The Annals of Statistics
1984, Vol. 12, No. 1, 301–310

INVARIANT CONFIDENCE SEQUENCES FOR SOME PARAMETERS IN A MULTIVARIATE LINEAR REGRESSION MODEL

BY B. K. SINHA AND S. K. SARKAR[1]

University of Pittsburgh and Temple University

Let $\mathbf{X}_1, \mathbf{X}_2, \cdots$ be independent p-variate normal vectors with $E\,\mathbf{X}_\alpha \equiv \beta\,\mathbf{Y}_\alpha$, $\alpha = 1, 2, \cdots$ and same p.d. dispersion matrix Σ. Here β: $p \times q$ and Σ are unknown parameters and $\mathbf{Y}_\alpha$'s are known $q \times 1$ vectors. Writing $\beta = (\beta_1' \beta_2')' = (\beta_{(1)}\beta_{(2)})$ with β_i: $p_i \times q(p_1 + p_2 = p)$ and $\beta_{(i)}$: $p \times q_i(q_1 + q_2 = q)$, we have constructed invariant confidence sequences for (i) β, (ii) $\beta_{(1)}$, (iii) β_1 when $\beta_2 = 0$ and (iv) $\sigma^2 = |\Sigma|$. This uses the basic ideas of Robbins (1970) and generalizes some of his and Lai's (1976) results. In the process alternative simpler solutions of some of Khan's results (1978) are obtained.

1. Introduction. The problem of deriving confidence sequences i.e., sequences of confidence regions which contain the true parameter for every sample size simultaneously at a specified confidence level, has been tackled by Robbins (1970) and Lai (1976) using likelihood ratio and generalized likelihood ratio martingales. Let, under P_θ, $\theta \in \Omega$, the random $p \times n$ matrix $X_{(n)} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)$ have the probability density $p_{n,\theta}(x_{(n)})$, $n \geq 1$, with respect to a σ-finite measure μ_n defined on the Borel sets of the space $\mathscr{X}_{(n)}$ of $X_{(n)}$. A family of subsets $\{R_n(X_{(n)}),\ n \geq m$ (some fixed positive integer)$\}$ of Ω is said to constitute a $(1-\alpha)$-level sequence of confidence sets for θ if $P_\theta\{\theta \in R_n(X_{(n)})$ for all $n \geq m\} \geq 1 - \alpha$, for all $\theta \in \Omega$. The construction of such a sequence of confidence sets is based on the following inequality due to Robbins (1970). Let F be a measure defined on Ω and

$$(1.1)\qquad Z_n = \begin{cases} \int_\Omega p_{n,\eta}(x_{(n)})\, dF(\eta)/p_{n,\theta}(x_{(n)}), & \text{if } p_{n,\theta}(x_{(n)}) > 0 \\ 0 & \text{if } p_{n,\theta}(x_{(n)}) = 0. \end{cases}$$

Then $\{Z_n, \mathscr{F}_n\}$ $n \geq 1$ is a P_θ-martingale ($\mathscr{F}_n$ is the Borel field generated by $X_{(n)}$) and satisfies the martingale inequality given by

$$(1.2)\qquad P_\theta\{Z_n \geq \delta \text{ for some } n \geq m\} \leq P_\theta\{Z_m \geq \delta\} + \delta^{-1} \int_{Z_{m<\delta}} Z_m\, dP_\theta \quad \text{for any} \quad \delta > 0.$$

Received March 1982; revised July 1983.

[1] The work of the authors is sponsored by the Air Force Office of Scientific Research under Contract F49620-82-K-001. Reproduction in whole or in part is permitted for any purpose of the United States Government.

AMS 1970 *subject classifications*. Primary 62F25, 62L10; secondary 62H99.

Key words and phrases. Confidence sequences, likelihood ratio martingales, multivariate normal distribution, maximal invariant.

To define invariant confidence sequences (Lai, 1976), let $\theta = (\theta_1, \theta_2)$ and let G be a group of tranformations on $\mathscr{X}_{(n)}$ leaving the family $\{P_{\theta_1,\theta_2}, (\theta_1, \theta_2) \in \Omega\}$ invariant. Let $\bar{G}$ be the induced group on Ω. We assume that G is such that $\bar{G}$ induces a transformation on the space of θ_1. Then, a sequence of confidence sets $\{R_n(x_{(n)}), n \geq m\}$ for θ_1 is said to be invariant under G if

$$R_n(g \cdot x_{(n)}) = \{\bar{g} \cdot \theta_1 : \theta_1 \in R_n(x_{(n)})\} \tag{1.3}$$

for all $n \geq m$, $x_{(n)} \in \mathscr{X}_{(n)}$ and $g \in G$.

