The Annals of Statistics
1990, Vol. 18, No. 3, 1389–1399

SEQUENTIAL CONFIDENCE REGIONS IN INVERSE REGRESSION PROBLEMS

By Jiunn T. Hwang[1] and Hung-kung Liu

Cornell University and SUNY at Stony Brook

In inverse regression problems (or more generally, the estimation of ratios of regression parameters) and errors-in-variables models, it has been shown by Gleser and Hwang that the length of any confidence interval with positive confidence level is infinite with positive probability. Therefore the confidence sets derived using asymptotic theory, although having correct asymptotic coverage probability, typically have zero confidence level when the sample size is fixed.

Is it possible to construct a sequential confidence interval with finite length and $1 - \alpha > 0$ confidence level? The answer is no for any *finite* stage sequential sampling. The answer is, however, yes for a *fully* sequential scheme, as demonstrated by Hwang and Liu. For the inverse regression problem, and more generally the set estimation of a ratio of regression parameters, we construct a $(1 - \alpha)$ confidence sequence. Applying such a confidence sequence, we can construct a $(1 - \alpha)$ sequential confidence interval with the length less than a prespecified quantity.

1. Introduction. In this paper, we deal with the problem of constructing sequential confidence intervals for ratios of regression parameters. Solutions to such problems can be applied to inverse regression problems (calibration problems) and biological assay problems. See Malley (1982) for extensive discussions about applications in a nonsequential setting. Often in biological assay, one is interested in comparing the potency of a new treatment to a standard treatment in terms of a ratio of parameters. Usually, the experiment is done in a sequential manner so that the experiment can be terminated as soon as enough evidence has been accumulated to ensure that the better treatment can be applied to patients early.

Gleser and Hwang (1987) considered the same models and other related models such as errors-in-variables models, except that they took a finite-sample (nonsequential) approach. They proved that any confidence sets having a positive confidence level $1 - \alpha$ will suffer the inevitable property that there is a positive probability of having infinite length. (A *confidence level* of a confidence set is the minimum coverage probability over the parameter space). Consequently, many confidence sets with asymptotic coverage probability $1 - \alpha$ have in fact zero confidence level (with respect to the finite-sample probability), since they usually have finite length for almost every observation.

In this paper, we address the question as to whether the sequential approach could be used to construct a nontrivial sequential confidence set, i.e., a

Received January 1989; revised August 1989.

[1]Research supported by NSF grant DMS 88-09016.

AMS 1980 *subject classifications*. Primary 62F25, 62L10; secondary 62H99, 62F11.

Key words and phrases. Confidence level, confidence region, stopping rule, calibration, principal components analysis, confidence sequences.

finite-diameter confidence set with positive confidence level based on a sampling that stops with probability 1.

As it turns out that it is still impossible to construct a nontrivial confidence set using a *two-stage* sequential procedure or, in fact, using any *finite-stage* sequential procedure. The results hold for a wide class of models, having the same generality as the main theorem in Gleser and Hwang (1987). These models include inverse regression models, multivariate linear and most nonlinear errors-in-variables models and estimation of principal component vectors.

What if one uses a *fully* sequential procedure? The answer is then yes for at least two models: structural multiple linear errors-in-variables and the estimation of ratios of multiple regression parameters. This existence result and the nonexistence result regarding finite-stage sequential rules (depicted in the last paragraph) are not reported here due to the limitation on space. See Hwang and Liu (1989). In the present paper, we construct, for the ratio problem, nontrivial confidence intervals based on a result of Sinha and Sarkar (1984).

Before discussing the significance of the construction, we mention some earlier results. A recent reference for the confidence set construction in the inverse multiple regression problem is Oman (1988). As implied by Gleser and Hwang (1987), these proposed intervals inevitably share the undesirable property that their diameters may be infinite. Earlier, realizing the problem, Perng and Tong (1974) considered a simple inverse regression model and used two fully sequential samples to construct a confidence interval of a prespecified length l. As $l \to 0$, the coverage probability was shown to converge to $1 - \alpha > 0$. Much later Levy and Samaranayake (1988) provided a multivariate generalization of the results of Perng and Tong (1974).

