + \alpha\nu/(\alpha+2) - p-2(\alpha+1)/(\alpha+2)^2 + O(1),$$

and the theorem follows. Here, ν is given by Theorem 2.2 of Woodroofe [9].

In the following theorem, we shall obtain second-order approximations for the "regret".

Theorem 3 : For $t = 1$ and $m > p+2\alpha/(\alpha+2)$, as $\sigma \to \infty$

$$\omega(\sigma) = C\alpha^2/2(\alpha+2) + O(1)$$

Proof : From (1.4) and (1.8), substituting the values of $v_{n_0}(\sigma)$ and $\bar{v}(\sigma)$ in (1.7), we get for $t = 1$,

$$\omega(\sigma) = (2C/\alpha) n_0^{(1+\alpha/2)} \ E(N^{-\alpha/2} - n_0^{-\alpha/2}) + CE(N-n_0) \tag{3.3}$$

Expanding $N^{-\alpha/2}$ about no by Taylor series expansion, we obtain for $|U - n_0| \le |N - n_0|$

$$\omega(\sigma) = (2C/\alpha) \, n_0^{(1+\alpha/2)} E\left[-(\alpha/2)(N-n_0)n_0^{-(1+\alpha/2)} + (1/2)(N-n_0)^2(\alpha/2)(\alpha/2+1) U^{-(\alpha/2+2)}\right] + CE(N-n_0)$$

$$= \{C(\alpha+2)/4n_0\} \, E\left[(N-n_0)^2 (n_0/U)^{(\alpha/2+2)}\right]$$

Denoting by P, the c.d.f. of N, we can write

$$\omega(\sigma) = I_1 + I_2 \tag{3.4}$$

where

$$I_1 = \{C(\alpha+2)/4n_0\} \int_{N \le n_0/2} (N-n_0)^2 (n_0/U)^{(\alpha/2+2)} \, dp$$

and

$$I_2 = \{C(\alpha+2)/4n_0\} \int_{N > n_0/2} (N-n_0)^2 (n_0/U)^{(\alpha/2+2)} \, dp$$

Since $(n_0/U) \xrightarrow{a.s.} 1$ as $\sigma \to \infty$, for sufficiently large σ, we have $(n_0/U)^{(\alpha/2+2)} \le K$, where K is any generic constant independent of σ. From Corollary 2 of Chaturvedi [2], $P(N \le n_0/2) = O(\sigma^{-(m-p)})$, as $\sigma \to \infty$. Thus

$$I_1 = O\left(\sigma^{2\alpha/(\alpha+2)+p-m}\right)$$

$$= O(1) \tag{3.5}$$

for all $m > p + 2\alpha/(\alpha+2)$. It follows from a result of Bhattacharya and Mallik [1] that the asymptotic distribution of $(N-n_0)/n_0^{1/2}$ is $N[0, 2\alpha^2/(\alpha+2)^2]$. Moreover, from Theorem 2.3 of Woodroofe [9], $(N-n_0)^2/n_0$ is uniformly integerable for all $m > p+2\alpha/(\alpha+2)$. Hence, we obtain for all $m > p+2\alpha/(\alpha+2)$, as $\sigma \to \infty$

$$I_2 = C\alpha^2/2(\alpha+2) \tag{3.6}$$

The theorem now follows on making substitutions from (3.5) and (3.6) in (3.4).

4. *Bounded Risk Point Estimation of* $\underline{\beta}$

Let the loss of estimating $\underline{\beta}$ by $\hat{\underline{\beta}}_n$ is

$$L^*(\underline{\beta}, \hat{\underline{\beta}}_n) = A\left[\frac{1}{n}(\hat{\underline{\beta}}_n - \underline{\beta})'(X'_n X_n)(\hat{\underline{\beta}}_n - \underline{\beta})\right]^{\alpha/2} \tag{4.1}$$

The risk associated with the loss (4.1) is

$$v_n^*(\sigma) = K(p,\alpha)\,\sigma^\alpha/n^{\alpha/2} \tag{4.2}$$

Here, A,a and K(p,a) are same as that defined in Section 1. For specified $W(>0)$, suppose one wishes that the risk (4.2) should not exceed W. It is easy to see that, for known σ, the sample size needed to achieve the goal is the smallest positive integer $n \ge n^*$, where $n^* = \{K(p,\alpha)/W\}^{2/\alpha}\sigma^2$. In the absence of any knowledge about σ, we adopt the following sequential procedure.