Our object in this paper is to derive invariant confidence sequences for some parameters in the following probability model. Let $\mathbf{X}_1, \mathbf{X}_2, \cdots$ be independent normal p-vectors with $E\,\mathbf{X}_\alpha = \beta\,\mathbf{Y}_\alpha$, $\alpha = 1, 2, \cdots$ and the same p.d. dispersion matrix Σ. Here $\beta : p \times q$ and Σ are unknown parameters and $\mathbf{Y}_\alpha$'s are $q \times 1$ vectors which are regarded either as known or else the above distribution is to be understood conditionally given the $\mathbf{Y}_\alpha$'s. Let us write $\beta = (\beta_1' \beta_2')' = (\beta_{(1)} \beta_{(2)})$, where $\beta_i : p_i \times q (p_1 + p_2 = p)$, $\beta_{(i)} : p \times q_i (q_1 + q_2 = q)$, $i = 1, 2$. We want to construct sequences of invariant confidence sets for (i) β, (ii) $\beta_{(1)}$, (iii) β_1 when $\beta_2 = 0$ and (iv) $\sigma^2 = |\Sigma|$.

Some special cases of the above problems have been treated in the literature. Khan (1978) considered problem (i) when (a) $p = 1$ and (b) $q = 1$, $Y_\alpha = 1$. Our solution to problem (i), when specialized to case (b), is much simpler than Khan's and provides a natural multivariate analogue of the results of Robbins (1970) and Lai (1976). The motivation for problem (iii) in case (b) can be found in Giri (1968). We offer two solutions in this case. Problem (ii) in case (a) relates to a subset of the regression coefficients in a linear model situation.

In order to construct invariant confidence sequences, we reduce the data first by sufficiency and then by invariance. The following version of Stein's theorem, due to Hall, Wijsman and Ghosh (1965), is often used.

THEOREM 1. *Let, for each n, U_n be sufficient for $X_{(n)} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)$ and let there be a group G of transformations on $\mathscr{X}_{(n)}$. Let $T_n \equiv T(U_n)$ be a maximal invariant under the induced group of transformations on U_n-space. Then, under certain assumptions, for each n, T_n is sufficient for $(T_1, \cdots, T_n)$.*

In the applications of this theorem to our problem, it is not difficult to verify Assumption C of Hall, Wijsman and Ghosh (1965), which is sufficient for the above theorem to hold. We also use the representation theorem due to Wijsman (1967) (see also Kariya (1978, 1981a, 1981b)).

Problems (i)–(iii) are discussed in Section 2 and problem (iv) is treated in Section 3.

2. Invariant confidence sequences for regression coefficients.

2.1 *Problem* (i). Let $P_{\beta,\Sigma}$ be the probability measure under which the density of $X_{(n)} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)$ with respect to the Lebesgue measure is

$$(2\pi)^{-np/2} |\Sigma|^{-n/2} \exp\{-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}(X_{(n)} - \beta Y_{(n)})(X_{(n)} - \beta Y_{(n)})'\} \tag{2.1}$$

where $Y_{(n)} = (\mathbf{Y}_1, \cdots, \mathbf{Y}_n)$. It is assumed that $Y_{(n)}$ is of rank q for each $n \geq q$.

Define

$$\hat{\beta}_n = X_{(n)} Y'_{(n)} C_n^{-1}, \quad \text{where} \quad C_n = Y_{(n)} Y'_{(n)} \tag{2.2}$$

and

$$S_n = X_{(n)} X'_{(n)} - \hat{\beta}_n C_n \hat{\beta}'_n.$$

Then, for each $n \geq q$, $(\hat{\beta}_n, S_n)$ is sufficient for the above family of distributions. We now consider the group G_1 whose elements are g_1, $p \times p$ nonsingular matrices. The induced action of g_1 on $(\hat{\beta}_n, S_n)$ is $g_1 \cdot (\hat{\beta}_n, S_n) = (g_1 \hat{\beta}_n, g_1 S_n g'_1)$ and on (β, Σ) is $\bar{g}_1 \cdot (\beta, \Sigma) = (g_1 \beta, g_1 \Sigma g'_1)$. For each $n \geq p + q$, a maximal invariant under G_1 is $T_n = \hat{\beta}'_n S_n^{-1} \hat{\beta}_n$ and the joint distribution of $T_{p+q}, T_{p+q+1}, \cdots, T_n$ will depend on (β, Σ) only through $\eta'\eta$ where $\eta = \Sigma^{-1/2}\beta$. Let $h_{n,\eta}(t_{p+q}, \cdots, t_n)$ denote the joint density of $T_{p+q}, \cdots, T_n$. Because of Theorem 1, we have

$$h_{n,\eta}(t_{p+q}, \cdots, t_n)/h_{n,0}(t_{p+q}, \cdots, t_n) = p_{n,\eta}(t_n)/p_{n,0}(t_n)$$

where $p_{n,\eta}(t_n)$ is the density of T_n. Hence, for any measure F on the space of η, if we define

$$Z_n = \int \frac{p_{n,\eta}(t_n)}{p_{n,0}(t_n)} \, dF(\eta), \tag{2.3}$$

then $\{Z_n, \mathscr{F}_n\}_{n \geq p+q}$ is a martingale under $P_{0,\Sigma}$, $\mathscr{F}_n$ being the Borel field generated by $(T_{p+q}, \cdots, T_n)$, $n \geq p + q$.