We consider, in the present paper, a standard multiple linear regression model

$$\underset{n\times 1}{y_n} = \underset{n\times p}{X_n} \underset{p\times 1}{\beta} + \underset{n\times 1}{\varepsilon}, \tag{1.1}$$

where $y_n = (y_1, \dots, y_n)'$, $\beta = (\beta_1, \dots, \beta_p)'$ and $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$, where ε_i are iid with mean 0 and variance σ_ε^2. Let $p_1 \le p$ and $\beta_{(1)} = (\beta_1, \dots, \beta_{p1})'$. Under the assumption that $b'\beta_{(1)} \neq 0$, we construct a confidence sequence for $a'\beta_{(1)}/b'\beta_{(1)}$, where a and b are p_1-dimensional vector. [We could have taken $\beta_{(1)} = \beta$, but considering $\beta_{(1)}$ gives us more flexibility]. We achieve three goals that were not previously achieved.

(I) We construct fully sequential confidence procedures of finite length or length less than a prespecified quantity with coverage probability at least $1 - \alpha > 0$. (Yes, this is true for any parameter configuration and this is an exact and nonasymptotic result).

(II) We actually discover a confidence sequence. A sequence of confidence intervals C_n for θ is called a $(1 - \alpha)$ confidence sequence if for some m,

$$P(\theta \in C_n, \forall\, n \ge m) \ge 1 - \alpha. \tag{1.2}$$

Here m is the prespecified minimum number of samples taken. Note that (1.2) implies that for any stopping rule N such that $P(m < N < \infty) = 1$, C_N will always have at least $1 - \alpha$ coverage probability. For discussion of confidence sequences, see Robbins (1970), Lai (1976), Kahn (1978) and Sinha and Sarkar (1984). See also Farrell (1962, 1964).

A confidence sequence is therefore very flexible, since we can implement according to any stopping rule, or even any unspecified or partially unspecified stopping rule. Although at the expense of the procedure's efficiency, the flexibility is important especially in clinical trials, since in such studies, there are many factors such as side effects and financial cost that affect the decision to terminate a study. For detailed discussions about other situations where it is difficult to follow a specific rule, see, for example, Jennison and Turnbull (1984, 1989).

(III) The confidence sequences in Section 2 cover not only one ratio of parameters but simultaneously ratios of any linear combinations. Therefore these intervals are valid even if they are used repeatedly for different a and b. This is a sequential version of Scheffé's (1959) simultaneous confidence intervals for our problem. A sharper confidence sequence for a given pair of a and b is given in Section 3.

2. Scheffé type simultaneous confidence sequences. In this section, for model (1.1), we construct a confidence sequence simultaneously for $a'\beta_{(1)}/b'\beta_{(1)}$, where a and b are any p_1-dimensional vectors with $b'\beta_{(1)} \neq 0$. We assume that the matrix X_n has a full rank p, and that ε_i is normally distributed with mean 0 and variance σ^2. Let $\hat{\beta}_n$ be the least squares estimator of β, i.e., $\hat{\beta}_n = (X_n'X_n)^{-1}X_n'y_n$. Also let $\hat{\sigma}_n^2$ be the unbiased estimator for σ^2. Hence $\hat{\sigma}_n^2 = |y_n - X_n\hat{\beta}_n|^2/(n-p)$. Let $\hat{\beta}_{(1)n}$ be the intuitive estimator of $\beta_{(1)}$, namely, $\hat{\beta}_n' = (\hat{\beta}_{(1)n}', \hat{\beta}_{(2)n}')$ and $\hat{\beta}_{(1)n}$ is the vector of the first p_1 components of $\hat{\beta}_n$. Write the covariance matrix of $\hat{\beta}_{(1)n}$ as $\sigma^2\Sigma_{11,n}$, which is the $p_1 \times p_1$ principle submatrix of the covariance matrix $\sigma^2(X_n'X_n)^{-1}$ of $\hat{\beta}_n$. Sinha and Sarkar (1984) exhibited a confidence sequence S_n for $\beta_{(1)}$ with

$$P_{\beta,\sigma}\{\beta_{(1)} \in S_n, \forall\ n \geq m\} \geq 1 - \alpha, \tag{2.1}$$

where $m \geq p + 1$ is a prechosen lower bound on the sample size n. Actually, the procedure S_n depends also on m, which is suppressed in the notation. Their confidence region is, $\forall\ n \geq m$,

