The Annals of Mathematical Statistics
1971, Vol. 42, No. 4, 1403–1411

ON FIXED-WIDTH CONFIDENCE BOUNDS FOR REGRESSION PARAMETERS[1]

By M. S. Srivastava

University of Toronto

0. Summary. In this paper Chow and Robbins' (1965) sequential theory has been extended to construct a confidence region with prescribed maximum width and prescribed coverage probability for the linear regression parameters under weaker conditions than Srivastava (1967), Albert (1966), and Gleser (1965). An extension to multivariate case has also been carried out.

1. Introduction. Consider $\{y_n\}(n = 1, 2, \cdots)$, a sequence of independent observations with

$$Ey_n' = \beta' X_n, \tag{1.1}$$

and

$$\operatorname{Var} \cdot y_i = \sigma^2, \qquad i = 1, 2, \cdots, n, \tag{1.2}$$

where $\mathbf{y}' = (y_1, \cdots, y_n)$, and $X_n = (\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(n)})$. We assume that X_p is of full rank. (This can be achieved by sampling until p linearly independent $x^{(i)}$ are obtained; this sample size will be denoted by n_0.) The problem is to find a confidence region R in p-dimensional Euclidean space such that $P(\beta \in R) = \alpha$ and such that the maximum diameter of $R \leqq 2d$. Since no fixed sample procedure will meet our requirements, we will consider sequential procedures whose sample size (random variable) N depends on d; $N(d) \to \infty$ a.s. as $d \to 0$. This problem has been considered by Gleser (1965) and Srivastava (1967) under the following generalized Noether's conditions given by Gleser (1965):

(1.3) (i) $\lim_{n\to\infty} n^{-1}(X_n X_n') = \Sigma$,

(ii) $\lim_{n\to\infty} \max_{1 \leqq i \leqq n} \mathbf{x}^{(i)\prime}\mathbf{x}^{(i)} = 0$.

Srivastava (1967) obtained two confidence regions; an ellipsoidal confidence region R_N and a spherical confidence region R_N', $R_N' \supset R_N$, satisfying

(1.3)* (i) $\lim_{d\to 0} P\{\beta \in R_N'\} \geqq \lim_{d\to 0} P\{\beta \in R_N\} = \alpha$;

(ii) the maximum diameter of R_N' (and hence of R_N) $\leqq 2d$.

Received September 30, 1968; revised November 9, 1970.

[1] Research supported by Canada Council.

Albert (1966) reconsidered the problem in more generality. While for spherical regions he also assumes conditions (i) and (ii) of (1.3), but obtains ellipsoidal regions under slightly weaker conditions:

$$\text{(1.4)}\quad \begin{array}{lll} \text{(i)} & \operatorname{tr}(X_nX_n')^{-1} \to 0 & \text{as } n \to \infty, \\ \text{(ii)} & \max_{1 \leq i \leq n}(\mathbf{x}^{(i)\prime}\mathbf{x}^{(i)})\operatorname{tr}(X_nX_n')^{-1} \to 0 & \text{as } n \to \infty, \\ \text{(iii)} & \text{lim-sup}_{n\to\infty}[\lambda_{\max}(X_nX_n')/\lambda_{\min}(X_nX_n')] < \infty, & \\ \text{(iv)} & \lim_{c\to 0}\sup_n |\{\sum_{i=1}^{n(1+c)}(\mathbf{x}^{(i)\prime}\mathbf{x}^{(i)})/\sum_{c=1}^{n}\mathbf{x}^{(i)\prime}x^{(i)}\}-1| = 0, & \\ \text{(v)} & \text{For some}\quad c \geqq 1, \lambda_{c(n+2)}-\lambda_{c(n+1)} \geqq \lambda_{c(n+1)}-\lambda_{cn}. & \\ \text{(vi)} & \lambda_{n-1} \leqq \lambda_n \to \infty \quad\text{and}\quad \lambda_n/\lambda_{n+1} \to 1; & \lambda_n = \lambda_{\min}(X_nX_n'). \end{array}$$

He showed that, *except for* (1.4 v) *and* (1.4 vi), his conditions are weaker than condition (1.3). Without an example it is not clear that the weaker conditions (1.4) permit a larger class of problems to be covered than by (1.3). Given below is an example which shows that an extensive class of problems meet neither the requirements in (1.3), nor in (1.4).