The density of $(\hat{\beta}_n, S_n)$, $n \geq p + q$, at $(\beta, \Sigma) = (\eta, I)$ is given by (Anderson, 1958, page 183)

$$K_n | C_n |^{p/2} \exp\{- ½ \operatorname{tr}(\hat{\beta}_n - \eta) C_n (\hat{\beta}_n - \eta)' - ½ \operatorname{tr} S_n\} \, | S_n |^{(n-p-q-1)/2} \tag{2.4}$$

where K_n stands for a generic constant. Using Wijsman's representation theorem (Wijsman, 1967), we get that

$$\frac{p_{n,\eta}(t_n)}{p_{n,0}(t_n)} \tag{2.5}$$

$$= \frac{\int_{A \in \mathscr{A}} \exp\{- ½ \operatorname{tr}(A\hat{\beta}_n - \eta) C_n (A\hat{\beta}_n - \eta)' - ½ \operatorname{tr} A S_n A'\} | AA' |^{(n-p)/2} \, dA}{\int_{A \in \mathscr{A}} \exp\{- ½ \operatorname{tr} A (S_n + \hat{\beta}_n C_n \hat{\beta}'_n) A'\} \, | AA' |^{(n-p)/2} \, dA}$$

where $\mathscr{A}$ is the group of $p \times p$ nonsingular matrices. Now, choosing F such that $dF(\eta) = (2\pi)^{-pq/2} \, | C_m |^{-p/2} \, d\eta$ for some fixed positive integer $m \geq p + q$ and integrating (2.5) with respect to this F, we have

$$Z_n = \frac{| C_m |^{p/2}}{| Cn |^{p/2}} \cdot \frac{\int_{\mathscr{A}} \exp\{- ½ \operatorname{tr} A \, S_n A'\} \, | AA' |^{(n-p)/2} \, dA}{\int_{\mathscr{A}} \exp\{- ½ \operatorname{tr} A (S_n + \hat{\beta}_n C_n \hat{\beta}'_n) A'\} \, | AA' |^{(n-p)/2} \, dA}$$

$$= \{| C_m |^{p/2} / | C_n |^{p/2}\} \cdot \{| S_n + \hat{\beta}_n C_n \hat{\beta}'_n |^{n/2} / | S_n |^{n/2}\} \tag{2.6}$$

$$= \{| C_m |^{p/2} / | C_n |^{p/2}\} \, | I + C_n T_n |^{n/2}.$$

Since Z_n satisfies the inequality (1.2) with $\boldsymbol{\theta} = (0, \Sigma)$, replacing $X_{(n)}$ by $X_{(n)} -$

$\beta Y_{(n)}$ in (2.6), we therefore get

$$P_{\beta,\Sigma}\{|I_q + C_n(\hat{\beta}_n - \beta)' S_n^{-1}(\hat{\beta}_n - \beta)| \geq \delta^{2/n}\{|C_n|^{p/n}/|C_m|^{p/n}\} \text{ for some } n \geq m \geq p + q\} \leq P_{0,\Sigma}\{Z_m \geq \delta\} + \delta^{-1}\int_{Z_m<\delta} Z_m \, dP_{0,\Sigma}, \quad \text{for any} \quad \delta > 0. \tag{2.7}$$

Under $P_{0,\Sigma}$, $Z_m^{2/m}$ is distributed as $U_{p,q,m-q}^{-1}$, $U_{p,m,n}$ being the random variable distributed as the product of p independent Beta variables, $B((n - i + 1)/2, m/2)$, $i = 1, \cdots, p$. (See, for example, Anderson, 1958, page 194). Hence, taking $\delta = (1/\rho)^{m/2}$, we can write the right hand side (rhs) in (2.7) as

$$P(U_{p,q,m-q} \leq \rho) + \rho^{m/2}\int_{u>\rho} u^{-m/2} f_{p,q,m-q}(u)\, du \tag{2.8}$$

where f is the density of U. Using (2.7) and (2.8), one can obtain a sequence of invariant confidence sets for β. It is interesting to observe that the regions determined by (2.7) have the same shape as the confidence region determined by Wilk's Λ criterion which, in the fixed sample size problem, has smallest expected volume among all fully invariant confidence sets; see Hooper (1982, page 1290). Moreover, our measure dF is the same as his measure $m(d\gamma, \theta)$. In the following, we consider some special values of p and q, and present the corresponding inequalities (2.7) in simplified forms.

$p = 1$. Since $U_{1,q,m-q} =$ Beta $((m - q)/2, q/2)$, the second term in the rhs of (2.8) reduces to $2\rho^{(m-q)/2}(1 - \rho)^{q/2}/\{qB((m - q)/2, q/2)\}$, yielding the following inequality:

$$P_{\beta,\Sigma}[(\hat{\beta}_n - \beta)C_n(\hat{\beta}_n - \beta)' \leq \{(\rho^{-m}|C_n|/|C_m|)^{1/n} - 1\} \cdot S_n, \text{ for all } n \geq m \geq q + 1] \geq 1 - I_\rho((m - q)/2, q/2) - 2\rho^{(m-q)/2}(1 - \rho)^{q/2}/\{q\, B((m - q)/2, q/2)\}, \quad \rho > 0 \tag{2.9}$$

where $I_x(p, q) = P\{B(p, q) \leq x\}$. This is precisely the same as obtained by Khan (1978).