$$S_n = \left\{\beta_{(1)}: \left(\hat{\beta}_{(1)n} - \beta_{(1)}\right)'\Sigma_{11,n}^{-1}\left(\hat{\beta}_{(1)n} - \beta_{(1)}\right) \leq c_{n,\alpha}\hat{\sigma}_n^2\right\}, \tag{2.2}$$

where

$$c_{n,\alpha} = (n-p)\left(\left(1 + \frac{p_1 a}{m(m-p)}\right)^{m/n}\left(\frac{|\Sigma_{11,m}|}{|\Sigma_{11,n}|}\right)^{1/n} - 1\right), \tag{2.3}$$

a is the unique solution of

$$\alpha = P\left(F_{p_1, m-p} \geq a\right) + \left(1 + \frac{p_1 a}{m-p}\right)^{-m/2} \times \int_0^a \left(1 + \frac{p_1 x}{m-p}\right)^{m/2} f_{p_1, m-p}(x)\, dx, \tag{2.4}$$

and $F_{p_1, m-p}$ has an F-distribution with p_1 and $m-p$ degrees of freedom and $f_{p_1, m-p}$ is its density function.

Based on S_n, we can define an interval I_n,

$$I_n = \left\{ \frac{a'\beta_{(1)}}{b'\beta_{(1)}} : \beta_{(1)} \in S_n \right\}. \tag{2.5}$$

COROLLARY 2.1.

$$P_{\beta,\sigma}\left\{ \frac{a'\beta_{(1)}}{b'\beta_{(1)}} \in I_n \ \textit{for all } a \textit{ and } b \textit{ such that } b'\beta_{(1)} \neq 0 \textit{ and } \forall\ n \geq m \right\} \geq 1 - \alpha. \tag{2.6}$$

PROOF. Obviously $\beta_{(1)} \in S_n$ implies $a'\beta_{(1)}/b'\beta_{(1)} \in I_n$ for all a, b as long as $b'\beta_{(1)} \neq 0$. □

It is also possible to obtain an explicit expression for I_n as in Theorem 2.2. In doing so, we consider only the case where $b'\beta_{(1)} \neq 0$ for every $\beta_{(1)} \in S_n$. (Here, a fully sequential stopping rule can be taken so that this will happen.) Note that by Scheffé's (1959) results (a statement which is repeated in Lemma A.3), this assumption is equivalent to

$$\text{zero is separated from } b'\hat{\beta}_{(1)n} \pm \hat{\sigma}_n\left(c_{n,\alpha} b'\Sigma_{11,n} b\right)^{1/2}. \tag{2.7}$$

Here, for any real numbers k_1 and $k_2 > 0$, $k_1 \pm k_2$ represents the interval $(k_1 - k_2, k_1 + k_2)$. Since b is a nonzero vector, we may assume without loss of generality that $b_1 \neq 0$. Let $a = (a_1, \dots, a_{p1})$, $b = (b_1, \dots, b_{p1})$, $a_{(2)} = (a_2, \dots, a_{p1})$ and $b_{(2)} = (b_2, \dots, b_{p1})$. We also define

$$A = A_n = \left(\hat{\beta}'_{(1)n}\Sigma_{11,n}^{-1}\hat{\beta}_{(1)n} - c_{n,\alpha}\hat{\sigma}_n^2\right)\Sigma_{11,n}^{-1} - \Sigma_{11,n}^{-1}\hat{\beta}_{(1)n}\hat{\beta}'_{(1)n}\Sigma_{11,n}^{-1}. \tag{2.8}$$

Partition A as

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where A_{11} is a scalar. Let

$$\begin{aligned} B &= \left(A_{11}/b_1^2\right) b_{(2)} b'_{(2)} - (2/b_1) b_{(2)} A_{12} + A_{22}, \\ v &= \left(A_{11}/b_1\right) b_{(2)} - (1/b_1) A_{21}, \\ d &= v'B^{-1}v - A_{11}/b_1^2 \quad \text{and} \quad a_* = a_{(2)} - (a_1/b_1) b_{(2)}. \end{aligned} \tag{2.9}$$

THEOREM 2.2. *Assume that* $b'\beta_{(1)} \neq 0$ *for every* $\beta_{(1)} \in S_n$. *Namely* (2.3) *holds. Then* $d > 0$ *and*

$$I_n = \left((a_1/b_1) + a'_* B^{-1} v\right) \pm \left(d a'_* B^{-1} a_*\right)^{1/2}. \tag{2.10}$$

Theorem 2.2 is proved in the Appendix.

