The Annals of Statistics
1977, Vol. 5, No. 6, 1212-1220

A ROBUSTNESS PROPERTY OF THE TESTS FOR SERIAL CORRELATION

BY TAKEAKI KARIYA
Hitotsubashi University

This paper shows that the UMP or UMPU tests for serial correlation, derived under the assumption of a normal distribution, are quite robust against departure from normality. In fact, the tests are still UMP or UMPU in much broader classes of distributions and the null distributions remain unchanged under these classes. The results will be applied to a linear model.

1. Introduction. Tests for serial correlation are usually developed under the assumption of a normal distribution. A typical formal treatment for an exact or approximated model may be illustrated as follows:

$$X \sim N_n(0, \gamma\Sigma(\lambda)), \quad \gamma > 0, \tag{1.1}$$

$$\Sigma(\lambda)^{-1} = I_n + \lambda A, \quad \text{and} \quad \lambda \in \Lambda \equiv \{\lambda \in R \mid \Sigma(\lambda)^{-1} > 0\}, \tag{1.2}$$

where means are subtracted (by invariance). Here A is an $n \times n$ known matrix and $\Sigma(\lambda)^{-1} > 0$ denotes the positive definiteness of $\Sigma(\lambda)^{-1}$. (See examples in Section 4 and the papers [3], [10], [14].) As is well known in this model, for testing the hypothesis H: $\lambda = 0$ versus the alternative K: $\lambda > 0$, the test with c.r. (critical region) $T \equiv X'AX/X'X < c$ is a UMP (uniformly most powerful) test, and for testing H: $\lambda = 0$ versus K: $\lambda \neq 0$, the test with c.r. $T < c_1$ or $T > c_2$ is a UMPU (UMP unbiased) test (Anderson [3]). Kariya and Eaton [11] showed that the UMP test for the one-sided alternative is UMP in a much broader class of distributions and that the null distribution under any member of the class is the same as that under normal distribution (1.1) with $\gamma = 1$ and $\lambda = 0$. However, for the two-sided problem the authors failed to prove the UMPU property in the class.

In this paper it is shown that the UMPU test under normality is UMPU in a much broader class of distributions although this enlarged class remains smaller than the class treated in [11]. The null distribution is also the same as that under normality so that the existing tables are utilized to determine the critical points. The results will be applied through invariance to a linear model $y = X\beta + u$ under the assumption that the column space of X ($n \times k$) is spanned by some k latent vectors of the matrix A in (1.2). This assumption is needed even under normality to establish the UMP invariant (UMPI) or the UMPU invariant (UMPUI) properties in the linear model. Consequently the tests derived under

Received July 1975; revised July 1976/March 1977.
AMS 1970 *subject classifications*. Primary 62G10; Secondary 62F05, 62G35, 62J05.
Key words and phrases. Robustness, UMP test, UMPU test, serial correlation, invariance, completeness, Durbin-Watson test, Anderson-Anderson test, von Neumann test, linear model.

the normal model are quite robust against departure from normality in the following sense: the null distributions remain unchanged and the UMP or UMPU properties are still guaranteed with or without invariance. Such tests as the Durbin–Watson test, the Anderson–Anderson test, etc., are shown as examples at the close of the paper.

As general references, the reader is referred to Kadiyala [10] and Press [14].

2. The problem and a result in [11]. To state the problem in this paper, we define three classes of pdf's (probability density functions) with respect to the Lebesgue measure on Euclidean n-space R^n. Let $\mathscr{F}$ be the class of all pdf's on R^n and for an $n \times n$ matrix $\Sigma > 0$, let

$$\mathscr{F}_0(\Sigma) = \{f \in \mathscr{F} \mid f(x) = |\Sigma|^{-\frac{1}{2}} q(x'\Sigma^{-1}x),\ q \text{ is a function on } [0, \infty)\}, \tag{2.1}$$

$$\mathscr{F}_1(\Sigma) = \{f \in \mathscr{F} \mid f(x) = |\Sigma|^{-\frac{1}{2}} q(x'\Sigma^{-1}x),\ q \text{ is a nonincreasing function on } [0, \infty)\}, \tag{2.2}$$

and

$$\mathscr{F}_2(\Sigma) = \{f \in \mathscr{F} \mid f(x) = |\Sigma|^{-\frac{1}{2}} q(x'\Sigma^{-1}x),\ q \text{ is a nonincreasing and convex function on } [0, \infty)\}. \tag{2.3}$$