$q = 1$. Note that $U_{p,1,m-1}$ is distributed as Beta $((m - p)/2, p/2)$ (see e.g. Anderson, 1958). Hence, proceeding as above, a sequence of invariant confidence sets for $\beta(p \times 1)$ is obtained as

$$P_{\beta,\Sigma}[(\hat{\beta}_n - \beta)' S_n^{-1}(\hat{\beta}_n - \beta) \leq \{\rho^{-m/n}(C_n/C_m)^{p/n} - 1\}C_n^{-1} \text{ for all } n \geq m \geq p + 1] \geq 1 - I_\rho((m - p)/2, p/2) - 2\,\rho^{(m-p)/2}(1 - \rho)^{p/2}/\{pB((m - p)/2, p/2)\}, \quad \rho > 0. \tag{2.10}$$

Taking $Y_{(n)} = (1, 1, \cdots, 1)$: $1 \times n$ in the above, we obtain a $(1-\alpha)$-level sequence of invariant confidence sets for μ, the mean vector of a p-variate normal population $N_p(\mu, \Sigma)$ as

$$(2.11) \qquad R_n = \{\mu: n(\bar{\mathbf{X}}_{(n)} - \mu)' S_n^{-1}(\bar{\mathbf{X}}_{(n)} - \mu) \le C_{n,\alpha}\}, \quad n \ge m \ge p$$

where

$$\bar{\mathbf{X}}_{(n)} = \textstyle\sum_1^n \mathbf{X}_i/n, \quad S_n = \sum_1^n \mathbf{X}_i\mathbf{X}_i' - n\,\bar{\mathbf{X}}_{(n)}\bar{\mathbf{X}}_{(n)}',$$

$$C_{n,\alpha} = \rho_\alpha^{-m/n}(n/m)^{p/n} - 1, \quad \rho_\alpha \text{ being the value of } \rho$$

which makes the rhs of (2.10) equal to $(1-\alpha)$.

This sequence of confidence sets for μ is different from and much simpler than that obtained by Khan (1978) and is a natural multivariate analogue of the known univariate result of Robbins (1970) and Lai (1976). A reasonable criterion to compare the two sequences of confidence sets would be to look at the limiting behavior of the two sequences of (if necessary normalized) volumes of the corresponding confidence sets. The evaluation of this limit is straightforward in our case but extremely complicated in Khan's case because of its dependence on confluent hypergeometric functions and we have not attempted to do it here.

2.2 *Problem* (ii). Let us partition $\hat\beta_n$ and C_n, in the manner of β, as

$$(2.12) \qquad \hat\beta_n = (\hat\beta_{(1)n}\hat\beta_{(2)n}), \quad C_n = \begin{pmatrix} C_{(11)n} & C_{(12)n} \\ C_{(21)n} & C_{(22)n} \end{pmatrix}$$

and define $C_{(11.2)n} = C_{(11)n} - C_{(12)n}C_{(22)n}^{-1}C_{(21)n}$. Consider now the group G_2 whose elements are $g_2 = (A, B)$ where A is a $p \times p$ nonsingular matrix and B is a $p \times q_2$ arbitrary matrix. The induced action of G_2 on $(\hat\beta_n \cdot S_n)$ is $g_2(\hat\beta_n, S_n) = ((A\hat\beta_{(1)n}: B + A\hat\beta_{(2)n}), AS_nA')$ and on (β, Σ) is $\bar g_2 \cdot (\beta, \Sigma) = ((A\beta_{(1)}: B + A\beta_{(2)}), A\Sigma A')$. It is clear that under the group G_2 of transformations the problem of constructing a confidence sequence for $\beta_{(1)}$ remains invariant. It is not difficult to show that, for each $n \ge p + q$, a maximal invariant is $T_n = \hat\beta_{(1)n}'S_n^{-1}\hat\beta_{(1)n}$, and the joint density of $(T_{p+q}, \cdots, T_n)$ depends on (β, Σ) only through $\eta'\eta$ where $\eta = \Sigma^{-1/2}\beta_{(1)}$. Now the density of $(\hat\beta_{(1)n}, S_n)$ at $(\beta, \Sigma) = ((\eta:0), I)$ is given by

$$(2.13) \qquad \begin{aligned} &K_n |C_{(11.2)n}|^{p/2} \\ &\exp\{-\tfrac12 \operatorname{tr}(\hat\beta_{(1)n} - \eta)C_{(11.2)n}(\hat\beta_{(1)n} - \eta)' - \tfrac12 \operatorname{tr} S_n\}\, |S_n|^{(n-p-q-1)/2}. \end{aligned}$$

Hence, as before, using Theorem 1 and Wijsman's representation theorem, we find that

$$(2.14) \qquad Z_n = \{|C_{(11.2)m}|/|C_{(11.2)n}|\}^{p/2}\, |I + C_{(11.2)n}T_n|^{n/2}, \quad n \ge m \ge p + q$$

is a martingale under $P_{(0,\beta_{(2)}),\Sigma}$ relative to $\mathscr{F}_n$ (the Borel field generated by $(T_m, \cdots, T_n)$, $n \ge m \ge p + q$).