As established in Corollary 2.1, I_n, $n \geq m$, is a confidence sequence. Therefore I_N, for any stopping rule N, $m \leq N \leq \infty$, will have guaranteed coverage probability. This is due to the obvious inequality that $P(a'\beta_{(1)}/b'\beta_{(1)} \in I_N)$ is greater or equal to the left-hand side of (2.6). In particular, we can use the following two stopping times:

N_1 = the first $n \geq m$ such that I_n has finite length, i.e., (2.7) holds, and

N_2 = the first $n \geq m$ such that I_n has length no greater than a prespecified length l, i.e., (2.7) holds and $2\,(d a'_* B^{-1} a_*)^{1/2} \leq l$.

It can be argued that N_1 and N_2 will stop with probability 1 for $b\beta_{(1)} \neq 0$ and under the assumptions on X_n that implies the consistency of $\hat{\beta}_{(1)n}$. As an example, a typical assumption is that $(X_n' X_n)^{-1} = O(n^{-1})$, i.e., each and every element of $(X_n' X_n)/n$ approaches a finite number as $n \to \infty$. In fact, Theorem 2.3 holds if

$$\left(X_n' X_n\right)^{-1} = O(n^{-\varepsilon}) \quad \text{for some } \varepsilon > 0. \tag{2.11}$$

THEOREM 2.3. *Assume that* $b'\beta_{(1)} \neq 0$. *Then* $P(N_1 < \infty) = P(N_2 < \infty) = 1$.

PROOF. Since $N_2 \geq N_1$, it suffices to prove that N_2 is finite almost surely. Obviously,

$$P(N_2 < \infty) \geq P(\mathrm{dia}(I_n) \leq l), \quad \forall\, n.$$

The lower bound approaches 1 as $n \to \infty$, since $\mathrm{dia}(I_n) \to 0$ a.s. due to the fact that the volume of S_n approaches 0, an assertion proved below. Note that the volume of S_n is proportional to

$$\begin{aligned}
&\hat{\sigma}_n^{p_1} |\Sigma_{11,n}|^{1/2} (c_{n,\alpha})^{p_1/2} \\
&\quad = \hat{\sigma}_n^{p_1} \left\{ n \left[\left(1 + \frac{p_1 a}{m-p}\right)^{m/n} |\Sigma_{11,m}|^{1/n} |\Sigma_{11,n}|^{1/p_1 - 1/n} - |\Sigma_{11,n}|^{1/p_1} \right] \right\}^{p_1/2} \\
&\quad = \hat{\sigma}_n^{p_1} \{ n (k^{1/n} n^{-p_1 \varepsilon (1/p_1 - 1/n)} - n^{-\varepsilon}) \}^{p_1/2} O(1),
\end{aligned}$$

where $k = (1 + p_1 a/(m-p))^m |\Sigma_{11,m}|$. The last displayed expression can be shown easily to approach 0 for $\varepsilon \geq 1$. When $\varepsilon < 1$, the same assertion can be established by using L'Hospital's rule. This completes the proof. □

3. A sharper confidence sequence. In the last section, we exhibit the Scheffé-type sequential confidence sequence for ratios of parameters. Even though it has the flexibility of dealing with infinitely many ratios simultaneously, it may be unnecessarily wide if one is only interested in a particular ratio. In this section we construct a sharper confidence sequence solely for $a'\beta/b'\beta$, where a and b are fixed vectors which are linearly independent. (The problem is trivial if a and b are linearly dependent.)

We make a transformation $\eta = P\beta$, where P is a nonsingular matrix with the first two rows being a' and b'. Therefore the first two components η_1 and η_2 of η are $a'\beta$ and $b'\beta$, respectively. Write $y_n = (X_n P^{-1})\eta + \varepsilon$. The problem is then to set a confidence limit for $\theta = a'\beta/b'\beta = \eta_1/\eta_2$.

