Statistica Sinica 2(1992), 69-90

ASYMPTOTIC APPROXIMATIONS FOR LIKELIHOOD RATIO TESTS AND CONFIDENCE REGIONS FOR A CHANGE-POINT IN THE MEAN OF A MULTIVARIATE NORMAL DISTRIBUTION

Barry James, Kang Ling James and David Siegmund

University of Minnesota, University of Minnesota and Stanford University

Abstract: For independent, d-dimensional normally distributed observations we give a tail approximation for the significance level of the likelihood ratio test of no change in the mean vector against the alternative of exactly one change. Assuming there is exactly one change in the mean vector, we obtain conditional likelihood ratio confidence regions for the change-point and joint regions for the change-point and size of the change. For the significance level we compare our approximation numerically with the improved Bonferroni upper bound of Srivastava and Worsley (1986). For our probability calculations we adapt the method of Woodroofe (1976, 1978).

Key words and phrases: Change-point, boundary crossing probability, likelihood ratio test.

1. Introduction

The subject of this paper is tests and confidence regions for a change-point in the mean vector of m independent d-dimensional normal random variables with an unknown but fixed covariance matrix Σ.

In testing the hypothesis of no change against the alternative of exactly one change, our main result is a tail approximation for the significance level of the likelihood ratio test, generalizing the one dimensional case of our earlier paper (James, James and Siegmund (1988)). The proof, however, is *not* a generalization of the method of our earlier paper. It is inspired by the method of Woodroofe (1976, 1978), which it resembles in overall design although the detailed calculations, e.g., in Lemma 4, are different. This approximation is proved in Section 2. Its numerical accuracy and a comparison to the Bonferroni-like upper bound of Srivastava and Worsley (1986) are the subject of Section 3.

Sections 4 and 5 are concerned with likelihood ratio confidence regions for a change-point and joint regions for the change-point and amount of change. For related recent research see Worsley (1986), Siegmund (1985, 1988), James, James and Siegmund (1988), and Kim and Siegmund (1989). The latter two papers are specifically concerned with changes in the mean of univariate normal observations having an unknown variance and hence face some of the same problems we encounter in this paper. However, we deal with these problems differently and, we believe, more satisfactorily. Like these earlier papers we also condition on statistics which are sufficient for unknown nuisance parameters when a specific value is hypothesized for the change-point, but unlike James, James and Siegmund (1988), who carry out their conditioning in two stages, we condition directly on the minimal sufficient statistic. In the case of joint confidence regions for the change-point and amount of change, we differ from Kim and Siegmund (1989) by conditioning on the minimal sufficient statistic for the unknown nuisance parameters and hence obtain a region which in principle is exact. Again we adapt Woodroofe's method for the calculations; Lemmas 10–13 appear to have no antecedents in Woodroofe's work.

It is interesting that although our computations are considerably more complicated than in the univariate case, the resulting approximations are not. However, because of the large amount of conditioning involved and the iterative nature of the determination of the confidence regions, it seems difficult to approach this problem numerically or by Monte Carlo methods without a substantial increase in computational speed over what is readily available today. See Worsley (1983, 1986) for numerical treatment of much simpler single parameter cases.

2. Approximate Tail Probabilities for the Modified Likelihood Ratio Test

In this section we obtain an approximation for the tail probabilities of the distribution of the likelihood-ratio test statistic under the null hypothesis of no change-point.

Assume $X_1, X_2, \ldots, X_m$ are independent d-dimensional random (column) vectors having a multivariate normal distribution with the same unknown covariance matrix: $X_i \sim N(\mu_i, \Sigma)$, $i = 1, \ldots, m$. We wish to test the null hypothesis that all of the means are equal to some unknown μ, $H : \mu_1 = \cdots = \mu_m = \mu$, versus the alternative of a single change-point, $K : \mu_1 = \cdots = \mu_\rho \neq \mu_{\rho+1} = \cdots = \mu_m$ for some $1 \leq \rho \leq m-1$.

A modification of the likelihood ratio test uses the test statistic $T_m = \max_{m_0 \leq n \leq m_1} \frac{m}{n(m-n)}(S_n - \frac{n}{m}S_m)'(U_m - \frac{1}{m}S_m S_m')^{-1}(S_n - \frac{n}{m}S_m)$, where $S_n = X_1 + \cdots + X_n$, $U_m = X_1 X_1' + \cdots + X_m X_m'$ and $1 \leq m_0 < m_1 \leq m$. The actual likelihood ratio test is based on T_m when $m_0 = 1$ and $m_1 = m-1$. Under the null hypothesis, the distribution of T_m does not depend on μ and Σ; thus, in the calculations below under P_0 we may assume that $\mu = 0$ and $\Sigma = I$. Furthermore, Basu's Theorem (Lehmann (1986), Theorem 5.2) tells us that under the null hypothesis T_m is independent of the complete sufficient statistics S_m and U_m. Therefore, if $0 < c < 1$,

$$\begin{aligned} P_0(T_m \geq c^2) &= P_0\{T_m \geq c^2 | S_m = 0, U_m = mI\} \\ &= P_0\Big\{ \max_{m_0 \leq n \leq m_1} \frac{m}{n(m-n)} S_n'(mI)^{-1} S_n \geq c^2 | S_m = 0, U_m = mI \Big\} \\ &= P_{0,mI}^{(m)} \Big\{ \max_{m_0 \leq n \leq m_1} \frac{\|S_n\|}{[n(1-\frac{n}{m})]^{\frac{1}{2}}} \geq cm^{\frac{1}{2}} \Big\}, \end{aligned}$$

where $P_{x,y}^{(m)}(A) = P_0(A|S_m = x, U_m = y)$ for events A determined by $X_1, \ldots, X_m$.

By breaking down the above probability according to the last index for which the term in the maximum exceeds $b = cm^{\frac{1}{2}}$, we obtain

$$\begin{aligned} P_0(T_m \geq c^2) = P_{0,mI}^{(m)} \Big\{ \frac{\|S_{m_1}\|}{[m_1(1-\frac{m_1}{m})]^{\frac{1}{2}}} \geq b \Big\} + \sum_{n=m_0}^{m_1-1} \int_{\|\xi\| \geq b\{n(1-n/m)\}^{\frac{1}{2}}} \\ P_{0,mI}^{(m)} \Big\{ \max_{n < i \leq m_1} \frac{\|S_i\|}{[i(1-\frac{i}{m})]^{\frac{1}{2}}} < b | S_n = \xi \Big\} P_{0,mI}^{(m)} \{S_n \in d\xi\}, \end{aligned} \tag{2.1}$$

where $P(S \in d\xi)$ means $f_S(\xi) d\xi_1 \cdots d\xi_d$, with f_S the probability density function of S. By developing approximations for the probabilities and the densities in the above integrals, we obtain the following theorem, whose proof will follow after a series of lemmas.

Theorem 1. *Assume $m_0/m \to t_0$ and $m_1/m \to t_1$ as $m \to \infty$, where $0 \leq t_0 < t_1 \leq 1$. Let ν be the function defined for $t > 0$ by*

$$\nu(t) = 2t^{-2} \exp\Big\{ -2 \sum_{n=1}^{\infty} n^{-1} \Phi(-tn^{\frac{1}{2}}/2) \Big\},$$

where Φ is the standard normal distribution function. Then for $0 < c < 1$ and $b = cm^{\frac{1}{2}}$,

$$\begin{aligned} &P_{0,mI}^{(m)} \Big\{ \max_{m_0 \leq n \leq m_1} \frac{\|S_n\|}{[n(1-\frac{n}{m})]^{\frac{1}{2}}} \geq b \Big\} \\ &\sim \frac{1}{\Gamma(\frac{d}{2})} \Big(\frac{mc^2}{2} \Big)^{\frac{d}{2}} (1-c^2)^{(m-d-3)/2} \int_{t_0}^{t_1} \frac{1}{t(1-t)} \nu \Big(\frac{c}{\{t(1-t)(1-c^2)\}^{\frac{1}{2}}} \Big) dt \end{aligned} \tag{2.2}$$

as $m \to \infty$.