EXAMPLE. Consider the problem of polynomial regression. For convenience of computation, we will consider the case

$$y_t = \alpha + \beta t + \varepsilon_t,$$

$t = 1, 2, \cdots$; where ε_t are independently distributed with mean zero and variance σ^2. Identifying it with the specifications in (1.2), we have $p = 2$, and

$$X_n = \begin{pmatrix} 1 & 1 & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & 2 & \cdot & \cdot & \cdot & \cdot & n \end{pmatrix}.$$

Hence

$$X_nX_n' = \begin{pmatrix} n & n(n+1)/2 \\ n(n+1)/2 & n(n+1)(2n+1)/6 \end{pmatrix}.$$

It is clear that $n^{-1}(X_nX_n') \nrightarrow a$ positive definite matrix as $n \to \infty$. Consequently the conditions (1.3) are not satisfied. In order to see that the conditions (1.4) are also not satisfied, let us find the characteristic roots of X_nX_n'. The two characteristic roots of X_nX_n' are given by

$$\lambda = \{n(2n^2+3n+7) \pm n[(2n^2+3n+7)^2 - 12(n+1)(n-1)]^{\frac{1}{2}}\}/12.$$

Hence

$$\lambda_{\max}(X_nX_n')/\lambda_{\min}(X_nX_n') \to \infty.$$

It is known (see, e.g., Roy (1957)) that X_n can be written as

$$\text{(1.5)}\qquad X_n = T_nL_n$$

where T_n is a $p \times p$ triangular matrix with positive diagonal elements (hence unique), and L_n is a $p \times n$ semiorthogonal matrix; $L_n L_n' = I_p$ and I_p is a $p \times p$ identity matrix. Hence

$$T_n^{-1} X_n = L_n = (l_n^{(1)}, \cdots, l_n^{(n)}) = (\ell_{ij}(n)). \tag{1.6}$$

Let

$$\lambda_n = \lambda_{\min}(X_n X_n'). \tag{1.7}$$

In this paper it is shown that the results hold under the following weaker conditions:

(C^*)

(i) $\lim_{n\to\infty} \max_{1 \leq i \leq n} l_n^{(i)\prime} l_n^{(i)} = 0$,

(ii) $\lim_{c\to 0} \sup_n \|\sum_{i=n+1}^{n^*} l_{n^*}^{(i)} l_{n^*}^{(i)\prime}\| = 0$, where $n^* = [n(1+c)]$, the smallest integer greater than $n(1+c)$, and $\|B\| = [\lambda_{\max}(B)]^{\frac{1}{2}}$.

(iii) (1.4)(v).

First we will show that the condition (C^*) is weaker than the condition (1.4). From (1.6) we get

$$\begin{aligned} l_n^{(i)\prime} l_n^{(i)} &= \lambda_{\max} l_n^{(i)} l_n^{(i)\prime} = \lambda_{\max} T_n^{-1} \mathbf{x}^{(i)} \mathbf{x}^{(i)\prime} T_n^{-1\prime} \\ &\leq [\lambda_{\max} B_n^{-1}][\lambda_{\max} \mathbf{x}^{(i)} \mathbf{x}^{(i)\prime}], \qquad \text{(see [7] page 139)} \\ &\leq \operatorname{tr}(B_n^{-1})(\mathbf{x}^{(i)\prime}\mathbf{x}^{(i)}), \end{aligned}$$

where $B_n = X_n X_n' = T_n T_n'$. Consequently (1.4 ii) implies C^* (i). Next we show that (1.4 iii) and (1.4 iv) imply C^*(ii).