Consider a fixed θ_0 and focus on, for now, the set estimation problem for $\eta_1 - \theta_0\eta_2$. Based on y_n, its least squares estimator is $\hat{\eta}_{1(n)} - \theta_0\hat{\eta}_{2(n)}$, where $\hat{\eta}_{1(n)}$ and $\hat{\eta}_{2(n)}$ are the least squares estimators for η_1 and η_2, respectively. From (2.2), a confidence sequence for $\eta_1 - \theta_0\eta_2$ is described by the inequality

$$\frac{\left(\hat{\eta}_{1(n)} - \theta_o\hat{\eta}_{2(n)} - (\eta_1 - \theta_0\eta_2)\right)^2}{(1, -\theta_0)\Sigma_n(1, -\theta_0)'} \le c_{n,\alpha}\hat{\sigma}_n^2, \tag{3.1}$$

where $\sigma^2\Sigma_n$ is the covariance matrix of $(\hat{\eta}_{1(n)}, \hat{\eta}_{2(n)})$. Hence $\Sigma_n = P_{(1)}(X_n'X_n)^{-1}P_{(1)}'$, where $P_{(1)} = (a, b)'$. Furthermore, α and $c_{n,\alpha}$ are defined in (2.3) and (2.4) with $p_1 = 1$ and $|\Sigma_{11,j}|$, $j = m$, or n, should now be replaced by $(1, -\theta_0)\Sigma_j(1, -\theta_0)'$. It then follows that for every η_1, η_2 and θ_0 with probability at least $1 - \alpha$ that (3.1) holds for $n \ge m$. Setting $\theta_0 = \eta_1/\eta_2 = \theta$, (3.1) now reduces to

$$\frac{\left(\hat{\eta}_{1(n)} - \theta\hat{\eta}_{2(n)}\right)^2}{\hat{\sigma}_n^2(1, -\theta)\Sigma_n(1, -\theta)'} \le c_{n,\alpha}, \tag{3.2}$$

where

$$c_{n,\alpha} = (n-p)\left[\left(1 + \frac{a}{m-p}\right)^{m/n}\left[\frac{(1, -\theta)\Sigma_m(1, -\theta)'}{(1, -\theta)\Sigma_n(1, -\theta)'}\right]^{1/n} - 1\right],$$

and a is such that (2.4) holds with $p_1 = 1$. The set of θ satisfying (3.2) is denoted by SQ_n. Here, Q stands for the quadratic form on the left-hand side of (3.2). This quadratic expression is the Fieller's pivot (1954). See also Zerbe (1978).

COROLLARY 3.1. *SQ_n is a confidence sequence of level $1 - \alpha$. That is $P(\theta \in SQ_n, \forall\ n \ge m) \ge 1 - \alpha$.*

The right-hand side of (3.2), in general, depends on θ. To get an interval of θ, one has to solve a high degree polynomial, which can only be accomplished by numerical computations. There are, however, situations in which a simpler solution is possible.

For example, if $\Sigma_n = k\Sigma_m$, where the constant k depends on m and n, then $c_{n,\alpha}$ will not depend on θ. This is the case for a two-sample problem: $X = (X_1, \dots, X_n)'$ and $Y = (Y_1, \dots, Y_n)'$ and $EX_i = \eta_1$ and $EY_i = \eta_2$. For such a situation, SQ_n is exactly a Fieller's set [except, of course, the cut-off point $C_{n,\alpha}$ in (3.2) is different from Fieller's cut-off point] and note again that the confidence sequence is exactly the one-dimensional t-confidence sequence of Sinha and Sarkar (1984). This is probably the best scalar invariant one-dimensional t-interval that can be derived using Robbins' (1970) techniques.

For a general situation, a slightly more conservative confidence set which has a simpler form can be derived. Let Maxeig(M) denote the maximum of the eigenvalues of an arbitrary matrix M. One notes that

$$c^*_{n,\alpha} \equiv \max_\theta c_{n,\alpha} = (n-p)\left\{\left(1 + \frac{a}{m-p}\right)^{m/n}\left[\text{Maxeig}\left(\Sigma_m \Sigma_n^{-1}\right)\right]^{1/n} - 1\right\}.$$

Replacing $c_{n,\alpha}$ by $c^*_{n,\alpha}$ in (3.2), we obtain a larger set SQ^*_n which is also a $(1-\alpha)$ confidence sequence due to Corollary 3.1.