Lemma 1. *For all ξ such that $\|\xi\| < \{n(m-n)\}^{\frac{1}{2}}$,*

$$f_{S_n}(\xi | S_m = 0, U_m = mI) = \{\pi n(m-n)\}^{-d/2} \frac{\Gamma_d(\frac{m-1}{2})}{\Gamma_d(\frac{m-2}{2})} \left(1 - \frac{\|\xi\|^2}{n(m-n)}\right)^{(m-d-3)/2},$$

where Γ_d is the d-variate gamma function defined by

$$\Gamma_d\left(\frac{1}{2}\alpha\right) = \pi^{d(d-1)/4} \prod_{j=1}^{d} \Gamma\left(\frac{1}{2}[\alpha + 1 - j]\right).$$

Proof. Obviously

$$P_{\eta,\lambda}^{(m)}\{S_n \in d\xi\} = \frac{P_0\{S_n \in d\xi, S_m \in d\eta, U_m \in d\lambda\}}{P_0\{S_n \in d\eta, U_m \in d\lambda\}}.$$

The numerator equals the joint density of the independent random variables $S_n, S_m - S_n, U_m - S_n S_n'/n - (S_m - S_n)(S_m - S_n)'/(m-n)$ evaluated at ξ, $\eta - \xi$, $\lambda - \xi\xi'/n - (\eta - \xi)(\eta - \xi)'/(m-n)$. A similar result holds for the denominator. By substitution of the appropriate normal and Wishart densities, we obtain

$$\begin{aligned} &f_{S_n}(\xi | S_m = \eta, U_m = \lambda) \\ &= \frac{\Gamma_d(\frac{m-1}{2})}{\Gamma_d(\frac{m-2}{2})} \left\{\frac{m}{\pi n(m-n)}\right\}^{\frac{d}{2}} \frac{[\det\{\lambda - \frac{1}{n}\xi\xi' - \frac{1}{(m-n)}(\eta - \xi)(\eta - \xi)'\}]^{(m-d-3)/2}}{\{\det(\lambda - \frac{1}{m}\eta\eta')\}^{(m-d-2)/2}}. \end{aligned}$$

Substitution of $\eta = 0$ and $\lambda = mI$ yields the result of the lemma.

Lemma 2. *If $\xi = [c\{n(m-n)\}^{\frac{1}{2}} + x]\alpha$ for $x > 0$ and α a fixed unit vector, and if $\frac{n}{m} \to t$ as $m \to \infty$, $0 < t < 1$, then*

$$\begin{aligned} &f_{S_n}(\xi | S_m = 0, U_m = mI) \\ &\sim \left(\frac{1}{2\pi m t(1-t)}\right)^{\frac{d}{2}} (1 - c^2)^{(m-d-3)/2} \exp\left[-\frac{cx}{(1-c^2)\{t(1-t)\}^{\frac{1}{2}}}\right] \end{aligned}$$

as $m \to \infty$.

Proof. $(1 - \|\xi\|^2/\{n(m-n)\})^{(m-d-3)/2} = (1 - [c\{n(m-n)\}^{\frac{1}{2}} + x]^2/\{n(m-n)\})^{(m-d-3)/2} \sim (1-c^2)^{(m-d-3)/2} \exp\{-cx/[(1-c^2)\{t(1-t)\}^{\frac{1}{2}}]\}$. Also,

$$\Gamma_d(\{m-1\}/2)/\Gamma_d(\{m-2\}/2) \sim (m/2)^{d/2}.$$

Substituting these relations into the result of Lemma 1 completes the proof.

Lemma 3. *If C_m and D_m are $d \times d$ matrices such that $C_m \to C$ and $D_m \to D$ as $m \to \infty$, where $I - C$ is invertible, then*

$$\lim_{m\to\infty} \left\{ \frac{\det(I - C_m - \frac{1}{m} D_m)}{\det(I - C_m)} \right\}^{m/2} = \exp\{-\mathrm{tr}[D(I - C)^{-1}]/2\}.$$

Proof. The proof follows at once from the observation that

$$\begin{aligned} \det\big(I - C_m - \frac{1}{m} D_m\big)/\det(I - C_m) &= \det\{I - \frac{1}{m} D_m (I - C_m)^{-1}\} \\ &= 1 - \frac{1}{m}\mathrm{tr}\{D_m (I - C_m)^{-1}\} + O\Big(\frac{1}{m^2}\Big). \end{aligned}$$

Because of conditioning on the pair (S_m, U_m), the following calculation is substantially more elaborate than the corresponding results of Woodroofe (1976, 1978). See also Lemma 8.

Lemma 4. *For each $k = 1, 2, \ldots, x > 0$ and α a unit vector, $\mathcal{L}(X_{n+1}, \ldots, X_{n+k} | S_n = [c\{n(m-n)\}^{\frac{1}{2}} + x]\alpha, S_m = 0, U_m = mI)$ converges to the law of k independent, identically distributed $N(-c(t/[1-t])^{\frac{1}{2}}\alpha, I - c^2\alpha\alpha')$ random vectors, with $t = \lim_{m\to\infty}(n/m)$.*

Proof. We may assume, without loss of generality (by sufficiency), that $X_1, \ldots, X_n$ are i.i.d. $N(0, I)$. For economy of notation, we write

$$\xi = [c\{n(m-n)\}^{\frac{1}{2}} + x]\alpha,$$

as in Lemma 2. Straightforward calculations show that

$$\begin{aligned} & f_{X_{n+1}\cdots X_{n+k}}(x_1, \ldots, x_k | S_n = \xi, S_m = 0, U_m = mI) \\ &= \left(\frac{m-n}{m-n-k}\right)^{\frac{d}{2}} 2^{dk/2} \frac{\Gamma_d(\frac{m-2}{2})}{\Gamma_d(\frac{m-k-2}{2})} \frac{\Pi_{j=1}^d [\Pi_{i=1}^k \varphi(x_{ij}) \varphi\{(\xi_j + \sum_{i=1}^k x_{ij})/(m-n-k)^{\frac{1}{2}}\}]}{\Pi_{j=1}^d \varphi\{\xi_j/(m-n)^{\frac{1}{2}}\}} \\ &\quad \cdot \frac{\exp\{\frac{1}{2}(\mathrm{tr}B_m - \mathrm{tr}A_m)\}(\det A_m)^{(m-k-d-3)/2}}{(\det B_m)^{(m-d-3)/2}}, \end{aligned} \tag{2.3}$$

where φ is the standard normal density,

$$A_m = mI - \frac{1}{n}\xi\xi' - \sum_{i=1}^k x_i x_i' - \frac{1}{m-n-k}\Big(\xi + \sum_{i=1}^k x_i\Big)\Big(\xi + \sum_{i=1}^k x_i\Big)',$$

and

$$B_m = mI - \frac{m}{n(m-n)}\xi\xi'.$$

To facilitate our calculations, we note that, by the symmetry of the problem, it suffices to prove the lemma for $\alpha = e_1 = (1, 0, \ldots, 0)'$. Since

$$\frac{1}{m}B_m = I - \frac{1}{n(m-n)}\xi\xi' \to I - c^2\alpha\alpha'$$

and

$$\det(I - c^2\alpha\alpha') = 1 - c^2,$$

$$\frac{(\det A_m)^{(m-k-d-3)/2}}{(\det B_m)^{(m-d-3)/2}} \sim \frac{1}{m^{dk/2}(1-c^2)^{\frac{k}{2}}}\left\{\frac{\det(\frac{1}{m}A_m)}{\det(\frac{1}{m}B_m)}\right\}^{(m-k-d-3)/2}$$

We now apply Lemma 3 with $C_m = \frac{1}{n(m-n)}\xi\xi'$, $C = c^2\alpha\alpha'$,

$$D_m = B_m - A_m = -\frac{1}{m-n}\xi\xi' + \sum_{i=1}^{k} x_i x_i' + \frac{1}{m-n-k}\Big(\xi + \sum_{i=1}^{k} x_i\Big)\Big(\xi + \sum_{i=1}^{k} x_i\Big)',$$

and

$$D = \frac{kc^2 t}{1-t}e_1 e_1' + \sum_{i=1}^{k} x_i x_i' + c\Big(\frac{t}{1-t}\Big)^{\frac{1}{2}}\Big\{\Big(\sum_{i=1}^{k} x_i\Big)e_1' + e_1\Big(\sum_{i=1}^{k} x_i\Big)'\Big\}.$$

This yields

$$\begin{aligned}\text{LHS}(2.3) &\sim \frac{1}{m^{dk/2}(1-c^2)^{\frac{k}{2}}}\exp\Big\{-\frac{1}{2}\text{tr}[D(I-C)^{-1}]\Big\}\\ &= \frac{1}{m^{dk/2}(1-c^2)^{\frac{k}{2}}}\exp\Big\{-\frac{1}{2}\Big[\frac{kc^2 t}{(1-c^2)(1-t)} + \frac{\sum_{i=1}^{k} x_{i1}^2}{1-c^2}\\ &\quad + \sum_{j=2}^{d}\sum_{i=1}^{k} x_{ij}^2 + \frac{2c}{1-c^2}\Big(\frac{t}{1-t}\Big)^{\frac{1}{2}}\sum_{i=1}^{k} x_{i1}\Big]\Big\}.\end{aligned}$$

Since

$$\Gamma_d(\{m-2\}/2)\Gamma_d(\{m-k-2\}/2) \sim (m/2)^{dk/2}$$

and also

$$\begin{aligned}&\prod_{j=1}^{d}\Big\{\varphi\Big(\{\xi_j + \sum_{i=1}^{k} x_{ij}\}/(m-n-k)^{\frac{1}{2}}\Big)\prod_{i=1}^{k}\varphi(x_{ij})\Big\}\exp\Big(-\frac{1}{2}\text{tr}A_m\Big)\\ =&(2\pi)^{-d(k+1)/2}\exp\Big\{-\frac{1}{2}\text{tr}(mI) + \frac{1}{2}\text{tr}\big(\frac{1}{n}\xi\xi'\big)\Big\}\end{aligned}$$

and

$$\prod_{j=1}^{d} \varphi\Big(\xi_j/(m-n)^{\frac{1}{2}}\Big) \exp\Big\{-\frac{1}{2}\operatorname{tr}(B_m)\Big\}$$
$$=(2\pi)^{-\frac{d}{2}} \exp\{-\operatorname{tr}(mI)/2+\operatorname{tr}(n^{-1}\xi\xi')/2\},$$

substitution into (2.3) yields the lemma when $\alpha = e_1$ and, therefore, in full generality.