Under $P_{(0,\beta_{(2)}),\Sigma}$, $Z_m^{2/m}$ is distributed as $U_{p,q_1,m-q}^{-1}$. So replacing $X_{(n)}$ by $X_{(n)} - (\beta_{(1)}:0)Y_{(n)}$ in (2.14) and using the basic martingale inequality (1.2), we obtain the following.

$$P_{\beta,\Sigma}\{|I + C_{(11.2)n}(\hat{\beta}_{(1)n} - \beta_{(1)})' S_n^{-1}(\hat{\beta}_{(1)n} - \beta_{(1)})| \tag{2.15}$$
$$\le \rho^{-m/n}(|C_{(11.2)n}|/|C_{(11.2)m}|)^{p/n} \text{ for all } n \ge m \ge p + q\}$$
$$\ge 1 - P\{U_{p,q_1,m-q} \le \rho\} - \rho^{m/2} \int_{u>\rho} u^{-m/2} f_{p,q_1,m-q}(u)\, du, \quad 0 < \rho < 1.$$

2.3. *Problem* (iii). We provide two solutions in this case. The first solution is direct though involved and is based on invariance consideration parallel to that in Giri (1968). Let us consider the group G_3 of transformations whose elements g_3 are given by $g_3 = \binom{A_{11}\ A_{12}}{0\ \ A_{22}}$, $A_{ii}{:}p_i \times p_i$ nonsingular, $i = 1, 2$. The induced action of G_3 on $(\hat{\beta}_n, S_n)$ is $g_3 \cdot (\hat{\beta}_n, S_n) = (A\hat{\beta}_n, A\, S_n A')$ and that on $(\binom{\beta_1}{0}, \Sigma)$ is $\bar{g}_3 \cdot (\binom{\beta_1}{0}, \Sigma) = (\binom{A_{11}\beta_1}{0}, A\Sigma A')$. It is clear that the problem remains invariant. Let us partition $\hat{\beta}_n$ and S_n in the manner of β as follows:

$$\hat{\beta}_n = \begin{pmatrix} \hat{\beta}_{1n} \\ \hat{\beta}_{2n} \end{pmatrix}, \quad S_n = \begin{pmatrix} S_{11n} & S_{12n} \\ S_{21n} & S_{22n} \end{pmatrix} \tag{2.16}$$

and define

$$\hat{\beta}_{1\cdot 2n} = \hat{\beta}_{1n} - S_{12n}S_{22n}^{-1}\hat{\beta}_{2n}, \quad S_{11\cdot 2n} = S_{11n} - S_{12n}S_{22n}^{-1}S_{21n}. \tag{2.17}$$

It is easy to verify that, for each $n \ge p + q$, a maximal invariant is $T_n = (\hat{\beta}'_{1\cdot 2n}S_{11\cdot 2n}^{-1}\hat{\beta}_{1\cdot 2n}, \hat{\beta}'_{2n}S_{22n}^{-1}\hat{\beta}_{2n})$ and a corresponding maximal invariant in the parameter space is $\eta'\eta$ where $\eta = \Sigma_{11.2}^{-1/2}\beta_1(\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$. Let $p_{n,\eta}(T_n)$ denote the density of T_n at $(\binom{\beta_1}{0}, \Sigma) = (\binom{\eta}{0}, I)$. Denoting by $\mathscr{A}_i$ the group of $p_i \times p_i$ nonsingular matrices, $i = 1, 2$, and by defining $A = \binom{A_{11}\ A_{12}}{0\ \ A_{22}}$, $A_{(1)} = (A_{11}A_{12})$, and using Wijsman's representation theorem, we see that $p_{n,\eta}(T_n)/p_{n,0}(T_n)$ is equal to

$$\frac{\int_{\mathscr{A}_1\times\mathscr{A}_2\times R^{p_1p_2}} \exp\{-\tfrac12 \operatorname{tr}(A\, \hat{\beta}_n - \binom{\eta}{0})C_n(A\, \hat{\beta}_n - \binom{\eta}{0})' - \tfrac12 \operatorname{tr} A\, S_nA'\} \,|A_{11}A'_{11}|^{(n-p)/2}|A_{22}A'_{22}|^{(n-p)/2}\, dA_{11}\, dA_{22}\, dA_{12}}{\int_{\mathscr{A}_1\times\mathscr{A}_2\times R^{p_1p_2}} \exp\{-\tfrac12 \operatorname{tr} A(S_n + \hat{\beta}_nC_n\hat{\beta}'_n)A'\} \,|A_{11}A'_{11}|^{(n-p)/2}\, |A_{22}A'_{22}|^{(n-p_2)/2}\, dA_{11}\, dA_{22}\, dA_{12}}$$