$$\begin{aligned} \|\textstyle\sum_{n+1}^{n^*} l_{n^*}^{(i)} l_{n^*}^{(i)\prime}\| &\leq \textstyle\sum_{n+1}^{n^*} \|l_{n^*}^{(i)\prime} l_{n^*}^{(i)}\| = \textstyle\sum_{n+1}^{n^*} \mathbf{x}^{(i)\prime} B_{n^*}^{-1} \mathbf{x}^{(i)} \\ &\leq \lambda_{\max}(B_{n^*}^{-1}) \textstyle\sum_{n+1}^{n^*} \mathbf{x}^{(i)\prime}\mathbf{x}^{(i)} = [\textstyle\sum_{n+1}^{n^*} \mathbf{x}^{(i)\prime}\mathbf{x}^{(i)}]/\lambda_{\min}(B_{n^*}) \\ &\leq k[(\textstyle\sum_{n+1}^{n^*} \mathbf{x}^{(i)\prime}\mathbf{x}^{(i)})/\operatorname{tr} B_{n^*}] \leq k[(\textstyle\sum_{n+1}^{n^*} \mathbf{x}^{(i)\prime}\mathbf{x}^{(i)})/\operatorname{tr} B_n]. \end{aligned}$$

This completes the proof that (1.4) implies C^*.

We will now verify that the condition C^* is satisfied for the polynomial regression problem. With $X_n X_n' = T_n T_n'$ as given above, we have

$$T_n^{-1} X_n = \begin{pmatrix} n^{-\frac{1}{2}}, \cdots & n^{-\frac{1}{2}} \\ O(n^{-\frac{1}{2}}), \cdots & O(n^{-\frac{1}{2}}) \end{pmatrix},$$

and C^* (i) is satisfied. Also

$$\begin{aligned} \|\textstyle\sum_{i=1}^{n^*} \ell_{n^*}^{(i)} \ell_{n^*}^{(i)\prime}\|^2 &\leq \textstyle\sum_{i=n+1}^{n^*} \ell_{n^*}^{(i)\prime} \ell_{n^*}^{(i)} \leq [nc] \max_{1 \leq i \leq n^*} \ell_{n^*}^{(i)\prime} \ell_{n^*}^{(i)} \\ &= [nc] O(n^{*-1}) \to 0 \qquad \text{as } c \to 0 \end{aligned}$$

uniformly in n. To verify the convexity condition C^* (iii) we proceed as follows.

Let $\lambda_n^* = 12\lambda_n$, $b = (2n^2+3n+7)$, $\varepsilon = 12(n+1)(n-1)/b^2$.

Then

$$\lambda_n^* = nb[1-(1-\varepsilon)^{\frac{1}{2}}] = nb[\tfrac{1}{2}\varepsilon+\tfrac{1}{8}\varepsilon^2-\tfrac{1}{16}\varepsilon^3+\cdots],$$

$$= \frac{6n(n^2-1)}{b}+\frac{3}{2}\frac{n(n^2-1)^2}{b^3}-\frac{3}{4}\frac{n(n^2-1)^3}{b^5}+\cdots = u_1+u_2-u_3+\cdots \quad \text{(say)}.$$

We get

$$u_1''(x) = \frac{6(84x^3+113x^2+322x-21)}{b^3} > 0$$

and

$$u_2''(x) = \frac{3}{2}\frac{8x^7-48x^6-442x^5+210x^4+1000x^3-252x^2+760x+42}{b^5}$$

$$> 0 \qquad \text{for all } x > 15.$$

Hence

$$\lambda_n^{*''} = \frac{6(84n^3+113n^2+322n-21)}{b^3}+\frac{3}{2}\frac{(8n^7-48n^6+\cdots)}{b^5}-O(n^{-5})+O(n^{-7})-\cdots.$$

Consequently there exists an n_0 such that for $n \geqq n_0$, $\lambda_n^{*''} \geqq 0$ which proves the convexity of λ_n.