Lemma 5. *Let P_{μ,σ^2} denote the probability under which $Y_1, Y_2, \cdots$ are* i.i.d. *$N(\mu,\sigma^2)$ random variables. Let $T_i = Y_1+\cdots+Y_i, \mu(t) = -c/[2\{t(1-t)\}^{\frac{1}{2}}]$, and $\sigma^2 = 1-c^2$. Then, under the conditions of Lemma 2, as $m\to\infty$*

$$P^{(m)}_{0,mI}\Bigg[\max_{n<i\le m_1}\frac{\|S_i\|}{\{i(1-\frac{i}{m})\}^{\frac{1}{2}}} < cm^{\frac{1}{2}}|S_n=\xi\Bigg] \to P_{\mu(t),\sigma^2}\Big(\max_{i\ge1} T_i \le -x\Big).$$

Proof. It can be checked via the conditional density of S_i given $S_n=\xi, S_m=0$ and $U_m = mI$, that

$$\lim_{k\to\infty}\limsup_{m\to\infty}\sum_{i=n+k+1}^{m_1} P^{(m)}_{0,mI}\Bigg[\frac{\|S_i\|}{\{i(1-\frac{i}{m})\}^{\frac{1}{2}}} \ge cm^{\frac{1}{2}}|S_n=\xi\Bigg] = 0. \tag{2.4}$$

Let $Z_1, Z_2, \ldots$ be i.i.d. $N(-c(t/\{1-t\})^{\frac{1}{2}}\alpha, I-c^2\alpha\alpha')$. By Lemma 4,

$$\lim_{m\to\infty} P^{(m)}_{0,mI}\Bigg[\max_{n<i\le n+k}\frac{\|S_i\|}{\{i(1-\frac{i}{m})\}^{\frac{1}{2}}} < cm^{\frac{1}{2}}|S_n=\xi\Bigg]$$
$$=\lim_{m\to\infty} P[(\xi+Z_1+\cdots+Z_i)'\alpha < c\{i(m-i)\}^{\frac{1}{2}} \text{ for all } 1\le i\le k]$$
$$=\lim_{m\to\infty} P\Big[(Z_1+\cdots+Z_i)'\alpha<c\Big(\{i(m-i)\}^{\frac{1}{2}}-\{n(m-n)\}^{\frac{1}{2}}\Big)-x \text{ for all } 1\le i\le k\Big]$$
$$=P_{\mu(t),\sigma^2}\Big(\max_{1\le i\le k} T_i < -x\Big).$$

This, together with (2.4), yields the lemma.

Lemma 6. *Suppose $\mu>0$. With the notation of Lemma 5,*

$$\int_0^\infty \exp(-2\mu x/\sigma^2)P_{\mu,\sigma^2}\Big(\min_{i\ge1} T_i \ge x\Big)dx = \mu\nu(2\mu/\sigma),$$

where ν is the function defined in Theorem 1.

Proof. See problem 8.13 of Siegmund (1985) for the key steps.

Proof of Theorem 1. The first term in RHS(2.1) is asymptotically negligible. Using Lemmas 2 and 5, then changing to polar coordinates, we obtain after substantial calculation

$$
\begin{aligned}
&\text{LHS}(2.2)\\
&\sim \sum_{m_0\le n<m_1} \int_{\|\xi\|\ge c\{n(m-n)\}^{\frac{1}{2}}} P_{0,mI}^{(m)}\Big[\max_{n<i\le m_1} \frac{\|S_i\|}{\{i(1-\frac{i}{m})\}^{\frac{1}{2}}} < cm^{\frac{1}{2}} | S_n=\xi\Big]\\
&\quad \cdot P_{0,mI}^{(m)}\{S_n\in d\xi\}\\
&\sim \sum_{m_0\le n<m_1} \int_{\|\xi\|\ge c\{n(m-n)\}^{\frac{1}{2}}} P_{\mu(\frac{n}{m}),1-c^2}\Big\{\max_{i\ge 1} T_i \le -x\Big\}\\
&\times\Big\{\frac{1}{2\pi n(1-n/m)}\Big\}^{\frac{d}{2}} (1-c^2)^{(m-d-3)/2}\exp\Big[-\frac{cxm}{(1-c^2)\{n(m-n)\}^{\frac{1}{2}}}\Big]d\xi_1\cdots d\xi_d\\
&=\frac{(1-c^2)^{(m-d-3)/2}}{m^{\frac{d}{2}}2^{\frac{d}{2}-1}\Gamma(\frac{d}{2})}\sum_{m_0\le n<m_1}\int_0^\infty P_{\mu(\frac{n}{m}),1-c^2}\Big(\max_{i\ge 1}T_i\le -z\Big)\Big\{\frac{m^2}{n(m-n)}\Big\}^{\frac{d}{2}}\\
&\times \exp\Big[-\frac{czm}{(1-c^2)\{n(m-n)\}^{\frac{1}{2}}}\Big]\cdot\Big[z+c\{n(m-n)\}^{\frac{1}{2}}\Big]^{d-1}dz\\
&\sim\frac{(1-c^2)^{(m-d-3)/2}m^{\frac{d}{2}-1}c^{d-1}}{2^{\frac{d}{2}-1}\Gamma(\frac{d}{2})}\sum_{m_0\le n<m_1}\Big\{\frac{m^2}{n(m-n)}\Big\}^{\frac{1}{2}}\\
&\times\int_0^\infty P_{\mu(\frac{n}{m}),1-c^2}\Big(\max_{i\ge 1}T_i\le -z\Big)\cdot\exp\Big\{-\frac{czm}{(1-c^2)\{n(m-n)\}^{\frac{1}{2}}}\Big\}dz.
\end{aligned}
$$

Using Lemma 6 and approximating the Riemann sum by an integral, we have

$$
\begin{aligned}
&\text{LHS}(2.2)\\
&\sim\frac{(1-c^2)^{(m-d-3)/2}m^{\frac{d}{2}-1}c^d}{2^{d/2}\Gamma(\frac{d}{2})}\sum_{m_0\le n<m_1}\frac{m^2}{n(m-n)}\cdot\nu\Big(\frac{cm}{\{(1-c^2)n(m-n)\}^{\frac{1}{2}}}\Big)\\
&\sim \text{RHS}(2.2),
\end{aligned}
$$

which completes the proof of Theorem 1.

3. Numerical Examples

In this section we compare the approximation given in Theorem 1 with simulations conducted by Srivastava and Worsley (1986) and with their improved Bonferroni inequality. This inequality gives an upper bound for the significance level and is quite accurate in small samples but tends to be less accurate for larger sample sizes.

As Table 1 below shows, the approximation (2.2) yields numerical results which usually are somewhat larger than the Monte Carlo estimates and are worse for larger values of d. To understand this situation better it is helpful to consider the related approximation (26) in the paper of James, James and Siegmund (1987). That approximation is concerned with a d-dimensional Brownian bridge, $W_0(t), 0 \leq t \leq 1$, and says that for any $0 < t_0 < t_1 < 1$, as b becomes infinitely large

$$P\left\{ \max_{t_0 \leq t \leq t_1} \frac{\|W_0(t)\|}{[t(1-t)]^{1/2}} > b \right\}$$
$$= \frac{b^d \exp(-b^2/2)}{2^{(d-2)/2}\Gamma(d/2)} \{2^{-1}(1 - db^{-2})\log(r) + 2b^{-2} + o(b^{-2})\}, \tag{3.1}$$

where $r = [t_1(1-t_0)]/[t_0(1-t_1)]$. In fact, under suitable conditions the left hand side of (3.1) is an approximation for the significance level of the likelihood ratio test of Section 2, but it is a good one only if the sample size is quite large. On the other hand one may readily verify that the right hand side of (3.1) provides accurate approximations for the left hand side at least up to $d = 4$ by comparison with the exact probabilities obtained numerically by DeLong (1981).

The approximation (3.1) is more precise than (2.2), in the sense that it contains second order terms which depend on the dimension, d, and reduce the first order approximation considerably when the dimension is large. It would be interesting, but appears difficult to derive a second order term for the problem of Theorem 1. See Woodroofe and Takahashi (1982) for the analogous calculation in a much simpler context.