COROLLARY 3.2. *SQ^*_n is a confidence sequence of level $1-\alpha$. That is $P(\theta \in SQ^*_n, \forall\ n \geq m) > 1 - \alpha$.*

Since $c^*_{n,\alpha}$ does not depend on θ, to get an interval of θ, one only needs to solve a quadratic inequality, an analytically easy task. This is exactly what was done in the nonsequential classical example of Fieller (1954). SQ^*_n is probably not much larger than SQ_n, especially when n is large. They are identical for the two-sample problem depicted in the last paragraph.

Based on SQ^*_n (and SQ_n), we can define a sequential confidence interval with a prespecified length l. Naturally, we take

N_1 = the first $n \geq m$ such that the length of SQ_n is less than l, or

N_2 = the first $n \geq m$ such that the length of SQ^*_n is less than l.

Under assumption (2.11), we can establish the following theorem.

THEOREM 3.3. *Assume that $b'\beta \neq 0$, then*

$$P(N_1 < \infty) = P(N_2 < \infty) = 1; \tag{3.3}$$

and

$$EN_1 \text{ and } EN_2 \text{ are finite}. \tag{3.4}$$

Furthermore $P(\theta \in SQ_{N_1})$ and $P(\theta \in SQ^*_{N_2})$ are bounded below by $1-\alpha$. Hence SQ_{N_1} and $SQ^*_{N_2}$ are nontrivial sequential confidence intervals with length no greater than l.

PROOF. We will prove

$$P(N_2 < \infty) = 1, \tag{3.5}$$

which implies that $P(N_1 < \infty) = 1$ since $N_1 \le N_2$. The fact that SQ_{N_1} and $SQ^*_{N_2}$ have coverage probability bounded below by $1 - \alpha$ follows directly from (3.3) and Corollaries 3.1 and 3.2. We omit the proof of (3.4), which is quite lengthy and is based on the fact that $\hat{\sigma}_n^2$ has an exponential tail.

To prove (3.5), we note that the left-hand side is bounded below by

$$P(\text{length of } SQ_n^* \le l) \tag{3.6}$$

for every n. It therefore suffices to prove that (3.6) approaches 1 as $n \to \infty$. Let ν_{ij}, $1 \le i \le 2$, $1 \le j \le 2$, denote the (i, j)th element of Σ_n. Hence $\theta \in SQ_n^*$ if and only if

$$\hat{\eta}_1^2 - 2\theta\hat{\eta}_1\hat{\eta}_2 + \theta^2\hat{\eta}_2^2 \le c^*_{n,\alpha}\hat{\sigma}_n^2(\nu_{11} - 2\theta\nu_{12} + \theta^2\nu_{22}),$$

or equivalently

$$\theta^2(\hat{\eta}_2^2 - c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{22}) + 2\theta(c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{12} - \hat{\eta}_1\hat{\eta}_2) + \hat{\eta}_1^2 - c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{11} \le 0.$$

The following argument relies heavily on the assertion that

$$c^*_{n,\alpha}\Sigma_n \to 0. \tag{3.7}$$

This assertion can be proved using (2.11) and the identity

$$[\text{Maxeig}(\Sigma_m\Sigma_n^{-1})]^{1/n} = [\text{Mineig}(\Sigma_n\Sigma_m^{-1})]^{-1/n},$$

where Mineig($\cdot$) represents the minimum eigenvalue. The detailed calculation is similar to the proof at the end of Theorem 2.3.

As an application of (3.7), we have as $n \to \infty$,

$$\hat{\eta}_2^2 - c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{22} \to \eta_2^2 = (b\beta')^2 > 0.$$

When $\hat{\eta}_2^2 - c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{22} > 0$, the solution of θ form an interval with half-length

$$\frac{\left[(c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{12} - \hat{\eta}_1\hat{\eta}_2)^2 - (\hat{\eta}_2^2 - c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{22})(\hat{\eta}_1^2 - c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{11})\right]_+^{1/2}}{\hat{\eta}_2^2 - c^*_{n,\alpha}\hat{\sigma}_n^2\nu_{22}}, \tag{3.8}$$

where, for any number a, a_+ denotes $\max(a, 0)$. By (3.7), the half-length in (3.8) approaches 0 and hence (3.6) approaches 1, completing the proof. □

One can use the idea to construct a confidence sequence simultaneously for a finite number of ratios.