It may be pointed out that several examples could be constructed satisfying the conditions (C^*) and not satisfying (1.3) and (1.4). Another example is with the design matrix

$$X_n = \begin{pmatrix} 0, & 1, & 0, & 2, & 0, & 3, & 0, & 4, & \cdots \\ 1, & 0, & 2^{\frac{1}{2}}, & 0, & 3^{\frac{1}{2}}, & 0, & 4^{\frac{1}{2}}, & 0, & \cdots \end{pmatrix}.$$

2. Solution. The least squares (l.s.) estimator of β and σ^2 are respectively given by

(2.1) $$\hat{\beta}(n) = T_n^{-1}L_n\mathbf{y}_n = (X_nX_n')^{-1}X_n\mathbf{y}_n$$

and

(2.2) $$\hat{\sigma}^2(n) = n^{-1}\mathbf{y}_n'[I_n-L_n'L_n]\mathbf{y}_n = n^{-1}y_n'[I_n-X_n'(X_nX_n')^{-1}X_n]\mathbf{y}_n.$$

Proceeding as in Srivastava (1967), we obtain regions R_n and R_n' satisfying (i) and (ii) of (1.3)*, as follows:

(I) We start by taking $n_0 \geqq p$ observations $y_1, \cdots, y_{n_0}$. We then sample one extra observation at a time, stopping according to the stopping variable N defined by

(2.3) $$N = \text{smallest} \quad k \geqq n_0 \quad \text{such that} \quad (\hat{\sigma}^2(k)+k^{-1}) \leqq d^2\lambda_k/a_k^2$$

where λ_n is the smallest characteristic root of (X_nX_n'), $P\{\chi_p^2 < a^2\} = \alpha$, and $a_n^2 \to a^2$.

(II) When sampling is stopped at $N = n$, construct the region R_n defined by

$$R_n = \{\mathbf{z}: (\mathbf{z}-\hat{\boldsymbol{\beta}}(n))'(X_n X_n')(\mathbf{z}-\hat{\boldsymbol{\beta}}(n)) \leqq d^2\lambda_n\}. \tag{2.4}$$

It follows from Srivastava (1967) that the maximum diameter of the region R_n is $\leqq 2d$. The spheroid region R_n' can similarly be obtained as in Srivastava (1967).

3. Asymptotic properties of class $\mathscr{C}$. In this section, we study the properties of procedures in Class $\mathscr{C}$ as $d \to 0$.

THEOREM 1. *If condition* (C*) *is satisfied, then*

(a) $N(d) < \infty$ a.s.;

(b) $\lim_{d\to 0} N(d) = \infty$ a.s.;

(c) $\lim_{d\to 0} d^2\lambda_{N(d)}/\sigma^2 a^2_{N(d)} = 1$ a.s.;

(d) $\lim_{d\to 0} N(d)/q(a^2\sigma^2/d^2) = 1$ a.s.;

where $q(t) = \max\{n: \lambda_n < t\}$.

For proof, refer to Albert (1966) along with the following

LEMMA 1. *Under condition* C^*(i)

(a) $\lambda_n \equiv \lambda_{\min}(X_n X_n') \to \infty$,

(b) $\lambda_n/\lambda_{n+1} \to 1$,

as $n \to \infty$.

The proof of this lemma will be given at the end of this section.

THEOREM 2. *Under condition* (C^*) $\lim_{d\to 0} P\{\boldsymbol{\beta} \in R_n\} = \alpha$.

PROOF. We have

$$P\{\boldsymbol{\beta} \in R_N\} = P\{(\mathbf{y}_n' - \boldsymbol{\beta}' X_n) L_n' L_n (\mathbf{y}_n - X_n'\boldsymbol{\beta}) \leqq d^2\lambda_n\}$$
$$= P\{\mathbf{u}_n' L_n' L_n \mathbf{u}_n/\sigma^2 \leqq d^2\lambda_n/\sigma^2\}$$

where $u_n' = (u_1, \cdots, u_n)$, and $u_1, u_2, \cdots, u_n$ are independent and identically distributed with mean zero and variance σ^2. Let

$$\mathbf{w}_n = L_n \mathbf{u}_n, \qquad z_n = \sigma^{-2}\mathbf{w}_n'\mathbf{w}_n. \tag{3.1}$$

From Srivastava (1967) or Gleser (1965), z_n has an asymptotic chi-square $(\chi_p{}^2)$ distribution with p degrees of freedom. Thus, either by verifying Anscombe's condition C2 (uniform continuity in probability), as was done in the earlier version of this paper, or from Theorem 4.2 of Gleser (1969), (this theorem applies because