A suitable interpretation of (3.1) suggests an *ad hoc* modification of (2.2) which appears to be more accurate in most cases. It is easy to see that the approximation (3.1) equals the dominant term in (3.1) multiplied by the factor $1-d/b^2$ and then added to (a tail approximation for) $2P\{\|W_0(t)\| \geq b[t(1-t)]^{1/2}\}$, which for $t = t_0$ or t_1 makes a boundary correction to the argument of Theorem 1. This suggests modifying the approximation of Theorem 1 through multiplication by $1 - d/b^2$ (recall that $b = cm^{\frac{1}{2}}$) and then addition of the boundary correction

$$2P_{0,mI}^{(m)}\{\|S_{m_1}\| > b[m_1(1 - m_1/m)]^{1/2}\}, \tag{3.2}$$

which is twice the first term on the right hand side of (2.1). It is not difficult to see that (3.2) equals

$$2P\{F > (m - d - 1)c^2/[d(1 - c^2)]\}, \tag{3.3}$$

where F is distributed as $F_{d,m-d-1}$.

Table 1 compares the improved Bonferroni bound of Srivastava and Worsley (1986) with the two approximations derived here and the outcomes of a Monte

Table 1. Comparison of the approximation (2.2) and the modified approximation suggested above with the improved Bonferroni bound of Srivastava and Worsley (1986).

		d		
m		2	4	6
	$\alpha = 0.01$			
20	b	3.40	3.74	4.00
	IB	0.010	0.011	0.010
	(2.2)	0.010	0.014	0.016
	(3.4)	0.009	0.011	0.011
40	b	3.63	4.12	4.51
	IB	0.015	0.013	0.010
	(2.2)	0.013	0.014	0.012
	(3.4)	0.012	0.012	0.009
	$\alpha = 0.05$			
20	b	3.08	3.52	3.83
	IB	0.057	0.068	0.054
	(2.2)	0.054	0.068	0.084
	(3.4)	0.052	0.054	0.052
40	b	3.25	3.82	4.20
	IB	0.074	0.060	0.060
	(2.2)	0.063	0.060	0.070
	(3.4)	0.058	0.049	0.051
	$\alpha = 0.10$			
20	b	2.92	3.40	3.73
	IB	0.117	0.131	0.111
	(2.2)	0.106	0.132	0.175
	(3.4)	0.099	0.103	0.115
40	b	3.07	3.65	4.04
	IB	0.139	0.125	0.125
	(2.2)	0.118	0.121	0.146
	(3.4)	0.107	0.096	0.103

Carlo experiment. The Monte Carlo results are obtained by taking the appropriate percentile from a sample of size 10,000, and are used to determine the values of b. These values, as well as the improved Bonferroni bounds, labeled IB in Table 1, are taken directly from Srivastava and Worsley (1986). In addition to the approximation (2.2) from Theorem 1, Table 1 contains the heuristic modification

$$(1 - d/b^2) \times (2.2) + (3.3). \tag{3.4}$$

This second approximation almost always improves on the first one, sometimes quite substantially. It is usually better than the improved Bonferroni bound. It also appears to be easier to evaluate. None of the three approximations is badly misleading, as, for example, direct application of (3.1) would be.

4. Confidence Sets for the Change-Point

In this section we obtain confidence sets for the change-point ρ by using the likelihood-ratio tests of the hypothesis of a change-point at ρ versus two-sided alternatives (change-point at a point different from ρ). That is, our level $1-\alpha$ confidence set for ρ will consist of those values of ρ for which we accept the hypothesis of a change-point at ρ at level α.

Let $X_1, \ldots, X_m$ be independent d-dimensional random (column) vectors having multivariate normal distributions with the same unknown covariance matrix and a change-point at some $j \in \{1, \ldots, m-1\}$. Then $X_1, \ldots, X_j$ are i.i.d. $N(\mu, \Sigma)$, and $X_{j+1}, \ldots, X_m$ are i.i.d. $N(\mu+\delta, \Sigma)$ for some unknown μ, Σ, and $\delta \neq 0$. We test the hypothesis $H_\rho : j = \rho$ versus $K_\rho : j \neq \rho$. If $f(x_1, \ldots, x_m | \rho, \mu, \delta, \Sigma)$ is the joint probability density function of $X_1, \ldots, X_m$ with the parameters ρ, μ, δ, Σ, the likelihood ratio test statistic of H_ρ versus K_ρ is given by

$$\begin{aligned} L_m &= \frac{\sup_{j \neq \rho, \mu, \delta, \Sigma} f(X_1, \ldots, X_m | j, \mu, \delta, \Sigma)}{\sup_{\mu, \delta, \Sigma} f(X_1, \ldots, X_m | \rho, \mu, \delta, \Sigma)} \\ &= \max_{j \neq \rho} \left(\frac{\det \widehat{\Sigma}_\rho}{\det \widehat{\Sigma}_j} \right)^{m/2}, \end{aligned}$$

where

$$\widehat{\Sigma}_j = \frac{1}{m} \left\{ \sum_{i=1}^{j} (X_i - \bar{X}_j)(X_i - \bar{X}_j)' + \sum_{i=j+1}^{m} (X_i - \bar{X}_j^*)(X_i - \bar{X}_j^*)' \right\},$$

with

$$\bar{X}_j = (X_1 + \cdots + X_j)/j \quad \text{and} \quad \bar{X}_j^* = (X_{j+1} + \cdots + X_m)/(m-j).$$

Since

$$\widehat{\Sigma}_j = \frac{1}{m} \left\{ \sum_{i=1}^{m} (X_i - \bar{X})(X_i - \bar{X})' - \frac{j(m-j)}{m^2} (\bar{X}_j - \bar{X}_j^*)(\bar{X}_j - \bar{X}_j^*)' \right\},$$

where $\bar{X} = (X_1 + \cdots + X_m)/m$, we have

$$L_m = \max_{j \neq \rho} \left[\frac{\det\left\{ \sum_{i=1}^{m} (X_i - \bar{X})(X_i - \bar{X})' - \frac{j(m-j)}{m} (\bar{X}_j - \bar{X}_j^*)(\bar{X}_j - \bar{X}_j^*)' \right\}}{\det\left\{ \sum_{i=1}^{m} (X_i - \bar{X})(X_i - \bar{X})' - \frac{\rho(m-\rho)}{m} (\bar{X}_\rho - \bar{X}_\rho^*)(\bar{X}_\rho - \bar{X}_\rho^*)' \right\}} \right]^{-m/2}$$

Since L_m is independent of $S_m = X_1 + \cdots + X_m$ by Basu's Theorem,

$$P_{\rho, \mu, \delta, \Sigma}(L_m \geq \gamma) = P_{\rho, \mu, \delta, \Sigma}(L_m > \gamma | S_m = 0).$$

Note that under H_ρ, the distribution of L_m depends on the parameters δ and Σ, but not on μ. By evaluating the significance level conditionally, conditioning on the remaining sufficient statistics S_ρ and $U_m = \sum_{i=1}^m X_i X_i'$, we obtain a procedure free of unknown nuisance parameters.

Given $S_m = 0$, $S_\rho = \xi$ and $U_m = \lambda$,

$$L_m = \max_{j\neq\rho}\left[\frac{\det\left\{\lambda - \frac{m}{j(m-j)}S_jS_j'\right\}}{\det\left\{\lambda - \frac{m}{\rho(m-\rho)}\xi\xi'\right\}}\right]^{-\frac{m}{2}} = \max_{j\neq\rho}\left\{\frac{1-\frac{m}{j(m-j)}S_j'\lambda^{-1}S_j}{1-\frac{m}{\rho(m-\rho)}\xi'\lambda^{-1}\xi}\right\}^{-\frac{m}{2}}.$$

A bit of algebra shows that an equivalent conditional test, given $S_m = 0$, $S_\rho = \xi$, $U_m = \lambda$, rejects when $\max_j mS_j'\lambda^{-1}S_j/\{j(m-j)\}$ is large. So, to carry out the conditional test, it is sufficient to find $c_0 = c_0(\alpha,\rho,\xi,\lambda)$ such that

$$P_\rho\{S_j'\lambda^{-1}S_j \geq c_0 j(1-j/m) \text{ for some } j\neq\rho | S_m = 0, S_\rho = \xi, U_m = \lambda\} = \alpha. \quad (4.1)$$

Approximations for this probability can be obtained by considering separately the cases where the crossing occurs before ρ and after ρ. Consider first

$$P_\rho\{S_j'\lambda^{-1}S_j \geq c_0 j(1-j/m) \text{ for some } 1\leq j<\rho | S_m = 0, S_\rho = \xi, U_m = \lambda\}.$$

If λ is positive definite, which is true of almost all possible values of U_m, let A be such that $\lambda = mAA'$. By multiplying all X_i on the left by A^{-1}, we see that this last probability equals

$$\begin{aligned} &P_{0,mI}^{(m)}\{S_j'S_j \geq c_0 j(m-j) \text{ for some } 1\leq j<\rho | S_\rho = A^{-1}\xi\} \\ &= \sum_{j=1}^{\rho-1}\int_{\|x\|^2\geq j(m-j)c_0} P_{0,mI}^{(m)}\left\{\frac{\|S_n\|^2}{[n(m-n)]} < c_0 \text{ for all } j<n<\rho | S_j = x, S_\rho = A^{-1}\xi\right\} \\ &\cdot P_{0,mI}^{(m)}\{S_j\in dx | S_\rho = A^{-1}\xi\}. \end{aligned} \quad (4.2)$$

As in Section 2, we develop approximations for the probabilities and densities in the above integrals. These approximations allow us to prove the following theorem concerning tail probabilities for the conditional test statistics, the proof of which will be given following a series of lemmas.