APPENDIX

PROOF OF THEOREM 2.2.

The proof is based on the following lemmas. Let C_n be the cone containing S_n; i.e.,

$$C_n = \bigcup_{\substack{t \in R \\ t \ne 0}} \{\beta_{(1)}/t \colon \beta_{(1)} \in S_n\}.$$

LEMMA A.1. *$C_n = \{x: x'Ax \le 0,\ x \ne 0\}$, where A is given in* (2.8).

PROOF. Obviously,

$$C_n = \{x:\ f(t,x) \le c_{n,\alpha}\hat{\sigma}_n^2 \text{ for some } t \in R\},$$

where $f(t,x) = (\hat{\beta}_{(1)n} - tx)'\Sigma_{11,n}^{-1}(\hat{\beta}_{(1)n} - tx)$. Due to the fact that for each fixed x, the infimum is attained for some t,

$$C_n = \left\{x:\ \inf_{-\infty<t<\infty} f(t,x) \le c_{n,\alpha}\hat{\sigma}_n^2\right\}.$$

Now $f(t,x)$ is minimized at $t = (x'\Sigma_{11,n}^{-1}\hat{\beta}_{(1)n})/(x'\Sigma_{11,n}^{-1}x)$. Hence

$$\inf_{-\infty<t<\infty} f(t,x) = -\frac{\left(x'\Sigma_{11,n}^{-1}\hat{\beta}_{(1)n}\right)^2}{x'\Sigma_{11,n}^{-1}x} + \hat{\beta}'_{(1)n}\Sigma_{11,n}^{-1}\hat{\beta}_{(1)n}.$$

Consequently,

$$C_n = \left\{x:\ x'\Sigma_{11,n}^{-1}\hat{\beta}_{(1)n}\hat{\beta}'_{(1)n}\Sigma_{11,n}^{-1}x \ge \left(\hat{\beta}'_{(1)n}\Sigma_{11,n}^{-1}\hat{\beta}_{(1)n} - c_{n,\alpha}\hat{\sigma}_n^2\right)x'\Sigma_{11,n}^{-1}x\right\},$$

which implies the lemma. □

Before attacking I_n, we consider a related set, an ellipsoid, $E = \{\beta_{(1)}/(b\beta_{(1)}):\ \beta_{(1)} \in S_n\}$. Note that $I_n = \{a'x:\ x \in E\}$. Using C_n, we can have a representation of E. This is also due to the statement that

$$x \in E \Leftrightarrow x \in C_n \text{ and } b'x = 1, \tag{A.1}$$

which can easily be proved by simple analytic arguments.

LEMMA A.2.

$$E = \left\{x:\ x_1 = \frac{1 - b'_{(2)}x_{(2)}}{b_1} \ \textit{and}\ \left(x_{(2)} - B^{-1}v\right)'B\left(x_{(2)} - B^{-1}v\right) \le d\right\},$$

where B, v and d are given in (2.9).

PROOF. From (A.1), we will consider the intersection of C_n and the plane determined by $b'x = 1$. The last equation is equivalent to $x_1 = (1 - b'_{(2)}x_{(2)})/b_1$. Using Lemma A.1 and substituting this for x in the inequality $x'Ax \le 0$, we obtain

$$\left(\frac{1 - b'_{(2)}x_{(2)}}{b_1},\ x'_{(2)}\right)\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix} \dfrac{1 - b'_{(2)}x_{(2)}}{b_1} \\ x_{(2)}\end{pmatrix} \le 0,$$

which, by completing the square, is equivalent to

$$\left(x_{(2)} - B^{-1}v\right)'B\left(x_{(2)} - B^{-1}v\right) \le d. \tag{A.2}$$

Here d is positive; otherwise E is an empty set or a point set which implies the same for S_n. But S_n is neither. □

Now we can complete the proof of Theorem 2.2 by resorting to Scheffé's (1959) lemma.