C^* (ii) $\rightarrow$ (4.3, Gleser (1969)), as is seen from the proof of Lemma 1(b)), $z_{N(d)}$ converges in law to a $\chi_n^{\,2}$ rv as $d \rightarrow 0$. Hence, from Theorem 1c

$$P\{\boldsymbol{\beta} \in R_N\} = P\{z_{N(d)} \leqq d^2 \lambda_{N(d)}/\sigma^2\}$$
$$\rightarrow P\{\chi_p^{\,2} \leqq a^2\} = \alpha \qquad \text{as } d \rightarrow 0.$$

This completes the proof of our Theorem 2. We now turn to the proof of Lemma 1.

PROOF OF LEMMA 1(a). Let

$$B_n = X_n X_n', \qquad \lambda_n = \lambda_{\min}(X_n X_n'). \tag{3.2}$$

Under the condition that $X_p = (\mathbf{x}^{(1)}, \cdots, x^{(p)})$ is of full rank (cf. Section 1), $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(p)}$ are linearly independent nonnull vectors of known constants and are *independent* of n. Hence

$$\operatorname{tr} \textstyle\sum_{i=1}^p \boldsymbol{l}_n^{(i)} \boldsymbol{l}_n^{(i)\prime} = \sum_{i=1}^p \boldsymbol{l}_n^{(i)\prime} \boldsymbol{l}_n^{(i)} = \operatorname{tr}[(\sum_{i=1}^p \mathbf{x}^{(i)} \mathbf{x}^{(i)\prime})(T_n^{-1\prime} T_n^{-1})]$$
$$= \operatorname{tr}(B_p B_n^{-1}) \geqq \lambda_{\min}(B_p) \operatorname{tr} B_n^{-1}.$$

Since $\sum_{i=1}^p \boldsymbol{l}_n^{(i)\prime} \boldsymbol{l}_n^{(i)} \rightarrow 0$ as $n \rightarrow \infty$, and since $\lambda_{\min}(B_p)$ is independent of n, we get $\operatorname{tr} B_n^{-1} \rightarrow 0$. Hence $\lambda_n \rightarrow \infty$.

PROOF OF LEMMA 1(b). Since

$$X_{n+1} X_{n+1}' = X_n X_n' + \mathbf{x}^{(n+1)} \mathbf{x}^{(n+1)\prime}, \tag{3.3}$$

we get

$$\lambda_{n+1} \geqq \lambda_n \quad \text{and} \quad \lambda_{n+1}/\lambda_n \geqq 1. \tag{3.4}$$

Taking the inverse of both sides of (3.3) and recalling that $B_n = (X_n X_n')$, $\lambda_{\max}(B) = ||B||^2$, we get

$$B_{n+1}^{-1} = B_n^{-\frac{1}{2}}[I - B_n^{-\frac{1}{2}} \mathbf{x}^{(n+1)} \mathbf{x}^{(n+1)\prime} B_n^{-\frac{1}{2}}] B_n^{-\frac{1}{2}}.$$

Hence

$$\begin{aligned}
\lambda_{n+1}^{-1} = \|B_{n+1}^{-1}\|^2 &\geqq \|B_n^{-1}\|^2[1 - \mathbf{x}^{(n+1)\prime} B_n^{-1} \mathbf{x}^{(n+1)}] \\
&= \lambda_n^{-1}[1 - \boldsymbol{l}_{n+1}^{(n+1)\prime} T_{n+1}' B_n^{-1} T_{n+1} \boldsymbol{l}_{n+1}^{(n+1)}] \\
&= \lambda_n^{-1}[1 - \boldsymbol{l}_{n+1}^{(n+1)\prime}(I - T_{n+1}^{-1} \mathbf{x}^{(n+1)} \mathbf{x}^{(n+1)\prime} T_{n+1}^{-1\prime})^{-1} \boldsymbol{l}_{n+1}^{(n+1)}] \\
&= \lambda_n^{-1}[1 - \boldsymbol{l}_{n+1}^{(n+1)\prime}(I - \boldsymbol{l}_{n+1}^{(n+1)} \boldsymbol{l}_{n+1}^{(n+1)\prime})^{-1} \boldsymbol{l}_{n+1}^{(n+1)}] \\
&\geqq \lambda_n^{-1}[1 - (\boldsymbol{l}_{n+1}^{(n+1)\prime} \boldsymbol{l}_{n+1}^{(n+1)} / (1 - \boldsymbol{l}_{n+1}^{(n+1)\prime} \boldsymbol{l}_{n+1}^{(n+1)}))].
\end{aligned}$$