Theorem 2. *Suppose $\xi = m\xi_0$, $\lambda = m\lambda_0$ and $\rho = m\pi_0$, where $0<\pi_0<1$, $0<c_0<1$, $\xi_0'\lambda_0^{-1}\xi_0/\{\pi_0(1-\pi_0)\} < c_0$, and λ_0 is positive definite. Then*

$$P_\rho\{S_j'\lambda^{-1}S_j \geq c_0 j(1-j/m) \text{ for some } 1\leq j<\rho | S_m = 0, S_\rho = \xi, U_m = \lambda\}$$
$$\sim\left\{\frac{c_0\pi_0(1-\pi_0)}{\xi_0'\lambda_0^{-1}\xi_0}\right\}^{\frac{d}{2}}\left\{\frac{1-c_0}{1-\frac{\xi_0'\lambda_0^{-1}\xi_0}{\pi_0(1-\pi_0)}}\right\}^{(m-d-3)/2}\nu\left[\frac{c_0(1-\pi_0)^2+\xi_0'\lambda_0^{-1}\xi_0}{(1-\pi_0)\{(1-c_0)\xi_0'\lambda_0^{-1}\xi_0\}^{\frac{1}{2}}}\right] \quad (4.3)$$

as $m\to\infty$.

With the help of Theorem 2 an approximate conditional likelihood ratio confidence region for ρ is obtained as follows. For each trial value of ρ, set c_0 equal to the observed value of

$$\max_j (S_j - jS_m/m)'(U_m - m^{-1}S_mS_m')^{-1}(S_j - jS_m/m)/[j(1-j/m)].$$

If this maximum is attained at $j = \rho$, include ρ in the confidence set. Otherwise evaluate approximately the probability in (4.3) and the corresponding probability for the range $\rho < j < m$. If the sum of these two terms exceeds α, include the trial value of ρ in the confidence set, and otherwise exclude it. (Note that it is not necessary to find the actual value of c_0 for which (4.1) holds.) In principle one must repeat this procedure for each ρ, but in practice the number of trial values which must be considered depends on the behavior of the data. For numerical examples in a slightly simpler context see Siegmund (1988).

The conditions of Theorem 2 are assumed to hold in the following lemmas. Lemma 7 is closely related to Lemma 1.

Lemma 7. *If the matrix A satisfies $mAA' = \lambda$, then for any $1 \le j < \rho$*

$$f_{S_j}(x|S_\rho = A^{-1}\xi, S_m = 0, U_m = mI)$$
$$= \left\{\frac{\rho}{\pi j(\rho - j)}\right\}^{\frac{d}{2}} \frac{\Gamma_d(\frac{m-2}{2})}{\Gamma_d(\frac{m-3}{2})}$$
$$\cdot \frac{\left[\det\left\{mI - \frac{1}{j}xx' - \frac{1}{\rho-j}(A^{-1}\xi - x)(A^{-1}\xi - x)' - \frac{1}{m-\rho}(A^{-1}\xi)(A^{-1}\xi)'\right\}\right]^{(m-d-4)/2}}{\left[\det\left\{mI - \frac{m}{\rho(m-\rho)}(A^{-1}\xi)(A^{-1}\xi)'\right\}\right]^{(m-d-3)/2}}$$

for any x such that the argument of the determinant in the numerator is positive definite.

Proof. The joint density of S_j, S_ρ, S_m and U_m can be obtained from that of the independent S_j, $S_\rho - S_j$, $S_m - S_\rho$ and $U_m - S_jS_j'/j - (S_\rho - S_j)(S_\rho - S_j)'/(\rho - j) - (S_m - S_\rho)(S_m - S_\rho)'/(m - \rho)$, with a similar, but simpler, procedure for the joint density of S_ρ, S_m and U_m. Substitution of the appropriate normal and Wishart densities yields the result.

The proofs of Lemmas 8 and 9 below are analogous to those of Lemmas 4 and 5.

Lemma 8. *Let $x = [\{c_0 j(m-j)\}^{\frac{1}{2}} + x_0]\alpha$ where $x_0 > 0$, α is a unit vector, and $\lim_{m\to\infty} j/m = t$, $0 < t < \pi_0$. If $\lambda = mAA'$, where A is a fixed positive definite matrix, then for each $k = 1, 2, \ldots,$*

$$\mathcal{L}(X_{j+1}, \ldots, X_{j+k}|S_j = x, S_\rho = A^{-1}\xi, S_m = 0, U_m = mI)$$

converges to the law of k independent, identically distributed $N(\mu_0(t), \Sigma_{t,\alpha})$ random vectors, where

$$\mu_0(t) = (\pi_0 - t)^{-1}[A^{-1}\xi - \{c_0 t(1-t)\}^{\frac{1}{2}}\alpha]$$

and

$$\Sigma_{t,\alpha} = I - (1-t)c_0\alpha\alpha' - (\pi_0 - t)\mu_0(t)\{\mu_0(t)\}' - (1-\pi_0)^{-1}(A^{-1}\xi_0)(A^{-1}\xi_0)'.$$

Remark 1. The above result is intuitively correct, because $\mu_0(t)$ and $\Sigma_{t,\alpha}$ are the limits of the conditional mean and covariance matrix of $X_{j+i}, i = 1, 2, \ldots$. For example, the conditional expectation of X_{j+i}, $1 \le i < \rho - j$, is $(A^{-1}\xi - x)/(\rho - j)$.

Lemma 9. *Let $\alpha_0 = A^{-1}\xi_0/\|A^{-1}\xi_0\|$ and*

$$t^* = \frac{\xi_0'\lambda_0^{-1}\xi_0}{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0}, \quad \mu^* = \frac{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0}{2(1-\pi_0)(\xi_0'\lambda_0^{-1}\xi_0)^{\frac{1}{2}}}.$$

Under the conditions of Lemma 8,

$$\lim_{\alpha\to\alpha_0, t\to t^*} \lim_{m\to\infty} P_{0,mI}^{(m)}\{\|S_n\|^2 < c_0 n(m-n) \text{ for all } j < n < \rho | S_j = x, S_\rho = A^{-1}\xi\}$$
$$= P_{\mu^*, 1-c_0}\Big(\max_{n\ge 1} T_n \le -x_0\Big),$$

where P_{μ,σ^2} has the same meaning as in Lemma 5.

Remark 2. The reason for considering the special values α_0 and t^* is that conditionally, given $S_\rho = A^{-1}\xi$, $S_m = 0$ and $U_m = mI$, if the process crosses the boundary it will tend to do it near $j = mt^*$ and in the direction of α_0. This claim is made precise by the following lemmas.

Lemma 10. *Let $\Sigma_0 = I - (A^{-1}\xi_0)(A^{-1}\xi_0)'/\{\pi_0(1-\pi_0)\}$ and $f(z) = (z\alpha - j(m\pi_0)^{-1}A^{-1}\xi_0)'\Sigma_0^{-1}(z\alpha - j(m\pi_0)^{-1}A^{-1}\xi_0)$. Under the conditions of Lemma 8, as $m \to \infty$,*

$$\left[\frac{\det\left\{mI - \frac{1}{j}xx' - \frac{1}{\rho-j}(A^{-1}\xi - x)(A^{-1}\xi - x)' - \frac{1}{m-\rho}(A^{-1}\xi)(A^{-1}\xi)'\right\}}{\det\left\{mI - \frac{m}{\rho(m-\rho)}(A^{-1}\xi)(A^{-1}\xi)'\right\}}\right]^{(m-d-4)/2}$$
$$\sim \left(1 - \frac{m\pi_0}{j(\pi_0 - j/m)} f\left[\left\{c_0\frac{j}{m}\left(1-\frac{j}{m}\right)\right\}^{\frac{1}{2}}\right]\right)^{(m-d-4)/2} \cdot \exp\left\{-\frac{\pi_0 x_0}{t(\pi_0 - t)}\right. \tag{4.4}$$
$$\left.\cdot \frac{\{c_0 t(1-t)\}^{\frac{1}{2}}\alpha'\Sigma_0^{-1}\alpha - \frac{t}{\pi_0}(A^{-1}\xi_0)'\Sigma_0^{-1}\alpha}{1 - \frac{\pi_0}{t(\pi_0-t)}\left[\{c_0 t(1-t)\}^{\frac{1}{2}}\alpha - \frac{t}{\pi_0}A^{-1}\xi_0\right]'\Sigma_0^{-1}\left[\{c_0 t(1-t)\}^{\frac{1}{2}}\alpha - \frac{t}{\pi_0}A^{-1}\xi_0\right]}\right\}.$$