$$= \frac{\int_{\mathscr{A}_1\times\mathscr{A}_2\times R^{p_1p_2}} \exp\{-\tfrac12 \operatorname{tr}(A_{(1)}\hat{\beta}_n - \eta)C_n(A_{(1)}\hat{\beta}_n - \eta)' - \tfrac12 \operatorname{tr} A_{(1)}S_nA_{(1)}'\} \,|\mathrm{A}_{11}A'_{11}|^{(n-p)/2}\, dA_{11}\, dA_{12}}{\int_{\mathscr{A}_1\times\mathscr{A}_2\times R^{p_1p_2}} \exp\{-\tfrac12 \operatorname{tr} A_{(1)}(S_n + \hat{\beta}_nC_n\bar{\beta}'_n)A_{(1)}'\} \,|A_{11}A'_{11}|^{(n-p)/2}\, dA_{11}\, dA_{12}} \tag{2.18}$$

Now, we choose F such that $dF(\eta) = (2\pi)^{-p_1q/2}\,|C_m|^{p_1/2}\, d\eta$. Then, the numerator in (2.18), after being integrated with respect to this F, becomes

$$\left(\frac{|C_m|}{|C_n|}\right)^{p_1/2} \int_{\mathscr{A}_1\times R^{p_1p_2}} \exp\left\{-\frac{1}{2}\operatorname{tr} A_{11}S_{11n}A_{11}' - \frac{1}{2}\operatorname{tr} A_{12}S_{22n}A_{12}' - \operatorname{tr} A_{11}S_{12n}A_{12}'\right\} |A_{11}A_{11}'|^{(n-p)/2}\,dA_{11}\,dA_{12}$$

$$(2.19)\qquad = (2\pi)^{(p_1p_2/2)}\left(\frac{|C_m|}{|C_n|}\right)^{p_1/2} |S_{22n}|^{-p_1/2}\,|S_{11.2n}|^{-(n-p_2)/2} \cdot \int_{\mathscr{A}_1} \exp\left\{-\frac{1}{2}\operatorname{tr} A_{11}A_{11}'\right\} |A_{11}A_{11}'|^{(n-p)/2}\,dA_{11}.$$

Letting $S_n^* = S_n + \hat{\beta}_n C_n \hat{\beta}_n'$ and partitioning S_n^* in the manner of S_n, we get similarly the denominator in (2.18) as

$$(2.20)\qquad (2\pi)^{p_1p_2/2}\,|S_{22n}^*|^{-p_1/2}\,|S_{11.2n}^*|^{-(n-p_2)/2} \int_{\mathscr{A}_1} \exp\left\{-\frac{1}{2}\operatorname{tr} A_{11}A_{11}'\right\} |A_{11}A_{11}'|^{(n-p)/2}\,dA_{11}.$$

So, from (2.18)–(2.20) and Theorem 1, it is seen that

$$(2.21)\qquad Z_n = (|C_m|/|C_n|)^{p_1/2}(|S_{22n}^*|/|S_{22n}|)^{p_1/2}(|S_{11.2n}^*|/|S_{11.2n}|)^{(n-p_2)/2}, \quad n \ge m \ge p+q$$

is a martingale under $P_{(\beta_1,0),\Sigma}$ relative to $\mathscr{F}_n$, the Borel field generated by $(T_m, \cdots, T_n)$. On simplification, Z_n is equal to

$$(2.22)\qquad (|C_m|/|C_n|)^{p_1/2}\,|I + C_n\hat{\beta}_{2n}'S_{22n}^{-1}\hat{\beta}_{2n}|^{p_1/2} \cdot |I + (I + C_n\hat{\beta}_{2n}'S_{22n}^{-1}\hat{\beta}_{2n})^{-1}C_n\hat{\beta}_{1.2n}'S_{11.2n}^{-1}\hat{\beta}_{1.2n}|^{(n-p_2)/2}.$$

It is not difficult to show that for each $n \ge m \ge p + q$, conditionally given $\hat{\beta}_{2n}'S_{22n}^{-1}\hat{\beta}_{2n}$, $\hat{\beta}_{1.2n}$ and $S_{11.2n}$ are independently distributed as $N_{p_1q}(0, \Sigma_{11.2} \otimes (C_n^{-1} + \hat{\beta}_{2n}' S_{22n}^{-1}\hat{\beta}_{2n})^{-1})$ and $W_{p_1q}(\Sigma_{11.2}, n - p_2 - q)$ respectively. Hence, for each $n \ge m \ge p + q$, Z_n is distributed as

$$(2.23)\qquad Z_n \sim (|C_m|/|C_n|)^{p_1/2}U_{p_1,q,n-p_2-q}^{-(n-p_2)/2} \cdot U_{p_2,q,n-q}^{p_1/2}$$

where $U_{p_1,q,n-p_2-q}$ and $U_{p_2,q,n-q}$ are independently distributed. (2.23) for $n = m$ can be used to get the rhs in (1.2). Finally, a confidence sequence for β_1 is obtained by replacing $X_{(n)}$ in (2.22) by $X_{(n)} - \binom{\beta_1}{0} Y_{(n)}$.