Hence, from C^* (i), it follows that

$$\lim_{n \rightarrow \infty} \lambda_{n+1}/\lambda_n \leqq 1. \tag{3.5}$$

Combining (3.4) and (3.5) we get $\lambda_{n+1}/\lambda_n \rightarrow 1$ as $n \rightarrow \infty$.

4. Multivariate case. The set up for multivariate linear regression model is an easy and direct extension of univariate case. In place of the specification (1.1)–(1.2), we consider the more general set up as follows:

Let $Y_n = (\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)})$ be a $k \times n$ matrix of observations with independently distributed column vectors and

$$EY_n = \beta X_n \tag{4.1}$$

$$\mathrm{Cov} \cdot \mathbf{y}^{(i)} = \Sigma, \qquad i = 1, 2, \cdots, n$$

where β is a $k \times p$ matrix of unknown parameters and $X_n = (\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(n)})$, a $p \times n, p \leqq n-k$ matrix of known constants. We assume that X_p is of full rank (see Section 1). It is easy to see the correspondence between the univariate and multivariate. The specification (4.1) can easily be changed to the univariate case. Let $\mathbf{a}$ be any nonnull k-vector. Then (4.1) can be changed to the univariate case as follows:

$$E\mathbf{a}'Y_n = \mathbf{a}'\beta X_n \tag{4.2}$$

$$\mathrm{Var}\, \mathbf{a}'\mathbf{y}^{(i)} = \mathbf{a}'\Sigma\mathbf{a}.$$

Thus if we have a confidence region for β, we can then obtain one for $\mathbf{a}'\beta$ also. We require

(i) that the confidence region R_N for β be such that the maximum diameter of the region for $\mathbf{a}'\beta \leqq 2d$ for every nonnull k-vector $\mathbf{a}$, and

(ii) $\lim_{d\to 0} P(\beta \in R_N) \geqq \alpha$.

Under condition C^*, this can be achieved sequentially.

5. Solution. As in the univariate case, we find that the l.s. estimates for β and Σ are respectively,

$$\hat{\beta}'(n) = T_n^{-1'} L_n Y_n' = (T_n T_n')^{-1} X_n Y_n' = (X_n X_n')^{-1} X_n Y_n', \tag{5.1}$$

and

$$\hat{\Sigma}(n) = n^{-1} Y_n [I_n - L_n' L_n] Y_n'. \tag{5.2}$$

Let C_t be a random variable distributed according to the maximum characteristic root of a $t \times t$ Wishart matrix with mean lI_t, where $t = \min(k, p)$ and $l = \max(k, p)$. Let $\{a_n^*\}$ be any sequence of positive numbers converging to a^* satisfying

$$P\{c_t \leqq a^{*2}\} = \alpha. \tag{5.3}$$

Let

$$\mu_n = \lambda_{\max}\hat{\Sigma}((n)) \quad \text{and} \quad \mu = \lambda_{\max}(\Sigma). \tag{5.4}$$

Then $\mu_n \to \mu$ a.s. under C^* (i) (cf. Gleser's correction note). The steps (I) and (II) for this procedure are the same as in Section 2 with $\hat{\sigma}^2(n)$ and a_n^2 replaced by