Proof. Since

$$\frac{\det(B - yy')}{\det(B)} = 1 - y'B^{-1}y$$

for invertible matrices B, we can set

$$y = \left\{\frac{\rho}{j(\rho - j)}\right\}^{\frac{1}{2}} \left(x - \frac{j}{\rho} A^{-1}\xi\right)$$

to obtain

$$\begin{aligned} \text{LHS}(4.4) &= \left\{\frac{\det(m\Sigma_0 - yy')}{\det(m\Sigma_0)}\right\}^{(m-d-3)/2} = \left(1 - \frac{1}{m} y'\Sigma_0^{-1} y\right)^{(m-d-3)/2} \\ &= \left(1 - \frac{m\pi_0}{j(\pi_0 - j/m)} f\left[\left\{c_0 \frac{j}{m}\left(1 - \frac{j}{m}\right)\right\}^{\frac{1}{2}}\right]\right)^{(m-d-3)/2} \\ &\cdot \left(\frac{1 - \frac{m\pi_0}{j(\pi_0 - j/m)} f\left[\{c_0 \frac{j}{m}(1 - \frac{j}{m})\}^{\frac{1}{2}} + \frac{x_0}{m}\right]}{1 - \frac{m\pi_0}{j(\pi_0 - j/m)} f\left[\{c_0 \frac{j}{m}(1 - \frac{j}{m})\}^{\frac{1}{2}}\right]}\right)^{(m-d-3)/2} \end{aligned}$$

To complete the proof, it is sufficient to show that the last factor above converges to the exponential term of (4.4). Since

$$f\left[\left\{c_0 \frac{j}{m}\left(1 - \frac{j}{m}\right)\right\}^{\frac{1}{2}} + \frac{x_0}{m}\right] - f\left[\left\{c_0 \frac{j}{m}\left(1 - \frac{j}{m}\right)\right\}^{\frac{1}{2}}\right] \sim \frac{x_0}{m} f'\left[\left\{c_0 \frac{j}{m}\left(1 - \frac{j}{m}\right)\right\}^{\frac{1}{2}}\right]$$

and

$$f'(z) = 2\left\{z\alpha'\Sigma_0^{-1}\alpha - \frac{j}{m\pi_0}(A^{-1}\xi_0)'\Sigma_0^{-1}\alpha\right\},$$

the lemma follows from the fact that $\lim_{m\to\infty}(1+\frac{a_m}{m})^m = e^a$ when $\lim_{m\to\infty} a_m = a$.

Remark 3. The exponential term in Lemma 10, evaluated at $\alpha = \alpha_0$ and $t = t^*$ (see Remark 2), equals

$$\exp\left[\frac{-x_0\{c_0(1 - \pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0\}}{(1 - c_0)(1 - \pi_0)(\xi_0'\lambda_0^{-1}\xi_0)^{\frac{1}{2}}}\right].$$

Lemma 11. *Let* $e_1 = (1, 0, \dots, 0)', 0 < t < \pi_0$ *and let* α *be a unit vector. The function*

$$\begin{aligned} g(t, \alpha) = 1 - &\frac{\pi_0}{t(\pi_0 - t)} \left[\{c_0 t(1 - t)\}^{\frac{1}{2}}\alpha - \frac{t}{\pi_0}\|A^{-1}\xi_0\| e_1\right]' \left\{I - \frac{\xi_0'\lambda_0^{-1}\xi_0}{\pi_0(1 - \pi_0)} e_1 e_1'\right\}^{-1} \\ &\cdot \left[\{c_0 t(1 - t)\}^{\frac{1}{2}}\alpha - (t/\pi_0)\|A^{-1}\xi_0\| e_1\right] \end{aligned}$$

has a unique maximum at $\alpha = e_1$ *and* $t = t^*$, *where* t^* *is defined in Lemma* 9. *Furthermore,* $g(t,\alpha)$ *is bounded away from* $g(t^*, e_1)$ *outside a neighborhood of* (t^*, e_1), *and*

$$g(t^*, e_1) = (1 - c_0)\Big/\left(1 - \frac{\xi_0' \lambda_0^{-1} \xi_0}{\pi_0(1-\pi_0)}\right).$$

Proof. For fixed t, one can use Lagrange multipliers to show that $g(t,\alpha)$ is maximized at $\alpha = e_1$. In this case,

$$g(t, e_1) = 1 - \frac{\pi_0}{t(\pi_0 - t)} \cdot \frac{\left[\{c_0 t(1-t)\}^{\frac{1}{2}} - (t/\pi_0)\|A^{-1}\xi_0\|\right]^2}{1 - (\xi_0' \lambda_0^{-1} \xi_0)/[\pi_0(1-\pi_0)]}.$$

This is maximized at $t = t^*$, with $g(t^*, e_1)$ as given above.

Lemma 12. *Let* $e_1, g(t,\alpha)$ *be as defined in Lemma* 11 *and* t^* *as in Lemma* 9. *Then*

$$\sum_{j=1}^{\rho-1} \left\{\frac{g(\frac{j}{m}, e_1)}{g(t^*, e_1)}\right\}^{(m-d-4)/2}$$
$$\sim \frac{2(1-\pi_0)^2[2m\pi c_0(1-c_0)(\xi_0'\lambda_0^{-1}\xi_0)\{c_0\pi_0(1-\pi_0) - \xi_0'\lambda_0^{-1}\xi_0\}]^{\frac{1}{2}}}{\{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0\}^2}$$

as $m \to \infty$.

Proof. By Lemma 11, one may restrict the summation to j such that j/m lies in a neighborhood of t^*. The sum may then be evaluated by using a Taylor series for $f(t) = g(t, e_1)/g(t^*, e_1)$ about $t = t^*$, since by Lemma 11 we have $f(t^*) = 1$ and $f'(t^*) = 0$. Thus,

$$\sum_{j=1}^{\rho-1} \left\{\frac{g(j/m, e_1)}{g(t^*, e_1)}\right\}^{(m-d-4)/2} \sim \sum_{|\frac{j}{m} - t^*| < \epsilon} \{f(j/m)\}^{(m-d-4)/2}$$
$$\sim \sum_{|\frac{j}{m} - t^*| < \epsilon} \left\{1 + \frac{1}{2}\left(\frac{j}{m} - t^*\right)^2 f''(t^*)\right\}^{(m-d-4)/2}$$
$$\sim m \int_{t^*-\epsilon}^{t^*+\epsilon} \left\{1 + \frac{1}{2} f''(t^*)(t - t^*)^2\right\}^{(m-d-4)/2} dt$$
$$= m^{\frac{1}{2}} \int_{-m^{\frac{1}{2}}\epsilon}^{m^{\frac{1}{2}}\epsilon} \left\{1 + \frac{x^2}{2m} f''(t^*)\right\}^{(m-d-4)/2} dx \sim 2\left[\frac{m\pi}{\{-f''(t^*)\}}\right]^{\frac{1}{2}}.$$

The lemma then follows from the fact that

$$f''(t^*) = -\frac{\{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0\}^4}{2c_0(1-c_0)(1-\pi_0)^4(\xi_0'\lambda_0^{-1}\xi_0)\{c_0\pi_0(1-\pi_0) - \xi_0'\lambda_0^{-1}\xi_0\}}.$$

Lemma 13. *Let* e_1 *and* $g(t,\alpha)$ *be as defined in Lemma 11. Assume* $\lim_{m\to\infty}(j/m) = t$, *where* $0 < t < \pi_0$. *Then*

$$\lim_{t\to t^*}\lim_{m\to\infty} m^{(d-1)/2}\int_{S^{d-1}}\left\{\frac{g(\frac{j}{m},\alpha)}{g(\frac{j}{m},e_1)}\right\}^{(m-d-4)/2} d\alpha$$
$$= \left[\frac{2\pi\{c_0\pi_0(1-\pi_0) - \xi_0'\lambda_0^{-1}\xi_0\}}{c_0(\xi_0'\lambda_0^{-1}\xi_0)}\right]^{(d-1)/2},$$

where S^{d-1} *means the unit sphere in* $\mathcal{R}^d$, *and* $d\alpha$ *means the differential of surface volume.*