Our second solution is obtained if we proceed in a different way, reducing this problem to one similar to Problem (ii). For this, let us partition $X_{(n)}$ as $X_{(n)} = (X_{1(n)}'X_{2(n)}')'$, and assume $X_{2(n)}$ to be fixed. Let $P_{\beta_1,\Sigma}^*$ be the underlying conditional probability. Thus, under $P^*_{\beta_1,\Sigma}$, columns of $X_{1(n)}$ are independently distributed as p_1-variate normal with common dispersion matrix $\Sigma_{11.2}$ and with $E(X_{1(n)}/X_{2(n)}) = (\beta_1 \Sigma_{12} \Sigma_{22}^{-1}) Y^*_{(n)}$, where $Y^*_{(n)} = (Y'_{(n)} X'_{2(n)})'$: $(q + p_2) \times n$. Let us define

$$C^*_n = Y^*_{(n)} Y^{*\prime}_{(n)} = \begin{pmatrix} C^*_{(11)n} & C^*_{(12)n} \\ C^*_{(21)n} & C^*_{(22)n} \end{pmatrix}, \quad C^*_{(11)n}: q \times q$$

$$C^*_{(11.2)n} = C^*_{(11)n} - C^*_{(12)n} C^{*-1}_{(22)n} C^*_{(21)n} \tag{2.24}$$

$$\hat\beta^*_n = X_{1(n)} Y^{*\prime}_{(n)} C^{*-1}_n = (\hat\beta^*_{(1)n} \hat\beta^*_{(2)n}), \quad \hat\beta^*_{(1)n}: p_1 \times q$$

$$S^*_n = X_{1(n)} X'_{1(n)} - \hat\beta^*_n C^*_n \hat\beta'^*_n.$$

Then, arguing in the same manner as we did in the case of Problem (ii), it is seen that the sequence

$$Z_n = (|C^*_{(11.2)m}|/|C^*_{(11.2)n}|)^{p_1/2} \, |I + C^*_{(11.2)n} \hat\beta^{*\prime}_{(1)n} S^{*-1}_n \hat\beta^*_{(1)n}|^{n/2}, \tag{2.25}$$
$$n \geq m \geq p + q,$$

is a martingale under $P^*_{0,\Sigma}$, the underlying sequence of Borel fields being $\mathscr{F}_n$ = Borel field generated by $\hat\beta^{*\prime}_{(1)n} S^{*-1}_n \hat\beta^*_{(1)n}$, $n \geq m \geq p + q$. Here, under $P^*_{0,\Sigma}$, $Z^{2/m}_m$ is distributed as $U^{-1}_{p_1,q,m-p_2-q}$, which, being independent of $X_{2(m)}$, is also the distribution under $P_{0,\Sigma}$. So, using (2.15), we get

$$P^*_{\beta_1,\Sigma}\Bigg\{ |I + C^*_{(11.2)n} (\hat\beta^*_{(1)n} - \beta_1)' S^{*-1}_n (\hat\beta^*_{(1)n} - \beta_1)| \leq \frac{1}{\rho^m} \left(\frac{|C^*_{(11.2)n}|}{|C^*_{(11.2)m}|} \right)^{p_1/n} \text{ for all } n \geq m \geq p + q \Bigg\} \tag{2.26}$$
$$\geq 1 - P(U_{p_1,q,m-p_2-q} \leq \rho) - \rho^{m/2} \int_{u>\rho} u^{-m/2} f_{p_1,q,m-p_2-q}(u)\, du.$$

To derive the unconditional confidence sequence for β_1, it is enough to note that the rhs in (2.26) is independent of $X_{(2)}$'s. Moreover, after a little bit of algebra, it can be checked that

$$\hat\beta^*_{(1)n} = \hat\beta_{1.2n}, \quad S^*_n = S_{11.2n} \quad \text{and} \quad C^{*-1}_{(11.2)n} = C^{-1}_n + \hat\beta'_{2n} S^{-1}_{22n} \hat\beta_{2n}. \tag{2.27}$$

Hence, integrating the left hand side of (2.26) with respect to the probability measure on the space of $X_{1(\infty)}$ and using (2.27), we find that a $(1 - \alpha)$-level sequence of confidence sets for β_1 when $\beta_2 = 0$ is given by

$$R_n = \{\beta_1: |I + (C^{-1}_n + \hat\beta'_{2n} S^{-1}_{22n} \hat\beta_{2n})^{-1} (\hat\beta_{1.2n} - \beta_1)' S^{-1}_{11.2n} (\hat\beta_{1.2n} - \beta_1)| \tag{2.28}$$
$$\leq (\rho^{-m}_\alpha (|C^{-1}_m + \hat\beta'_{2m} S^{-1}_{22m} \hat\beta_{2m}|/|C^{-1}_n + \hat\beta'_{2n} S^{-1}_{22n} \hat\beta_{2n}|)^{p_1})^{1/n}\},$$

ρ_α being the value of ρ which makes the rhs of (2.26) equal to $1 - \alpha$.