μ_n and $a_n{}^{*2}$ respectively. The confidence regions R_n and R_n', $R_n \subset R_n'$, are respectively given by

$$R_n = \{Z: \lambda_{\max}[(Z-\hat{\beta}(n))'(X_nX_n')(Z-\hat{\beta}(n))] \leqq d^2\lambda_n\}$$

and

$$R_n' = \{Z: \lambda_{\max}[(Z-\hat{\beta}(n))'(Z-\hat{\beta}(n))] \leqq d^2\}.$$

It is simple to show that properties of Theorem 1 hold here also with σ^2 and a^2 replaced by μ and a^{*2} respectively. The results similar to Theorem 2 are contained in the following

THEOREM 3.

$$\lim_{d\to 0} P\{\beta \in R_n'\} \geqq \lim_{d\to 0} P\{\beta \in R_n \geqq \alpha.$$

PROOF. Following Section 3, we have

$$P\{\beta \in R_n\} = P\{\max_{\mathbf{a}:\mathbf{a}'\mathbf{a}=1} [\mathbf{a}'(\hat{\beta}(n)-\beta)(X_nX_n')(\hat{\beta}(n)-\beta)\mathbf{a}] \leqq d^2\lambda_n\}$$
$$= P\{\max_{\mathbf{a}:\mathbf{a}'\mathbf{a}=1} \mathbf{a}'U_nL_n'L_nU_n'\mathbf{a} \leqq d^2\lambda_n\},$$

where $U_n = (\mathbf{u}^{(1)}, \cdots, \mathbf{u}^{(n)})$ is a $k \times n$ matrix with independent column vectors with $E(\mathbf{u}^{(i)}) = 0$ and $\text{Cov}\cdot\mathbf{u}^{(i)} = \Sigma$, $i = 1, 2, \cdots, n$. Using the result (see [7] page 142) that $\sup_{a:a'a=1} [\mathbf{a}'A\mathbf{a}/\mathbf{a}'B\mathbf{a}] = \lambda_{\max} AB^{-1}$, B pd, and following as in [8], we get

$$P\{\beta \in R_n\} = P\{\lambda_{\max}\Sigma^{-1}U_nL_n'L_nU_n' \leqq d^2\lambda_n/\mu\} = P\{z_n \leqq d^2\lambda_n/\mu\}$$

where

$$z_n = \lambda_{\max}(\Sigma^{-1}U_nL_n'L_nU_n') = \lambda_{\max}(W_nW_n'); \quad W_n = \Sigma^{-\frac{1}{2}}U_nL_n.$$

Without any loss of generality, we assume that $k \leqq p$, i.e., $t = \min(k, p) = k$ and $l = \max(k, p) = p$. Hence it follows as in Section 3 that W_nW_n' has asymptotically Wishart distribution with mean pI_k. Thus, either by verifying Anscombe's condition C2, as was done in the earlier version of this paper, or from Theorem 4.4 of Gleser (1969), it follows that $W_{N(d)}W_{N(d)}'$ has asymptotically Wishart distribution with mean pI_k.

Acknowledgment. The author wishes to thank the referees for their helpful suggestions and for catching some errors.

REFERENCES

[1] ALBERT, A. (1966). Fixed size confidence ellipsoids for linear regression parameters. *Ann. Math. Statist.* **37** 1602–1630.

[2] ANSCOMBE, F. J. (1952). Large sample theory of sequential estimation. *Proc. Cambridge Philos. Soc.* **48** 600–608.

[3] CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[4] GLESER, L. J. (1965). On the asymptotic theory of fixed-size sequential confidence bounds for linear regression parameters. *Ann. Math. Statist.* **36** 463–467. (See Correction Note, *Ann. Math. Statist.* **37** 1053–55 (1966).)

[5] GLESER, L. J. (1969). On limiting distributions for sums of a random number of independent random vectors. *Ann. Math. Statist.* **40** 935–941.

[6] LOÈVE, M. (1963). *Probability Theory*, 3rd ed. Van Nostrand, New York.

[7] ROY, S. N. (1957). *Some Aspects of Multivariate Analysis.* Wiley, New York.

[8] SRIVASTAVA, M. S. (1967). On fixed-width confidence bounds for regression parameters and mean vector. *J. Roy. Statist. Soc. Ser. B* **29** 132–140.