Proof. By writing $\alpha = \alpha_1 e_1 + (1-\alpha_1^2)^{\frac{1}{2}}e$, where $ee_1' = 0$, we can evaluate the integrand above and show that it depends on α only through α_1. Letting $\alpha_1 = \cos\theta$, where θ is the angle between α and e_1, we can write the integrand as a function of θ; that is,

$$h(\theta) = h(\theta,t,m) = \left\{\frac{g(t,\alpha)}{g(t,e_1)}\right\}^{(m-d-4)/2} = \Bigg[1 - \frac{1}{t(\pi_0 - t)}$$
$$\times\,[2t(1-\cos\theta)\{c_0t(1-t)(\xi_0'\lambda_0^{-1}\xi_0)\}^{\frac{1}{2}} - c_0t(1-t)(\xi_0'\lambda_0^{-1}\xi_0)(1-\pi_0)^{-1}(1-\cos^2\theta)]$$
$$\times\left\{1 - \frac{\xi_0'\lambda_0^{-1}\xi_0}{\pi_0(1-\pi_0)} - \frac{\pi_0}{t(\pi_0-t)}[\{c_0t(1-t)\}^{\frac{1}{2}} - t\pi_0^{-1}(\xi_0'\lambda_0^{-1}\xi_0)^{\frac{1}{2}}]^2\right\}^{-1}\Bigg]^{(m-d-4)/2}$$

Integrating over the hypersphere corresponding to fixed θ, we obtain

$$\int_{S^{d-1}}\left\{\frac{g(j/m,\alpha)}{g(j/m,e_1)}\right\}^{(m-d-4)/2} d\alpha = \int_0^\pi h(\theta)(\sin\theta)^{(d-2)/2}\frac{2\pi^{(d-1)/2}}{\Gamma[(d-1)/2]}d\theta.$$

The conclusion of the lemma now follows by a calculus argument.

Proof of Theorem 2. By (4.2) and circular symmetry, it is sufficient to prove the theorem when $(A^{-1}\xi_0)/\|A^{-1}\xi_0\| = e_1 = (1,0,\ldots,0)'$; therefore, this is assumed below.

Let $x = [\{c_0j(m-j)\}^{\frac{1}{2}} + x_0]\alpha$, where $x_0 > 0$ and α is a unit vector. Formula

(4.2) and Lemmas 7 through 11, together with Remark 3, imply that, as $m \to \infty$,

$$\begin{aligned} \text{LHS}(4.3) \sim & \frac{\Gamma_d(\frac{m-2}{2})}{\Gamma_d(\frac{m-3}{2})} m^{-\frac{d}{2}} \left[\det\left\{I - \frac{(A^{-1}\xi_0)(A^{-1}\xi_0)'}{\pi_0(1-\pi_0)}\right\}\right]^{-\frac{1}{2}} \\ & \cdot \sum_{j=1}^{\rho-1} \left(\left\{\frac{\rho}{\pi j(\rho-j)}\right\}^{\frac{d}{2}} \int_{\|x\|^2 \geq c_0 j(m-j)} \left\{g\left(\frac{j}{m}, \alpha\right)\right\}^{(m-d-4)/2} \right. \\ & \left. \cdot P_{\mu^*,1-c_0}\left(\max_{n\geq 1} T_n \leq -x_0\right) \exp\left[-\frac{x_0\{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0\}}{(1-c_0)(1-\pi_0)(\xi_0'\lambda_0^{-1}\xi_0)^{\frac{1}{2}}}\right] dx_1 \cdots dx_d \right). \end{aligned} \tag{4.5}$$

The ratio of the gamma functions is asymptotically equivalent to $(m/2)^{d/2}$, and

$$\det\left\{I - \frac{(A^{-1}\xi_0)(A^{-1}\xi_0)'}{\pi_0(1-\pi_0)}\right\} = 1 - \frac{\xi_0'\lambda_0^{-1}\xi_0}{\pi_0(1-\pi_0)}.$$

Now change to polar coordinates, to obtain

$$\begin{aligned} \text{LHS}(4.3) \sim & \frac{1}{2^{d/2}} \left\{1 - \frac{\xi_0'\lambda_0^{-1}\xi_0}{\pi_0(1-\pi_0)}\right\}^{-\frac{1}{2}} \\ & \cdot \sum_{j=1}^{\rho-1} \left\{\frac{\rho}{\pi j(\rho-j)}\right\}^{\frac{d}{2}} \int_{S^{d-1}} \int_0^\infty \left\{g\left(\frac{j}{m}, \alpha\right)\right\}^{(m-d-4)/2} P_{\mu^*,1-c_0}\left(\max_{n\geq 1} T_n \leq -x_0\right) \\ & \cdot \exp\left[-\frac{x_0\{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0\}}{(1-c_0)(1-\pi_0)(\xi_0'\lambda_0^{-1}\xi_0)^{\frac{1}{2}}}\right] [\{c_0 j(m-j)\}^{\frac{1}{2}} + x_0]^{d-1} dx_0 d\alpha \end{aligned}$$

where the integration over the sphere is as in Lemma 11. By Lemma 11, we can substitute for $[\{c_0 j(m-j)\}^{\frac{1}{2}} + x_0]^{d-1}$ by $[m\{c_0 t^*(1-t^*)\}^{\frac{1}{2}}]^{d-1}$ and $[\rho/\{\pi j(\rho - j)\}]^{\frac{d}{2}}$ by $[\pi_0/\{\pi m t^*(\pi_0 - t^*)\}]^{\frac{d}{2}}$ in the asymptotic expression. By Lemma 6,

$$\begin{aligned} & \int_0^\infty P_{\mu^*,1-c_0}\left(\max_{n\geq 1} T_n \leq -x_0\right) \exp\left[\frac{-x_0\{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0\}}{(1-c_0)(1-\pi_0)(\xi_0'\lambda_0^{-1}\xi_0)^{\frac{1}{2}}}\right] dx_0 \\ & = \mu^*\nu\left\{\frac{2\mu^*}{(1-c_0)^{\frac{1}{2}}}\right\} = \frac{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0}{2(1-\pi_0)(\xi_0'\lambda_0^{-1}\xi_0)^{\frac{1}{2}}} \nu\left\{\frac{c_0(1-\pi_0)^2 + \xi_0'\lambda_0^{-1}\xi_0}{(1-c_0)^{\frac{1}{2}}(1-\pi_0)(\xi_0^{-1}\lambda_0^{-1}\xi_0)^{\frac{1}{2}}}\right\}, \end{aligned}$$

which is independent of α. Therefore, we have

$$\begin{aligned} \text{LHS}(4.3) \sim & \left\{\frac{m\pi_0 c_0(1-t^*)}{2\pi(\pi_0 - t^*)}\right\}^{\frac{d}{2}} \left[m c_0 t^*(t - t^*)\left\{1 - \frac{\xi_0'\lambda_0^{-1}\xi_0}{\pi_0(1-\pi_0)}\right\}\right]^{-\frac{1}{2}} \\ & \cdot \mu^*\nu\left\{\frac{2\mu^*}{(1-c_0)^{\frac{1}{2}}}\right\} \sum_{j=1}^{\rho-1} \int_{S^{d-1}} \left\{g\left(\frac{j}{m}, \alpha\right)\right\}^{(m-d-4)/2} d\alpha. \end{aligned} \tag{4.6}$$

The sum of integrals in (4.6) can be handled by writing

$$\sum_{j=1}^{\rho-1}\int_{S^{d-1}}\left\{g\left(\frac{j}{m},\alpha\right)\right\}^{(m-d-4)/2}d\alpha=\{g(t^*,e_1)\}^{(m-d-4)/2}$$

$$\cdot\left[\sum_{j=1}^{\rho-1}\left\{\frac{g(\frac{j}{m},e_1)}{g(t^*,e_1)}\right\}^{(m-d-4)/2}\right]\cdot\int_{S^{d-1}}\left\{\frac{g(\frac{j}{m},\alpha)}{g(\frac{j}{m},e_1)}\right\}^{(m-d-4)/2}d\alpha.$$

The proof of Theorem 2 is then completed by using Lemmas 12 and 13 and evaluating the constants via Lemma 11 and the definition of t^*.