The stopping time is defined by

$$N = \inf[n \ge m : n \ge \{K(p,\alpha)\,|\,W\}^{2/\alpha}.\hat{\sigma}_n^2] \tag{4.3}$$

Estimate $\underline{\beta}$ by $\hat{\underline{\beta}}_N$

For the sequential procedure (4.3), we state the following theorem, the proof of which can be obtained exactly along the lines of that of Theorem 2 after necessary modifications at various places.

Theorem 4 : For all m> p+2, as $\sigma \to \infty$,

$$E(N) = n^* + v^* - p - 2 + 0(1),$$

where v^* is specified.

The following theorem gives second-order approximations for the expected loss of the sequential procedure (4.3).

Theorem 5: For all m > p+2, as $\sigma \to \infty$

$$E[\,L^*(\underline{\beta}\ \hat{\underline{\beta}}_N)\,] = W[1-(\alpha/2n^*)\left\{v^* - p - (\alpha+2)^2/2\right\}] + O(1)$$

Proof: Expanding $N^{-\alpha/2}$ about n^* by Taylor series, we get

For $\quad |U-n^*| \le |N-n^*|$

$$E[\,L^*(\underline{\beta}, \hat{\underline{\beta}}_N)\,]$$

$$= WE(n^*/N)^{\alpha/2}$$

$$= W\left[1-(\alpha/2n^*)\,E\,(N-n^*) + \left\{\alpha(\alpha+2)/8n^*\right\} E\left\{\frac{(N-n^*)^2}{n^*}\right\}\left(\frac{n^*}{U}\right)^{(\alpha/2+2)}\right]$$

Now using Theorem 4, the results $(N-n^*)/(n^{1/2}) \xrightarrow{\mathcal{L}} N(0,2)$ as $\sigma \to \infty$, $(N-n^*)^2/n^*$ is uniformly integrable for all m > p+2 and the arguments similar to those in the proof of Theorem 3, we can obtain the desired results. The details are omitted for brevity.

ACKNOWLEDGEMENT

We are thankful to the refree for his comments.

REFERENCES

[1] Bhattacharya, P.K. and Mallik, A., 1973. Asymptotic normality of the stopping times of some sequential procedures. *Ann. Statist.* 1, 1203-1211.

[2] Chaturvedi, A., 1987. Sequential point estimation of regression parameters in a linear model. *Ann. Inst. Statist. Math.*, A39, 55-67.

[3] Hayre, L.S., 1983. Sequential estimation of the difference between the means of two normal

populations. *Metrika*, **30**, 101-108.

[4] Judge, George, G. and Bock, M.E. 1978. The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics. *North-Holland Publishing Company, New York.*

[5] Khan, R.A., 1969. A general method of determining fixed- width confidence intervals. *Ann. Math. Statist.*, **40**, 704-709.

[6] Starr, N, 1966. On the asymptotic efficiency of a sequential procedure for estimating the mean. *Ann. Math. Statist.*, **37**, 1173- 1185.

[7] Starr, N. and Woodroofe, M., 1969. Remarks on sequential point estimation. *Proc. Nat. Acad. Sci., USA*, **63**, 285-288.

[8] Swanepoel, J.W.H. and van Wyk, J.W.J., 1982. Fixed width confidence intervals for the location parameter of an exponential distribution. *Commun. Statist.*, **11**(11), 1279-1289.

[9] Woodroofe, M. 1977. Second order approximations for sequential point and interval estimation. *Ann. Statist* **5**, 994- 995.