Clearly $\mathscr{F}_2(\Sigma) \subset \mathscr{F}_1(\Sigma) \subset \mathscr{F}_0(\Sigma)$. If $f(x) = |\Sigma|^{-\frac{1}{2}} q(x'\Sigma^{-1}x)$ belongs to $\mathscr{F}_2(\Sigma)$, then

$$g(x) = \int_0^\infty a^{-n/2} |\Sigma|^{-\frac{1}{2}} q(x'\Sigma^{-1}x/a)\, dG(a) \tag{2.4}$$

also belongs to $\mathscr{F}_2(\Sigma)$, where G is a distribution function on $(0, \infty)$. Hence $\mathscr{F}_2(\Sigma)$ contains the contaminated normal distribution, the multivariate t-distribution (multivariate Cauchy distribution) etc., as well as $N(0, \Sigma)$. It also contains certain distributions with bounded supports.

Now suppose that $X \in R^n$ is a random vector with a pdf h. Consider the problem of testing $H_0\colon h \in \mathscr{F}_0(\gamma I_n)$, $\gamma > 0$ versus the alternatives

$$K_1\colon h \in \mathscr{F}_1(\gamma\Sigma(\lambda)), \quad \gamma > 0, \quad \lambda > 0, \tag{2.5}$$

and

$$K_2\colon h \in \mathscr{F}_2(\gamma\Sigma(\lambda)), \quad \gamma > 0, \quad \lambda \neq 0, \tag{2.6}$$

where $\Sigma(\lambda)^{-1}$ is given in (1.2). Here it is noted that the domain Λ of λ in (1.2) is an open interval including $\lambda = 0$. The problem of testing H_0 versus K_1 has been treated in [11] and the result is summarized as

THEOREM 1. *For testing H_0 versus K_1 the test which rejects H_0 for small values of*

$$T = X'AX/X'X \tag{2.7}$$

is a UMP *test and the null distribution of T is the same as that under $N(0, I_n)$.*

This paper treats the problem of testing H_0 versus K_2, and shows that the test with c.r. $T < c_1$ or $T > c_2$ is UMPU. Since the test with c.r. $T < c$ is UMP for K_1 as in Theorem 1, one may conjecture that the test with c.r. $T < c_1$ or $T > c_2$

is also UMPU for K_1 with $\lambda \neq 0$ instead of $\lambda > 0$. But the author has been unable to show it. In Section 4, the results are applied to a linear model through invariance.

Finally we remark that $\Sigma(\lambda)^{-1}$ in (1.2) can be replaced by a matrix of the form $\Phi(\lambda)^{-1} = S + \lambda A$ where $S > 0$, in which case H_0 is replaced by H_0': $h \in F_0(\gamma S^{-1})$. Then the problem is to show the UMPU property of the test $X'AX/X'SX < c_1$ or $X'AX/X'SX > c_2$. But this problem is clearly reduced to the above problem.

3. The main results. Let $\mathscr{D}_\alpha$ be the class of level α test functions which are unbiased for the two-sided problem stated in Section 2. Let $W = X'X = ||X||^2$.

LEMMA 1. *The pair* (T, W) *is a complete sufficient statistic for the family* $\{\mathscr{F}_2(\gamma\Sigma(\lambda)), \gamma > 0, \lambda \in \Lambda\}$. *Furthermore,* W *is a complete sufficient statistic for* $\{\mathscr{F}_2(\gamma\Sigma(0)) = \mathscr{F}_2(\gamma I_n), \gamma > 0\}$, *and so for* $\{\mathscr{F}_0(\gamma I_n), \gamma > 0\}$.

PROOF. Both sufficiency assertions follow from the factorization theorem. Since $(V, W) = (X'AX, X'X)$ is complete for the normal family $\{N(0, \gamma\Sigma(\lambda)), \gamma > 0, \lambda \in \Lambda\}$ and since $\mathscr{F}_2(\gamma\Sigma(\lambda))$ contains $N(0, \gamma\Sigma(\lambda))$, (V, W) is complete for $\{\mathscr{F}_2(\gamma\Sigma(\lambda)), \gamma > 0, \lambda \in \Lambda\}$ and so is $(T, W) = (V/W, W)$. The completeness of W follows by similar consideration.