5. Joint Confidence Regions

This section is concerned with joint confidence regions for the change-point ρ and $\delta=\mu_2-\mu_1$. The case of known Σ was discussed by Siegmund (1988). The present case is substantially more complicated. See also Kim and Siegmund (1989) and Knowles, Siegmund and Zhang (1991). It is convenient to introduce the notation

$$S=\sum_{i=1}^{m}(X_i-\bar{X})(X_i-\bar{X})', \tag{5.1}$$

$$W_j=[jS_m/m-S_j]/[j(1-j/m)]^{1/2}, \tag{5.2}$$

$$\Lambda_j=[\max_i W_i'S^{-1}W_i-W_j'S^{-1}W_j]/[1-W_j'S^{-1}W_j], \tag{5.3}$$

$$W_j(\delta)=[jS_m/m-S_j-j(1-j/m)\delta]/[j(1-j/m)]^{1/2}, \tag{5.4}$$

and

$$\Lambda_j(\delta)=W_j'(\delta)[S-W_jW_j'+W_j(\delta)W_j'(\delta)]^{-1}W_j(\delta). \tag{5.5}$$

The log likelihood ratio statistic for testing the specific values (ρ,δ) against a general alternative is proportional to

$$-\log(1-\Lambda_\rho)-\log(1-\Lambda_\rho(\delta)). \tag{5.6}$$

Large values of the first term provide evidence against the hypothesized value of ρ; and given a specific ρ, large values of the second term provide evidence against the hypothesized value of δ. The set of values not rejected by an α level test are a $(1-\alpha)$ confidence region. The significance level will not depend on the unknown nuisance parameter Σ provided we evaluate it conditionally given the sufficient statistic when (ρ,δ) is known, to wit, S_m and

$$V_\rho(\delta)=S-W_\rho W_\rho'+W_\rho(\delta)W_\rho'(\delta). \tag{5.7}$$

Since (5.6) is invariant with respect to changes in location, we can condition on any convenient value of S_m, e.g. $S_m = 0$.

Assuming the true change-point is ρ, we find, from standard multivariate analysis and Basu's Theorem (Lehmann (1986), Theorem 5.2), that (5.5) at $j = \rho$ and (5.7) are stochastically independent and hence

$$P_{\rho,\delta}\{\Lambda_\rho(\delta) > c | V_\rho(\delta), S_m = 0\} = P\{B_{d/2,(m-d-1)/2} > c\}, \qquad (5.8)$$

where $B_{d/2,(m-d-1)/2}$ has a Beta distribution with the indicated parameters.

Simple algebra shows that the statistic in (5.6) exceeds b if and only if $\Lambda_\rho + \Lambda_\rho(\delta) - \Lambda_\rho\Lambda_\rho(\delta)$ exceeds $c = 1 - \exp(-b)$. Hence we seek to evaluate

$$P_{\rho,\delta}\{\Lambda_\rho + \Lambda_\rho(\delta) - \Lambda_\rho\Lambda_\rho(\delta) > c | V_\rho(\delta), S_m = 0\} = P_{\rho,\delta}\{\Lambda_\rho(\delta) > c | V_\rho(\delta), S_m = 0\}$$
$$+ E_{\rho,\delta}\Big[P_\rho\Big\{\Lambda_\rho > \frac{c - \Lambda_\rho(\delta)}{1 - \Lambda_\rho(\delta)} | V_\rho(\delta), S_m = 0, S_\rho\Big\};\ \Lambda_\rho(\delta) < c | V_\rho(\delta), S_m = 0\Big]. \quad (5.9)$$

The first term on the right hand side of (5.9) is given in (5.8). The conditional probability in the second term can be evaluated approximately by the results of Theorem 2, and the resulting expression integrated numerically to obtain an approximation to the desired probability.

To be more specific, by (5.1)–(5.5) the conditional probability in (5.9) can be rewritten

$$P_\rho\Big\{\max_j W_j' S^{-1} W_j > W_\rho' S^{-1} W_\rho + (1 - W_\rho' S^{-1} W_\rho)\Big[\frac{c - \Lambda_\rho(\delta)}{1 - \Lambda_\rho(\delta)}\Big] | S_\rho, S, S_m = 0\Big\},$$

which is in the form addressed by Theorem 2 if we put

$$c_0 = W_\rho' S^{-1} W_\rho + (1 - W_\rho' S^{-1} W_\rho)\Big[\frac{c - \Lambda_\rho(\delta)}{1 - \Lambda_\rho(\delta)}\Big].$$

If $S_\rho = m\xi_0$, $S_m = 0$, $S = m\lambda_0$ and $\rho = m\pi_0$, then $W_\rho' S^{-1} W_\rho = \xi_0' \lambda_0^{-1} \xi_0 / \pi_0(1 - \pi_0)$ and hence the factor

$$\{(1 - c_0)/(1 - \xi_0' \lambda_0^{-1} \xi_0 / [\pi_0(1 - \pi_0)]\}^{(m-d-3)/2}$$

appearing on the right hand side of (4.3) equals

$$[(1 - c)/(1 - \Lambda_\rho(\delta))]^{(m-d-3)/2}. \qquad (5.10)$$

To take the expectation in (5.9), we must integrate with respect to

$$P_{\rho,\delta}\{\Lambda_\rho(\delta) \in dy | V_\rho(\delta), S_m = 0\} P_{\rho,\delta}\{W_\rho(\delta) \in dw | \Lambda_\rho(\delta) = y, V_\rho(\delta), S_m = 0\}.$$

By (5.8) the conditional probability density function of $\Lambda_\rho(\delta)$ equals

$$[B(d/2,(m-d-1)/2)]^{-1}y^{(d-2)/2}(1-y)^{(m-d-3)/2}, \tag{5.11}$$

the final factor of which cancels with the denominator of (5.10).

It follows from standard arguments that given $V_\rho(\delta) = v$ and $S_m = 0$, the conditional distribution of $U = v^{-1/2}W_\rho(\delta)$ is that of the first d coordinates of a point uniformly distributed on the unit sphere S^{m-2} in $m-1$ dimensional Euclidean space. Since $\|U\|^2 = W_\rho(\delta)'v^{-1}W_\rho(\delta) = \Lambda_\rho(\delta)$, the conditional distribution of $U/\|U\|$ given $V_\rho(\delta) = v$, $S_m = 0$ and $\Lambda_\rho(\delta)$ is uniform on S^{d-1}. Hence unconditioning (5.9) involves d-dimensional integration which can be expressed in terms of a product measure involving (5.11) and surface measure on S^{d-1}. By virtue of the cancellation between (5.10) and (5.11) the result is the product of a numerical constant depending on m, $(1-c)^{(m-d-3)/2}$, and a d-dimensional integral which does not involve the sample size m. We omit the details. For values of d up to about $d = 4$, the integral arising in this argument can be easily evaluated numerically, but for larger values of d it would be useful to have a comparatively simple approximation.

With the help of the preceding approximation one can evaluate (5.9) approximately and hence determine an approximate joint confidence region for ρ and δ. As in the case of a confidence region for ρ alone, numerical implementation involves selection of trial values and evaluation of (5.9) with c set equal to the observed value of $\Lambda_\rho + \Lambda_\rho(\delta) - \Lambda_\rho\Lambda_\rho(\delta)$. The confidence region consists of all pairs of parameter values yielding a probability greater than α. For efficient selection of trial values it is useful first to compute a confidence region for ρ and confidence regions for δ at a few values of ρ, assumed known. A numerical example in a simpler context is given by Siegmund (1988).

Acknowledgement

This research was partially supported by AFOSR Grant 89-0355, National Security Agency Grant MDA 904-89-H-2040, and NSF Grant DMS 89-0574.

References

DeLong, D. M. (1981). Crossing probabilities for a square root boundary by a Bessel process. *Comm. Statist. Theory Methods* **10**, 2197-2213.

James, B., James, K. L. and Siegmund, D. (1987). Tests for a change-point. *Biometrika* **74**, 71-83.

James, B., James, K. L. and Siegmund, D. (1988). Conditional boundary crossing probabilities with applications to change-point problems. *Ann. Probab.* **16**, 825-839.

Kim, H.-J. and Siegmund, D. (1989). The likelihood ratio test for a change-point in simple linear regression. *Biometrika* **76**, 409-423.

Knowles, M., Siegmund, D. and Zhang, H. (1991). Confidence regions in semilinear regression. *Biometrica* **78**, 15-31.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd edition. John Wiley, New York.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals.* Springer-Verlag, New York-Heidelberg-Berlin.

Siegmund, D. (1988). Confidence sets in change-point problems. *Internat. Statist. Rev.* **56**, 31-48.

Srivastava, M. S. and Worsley, K. J. (1986). Likelihood ratio tests for a change in the multivariate normal mean. *J. Amer. Statist. Assoc.* **81**, 199-204.

Woodroofe, M. (1976). Frequentist properties of Bayesian sequential tests. *Biometrika* **63**, 101-110.

Woodroofe, M. (1978). Large deviations of the likelihood ratio statistic with applications to sequential testing. *Ann. Statist.* **6**, 72-84.

Woodroofe, M. and Takahashi, H. (1982). Asymptotic expansions for the error probabilities of some repeated significance tests. *Ann. Statist.* **10**, 895-908.

Worsley, K. J. (1983). The power of likelihood ratio and cumulative sum tests for a change in a binomial probability. *Biometrika* **70**, 455-464.

Worsley, K. J. (1986). Confidence regions and tests for a change-point in a sequence of exponential family random variables. *Biometrika* **73**, 91-104.

Department of Mathematics & Statistics, University of Minnesota, Duluth, MN 55812-2496, U.S.A.

Department of Statistics, University of Minnesota, Duluth, MN 55812-2496, U.S.A.

Department of Statistics, Stanford University, Stanford, CA 94305-4065, U.S.A.

(Received January 1991; accepted July 1991)