3. Invariant confidence sequence for $|\Sigma|$. Consider the group G_4 with elements $g_4 = (B, T)$ where B is a $p \times q$ matrix and T is a nonsingular lower triangular matrix such that $|T|^2 = 1$. The induced action of G_4 on $(\hat{\beta}_n, S_n)$ is $g_4 \cdot (\hat{\beta}_n, S_n) = (B + \hat{\beta}_n, T S_n T')$ and the corresponding action on (β, Σ) is $\bar{g}_4 \cdot (\beta, \Sigma) = (B + \beta, T \Sigma T')$. For each $n \geq p + q$, a maximal invariant under G_4 can be shown to be $T_n = |S_n|$ (see Eaton, 1967), the corresponding maximal invariant in the parameter space being $\sigma^2 = |\Sigma|$. From Anderson (1958), we know that $T_n \sim \sigma^2 \prod_1^p \chi^2_{v-i+1}$, χ^2's being independent, $v = n - q$. Define for some measure F over $(0, \infty)$

$$Z_n = \int_0^\infty \frac{p_{n,\sigma}(T_n)}{p_{n,1}(T_n)}\, dF\,(1/\sigma). \tag{3.1}$$

Then, by Theorem 1, $\{Z_n, \mathscr{F}_n\}_{n \geq p+q}$ is a martingale under $P_{0,I}$ where $\mathscr{F}_n$ is the Borel field generated by $(T_{p+q}, \cdots, T_n)$, $n \geq p + q$. The martingale inequality (1.2), upon replacing T_n by T_n/σ^2 in the expression for Z_n in its right hand side, can be used to obtain a sequence of confidence intervals for σ^2. We consider below a special case $p = 2$.

When $p = 2$, $T_n \sim (\sigma^2/4)\chi^4_{2v-2}$ (see e.g., Srivastava and Khatri, 1979), with the density

$$p_{n,\sigma^2}(T_n) = \frac{1}{2\Gamma(v-1)} \left(\frac{1}{\sigma}\right)^{v-1} \exp(-\sqrt{T_n}/\sigma)\, \frac{v-3}{T_n^2}. \tag{3.2}$$

We choose F such that $dF(1/\sigma) = (1/\Gamma(m-q))\, d(1/\sigma)$ for some fixed positive integer $m \geq q + 2$.

Let $t_1(n, m, \delta)$ and $t_2(n, m, \delta)$ $(t_1 < t_2)$ be the solutions of

$$\exp(t)t^{-n+q} = \delta\Gamma(m-q)/\Gamma(n-q), \quad n \geq m \geq q+2, \quad \delta > 0. \tag{3.3}$$

Then after straightforward simplification we have the following:

$$\begin{aligned} P_{\beta,\Sigma}\{T_n/t_2^2(n, m, \delta) \leq \sigma^2 \leq T_n/t_2^2(n, m, \delta),\ \forall\, n \geq m \geq q+2\} \\ \geq P\{\chi^2_{2m-2q-2} \leq 2\, t_2(m, m, \delta)\} - P\{\chi^2_{2m-2q-2} \leq 2\, t_1(m, m, \delta)\} \\ - (1/\delta(m-q-1))\{1/t_1(m, m, \delta) - 1/t_2(m, m, \delta)\}, \quad \delta > 0. \end{aligned}$$

Acknowledgment. We sincerely thank the referee for many helpful comments leading to elimination of an incorrect result and improved presentation.

REFERENCES

ANDERSON, T. W. (1958). *Introduction to Multivariate Statistical Analysis.* Wiley, New York.

EATON, M. L. (1967). The generalized variance: Testing and ranking problem. *Ann. Math. Statist.* **38** 941–943.

GIRI, N. (1968). Locally and asymptotically minimax tests of a multivariate problem. *Ann. Math. Statist.* **39** 171–178.

HALL, W. J., WIJSMAN, R. A. and GHOSH, J. K. (1965). The relationship between sufficiency and invariance with application in sequential analysis. *Ann. Math. Statist.* **36** 575–614.

HOOPER, P. M. (1982). Invariant confidence sets with smallest expected measure. *Ann. Statist.* **10** 1283–1294.

KARIYA, T. (1978). The general MANOVA problem. *Ann. Statist.* **6** 200–214.

KARIYA, T. (1981a). Tests for the independence between two seemingly unrelated regression equations. *Ann. Statist.* **9** 381–390.

KARIYA, T. (1981b). Robustness of multivariate tests. *Ann. Statist.* **9** 1267–1275.

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

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

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

SRIVASTAVA, M. S. and KHATRI, C. G. (1979). *An Introduction to Multivariate Statistics.* North Holland, New York.

WIJSMAN, R. A. (1967). Cross-sections of orbits and their applications to densities of maximal invariants. *Fifth Berkeley Sym. Math. Statist. Probab.* **1** 389–400.

DEPARTMENT OF MATHEMATICS AND STATISTICS
FACULTY OF ARTS AND SCIENCES
UNIVERSITY OF PITTSBURGH
PITTSBURGH, PENNSYLVANIA 15260

DEPARTMENT OF STATISTICS
TEMPLE UNIVERSITY
PHILADELPHIA, PENNSYLVANIA 19122