LEMMA A.3. *Let S be a set consisting of x such that $(x - v_0)'D(x - v_0) \leq c$, where v_0 is a fixed vector, D is a fixed positive definite matrix and c is a positive constant. Then for a fixed vector l, the set of lx, $x \in S$, is $l'v_0 \pm (cl'D^{-1}l)^{1/2}$.*

PROOF OF THEOREM 2.2. Now the interval I_n consists of $a'x$, where $x \in E$. Note that

$$a'x = a_1 \frac{1 - b'_{(2)}x_{(2)}}{b_1} + a'_{(2)}x_{(2)} = \frac{a_1}{b_1} + a'_* x_{(2)},$$

where a_* *is* given in (2.9). Since $x_{(2)}$ satisfies (A.2), by Lemma A.3, the set of $a'_* x_{(2)}$ is

$$a'_* B^{-1} v \pm \left(d a'_* B^{-1} a_*\right)^{1/2}.$$

Equation (2.10) then follows. □

Acknowledgments. The authors thank professors T. L. Lai and David Siegmund for their encouragement which led to this version, improving upon an earlier one. They also thank professors L. D. Brown, Roger Farrell, R. Kahn, B. K. Sinha, B. Turnbull and C. Zhang for helping the authors in understanding the relevant sequential analysis results. We appreciate professor William Strawderman's many useful editorial comments.

REFERENCES

FARRELL, R. H. (1962).. Bounded length confidence intervals for the zero of a regression function. *Ann. Statist.* **33** 237–247.

FARRELL, R. H. (1964). Asymptotic behavior of expected sample size in certain one sided tests. *Ann. Statist.* **35** 36–72.

FIELLER, E. C. (1954). Some problems in interval estimation. *J. Roy. Statist. Soc. Ser. B* **16** 175–185.

GLESER, L. J. and HWANG, J. T. (1987). The nonexistence of $100(1 - \alpha)\%$ confidence sets of finite expected diameter in errors-in-variables and related models. *Ann. Statist.* **15** 1351–1362.

HWANG, J. T. and LIU, H. K. (1989). Existence and nonexistence theorems of finite diameter sequential confidence regions for errors-in-variables models. Technical Report, Cornell Statist. Center.

JENNISON, C. and TURNBULL, B. W. (1984). Repeated confidence intervals for group sequential clinical trials. *Controlled Clin. Trials.* **5** 33–45.

JENNISON, C. and TURNBULL, B. W. (1989). Interim analysis: Repeated confidence interval approach (with Discussion). *J. Roy. Statist. Soc. Ser. B* **51** 305–361.

KAHN, R. (1978). On the confidence sequences for the mean vector of a multivariate normal distribution. *J. Multivariate Anal.* **8** 550–558.

LAI, T. L. (1976). On confidence sequences. *Ann. Statist.* **4** 265–280.

LEVY, M. and SAMARANAYAKE, V. A. (1988). Fixed width interval estimation for the multiple response calibration problem. Technical Report, Dept. Quantitative Anal. and Inform. Systems, Univ. Cincinnati.

MALLEY, J. (1982). Simultaneous confidence intervals for ratios of normal means. *J. Amer. Statist. Assoc.* **77** 170–176.

OMAN, S. D. (1988). Confidence region in multivariate calibration. *Ann. Statist.* **16** 174–187.

PERNG, S. K. and TONG, Y. L. (1974). A sequential solution to the inverse linear regression problem. *Ann. Statist.* **2** 535–539.

ROBBINS, H. (1970). Statistical methods related to the law of the iterated logarithm. *Ann. Math. Statist.* **41** 1397–1409.

SCHEFFÉ, H. (1959). *The Analysis of Variance*. Wiley, New York.

SINHA, B. K. and SARKAR, S. K. (1984). Invariant confidence sequences for some parameters in a multivariate linear regression model. *Ann. Statist.* **12** 301–310.

ZERBE, G. O. (1978). On Fieller's theorem and the general linear model. *Amer. Statist.* **32** 103–105.

DEPARTMENT OF MATHEMATICS
WHITE HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

DEPARTMENT OF APPLIED MATHEMATICS
AND STATISTICS
STATE UNIVERSITY OF NEW YORK
STONY BROOK, NEW YORK 11794-3600