From this lemma, it is sufficient to consider test functions in $\mathscr{D}_\alpha$, based on (T, W) only. Let $\mathscr{O}(n)$ denote the set of $n \times n$ orthogonal matrices and let $d_1 \leqq d_2 \leqq \cdots \leqq d_n$ be the latent roots of A. Without loss of generality, $d_1 \neq d_n$ is assumed.

LEMMA 2. *Under* H_0, i.e., $h(x) = \gamma^{-n/2}q(x'x/\gamma)$, *$T$ has a* pdf, *say* $f_1(t)$, *on* $[d_1, d_n]$ *which does not depend on q and γ. Also under* H_0, *the* pdf *of W is given by*

$$f_2(w) = c(n)\gamma^{-n/2}q(w/\gamma)w^{(n/2)-1}, \qquad 0 < w < \infty \tag{3.1}$$

where $c(n) = [\Gamma(\frac{1}{2})]^n/\Gamma(n/2)$. *Further, T and W are independent under* H_0.

PROOF. Under H_0, $X/||X||$ has a uniform distribution on $\{x \in R^n \mid ||x|| = 1\}$, so does $Z/||Z||$ where $Z = QX$ with $Q \in \mathscr{O}(n)$. Further, $Z_i^2/||Z||^2$'s $(i = 1, \cdots, n)$ jointly have a Dirichelet distribution. Since T is expressed as $T = \sum_{i=1}^n d_i Z_i^2/||Z||^2$ for some $Q \in \mathscr{O}(n)$, the distribution of T does not depend on q and γ. The pdf of W can be derived by changing to polar coordinates in R^n. Since the distribution of T does not depend on a particular distribution in H_0, the independence of T and W under H_0 follows from a result due to Basu [4]. This completes the proof.

THEOREM 2. *If X has a* pdf *of the form*

$$h(x) = \gamma^{-n/2}|I + \lambda A|^{\frac{1}{2}}q([x'x + \lambda x'Ax]/\gamma) \tag{3.2}$$

(*that is,* $h \in \mathscr{F}_0(\gamma\Sigma(\lambda))$), *then the joint* pdf *of T and W with respect to the Lebesgue measure is* $f(t, w : \gamma, \lambda) = g(t, w : \gamma, \lambda)f_1(t)$, *where* $f_1(t)$ *is given in Lemma 2 and*

$$g(t, w : \gamma, \lambda) = c(n)\gamma^{-n/2}|I + \lambda A|^{\frac{1}{2}}q([w + \lambda tw]/\lambda)w^{n/2-1}. \tag{3.3}$$

PROOF. Choose $P \in \mathscr{O}(n)$ such that $P'AP = \text{diag}\{d_2, d_3, \cdots, d_{n-1}, d_1, d_n\} = D =$ diagonal matrix with diagonal elements $d_2, d_3, \cdots, d_{n-1}, d_1, d_n$ in this order. Let $Z = PX$. Then the pdf of Z is $\gamma^{-n/2}|I + \lambda A|^{\frac{1}{2}} q([z'z + \lambda z'Dz]/\gamma)$. Changing $z = (z_1, \cdots, z_n)'$ to the polar coordinates $z_i = w^{\frac{1}{2}} s_i(\theta)$ $(i = 1, \cdots, n)$ yields the pdf of $(W, \theta_1, \cdots, \theta_{n-1})$

$$\gamma^{-n/2}|I + \lambda A|^{\frac{1}{2}} q([w + \lambda t(\theta) w]/\gamma) w^{n/2-1} J(\theta) ,$$

where $s_i(\theta) = \cos\theta_1 \cos\theta_2 \cdots \cos\theta_{i-1} \sin\theta_i$ $(i = 1, \cdots, n-1)$, $s_n(\theta) = \cos\theta_1 \cos\theta_2 \cdots \cos\theta_{n-1}$, $t(\theta) = d_2 s_1(\theta)^2 + d_3 s_2(\theta)^2 + \cdots + d_{n-1} s_{n-2}(\theta)^2 + d_1 s_{n-1}(\theta)^2 + d_n s_n(\theta)^2$ and $J(\theta) = \cos^{n-2}\theta \cos^{n-3}\theta_2 \cdots \cos\theta_{n-1}$ $(-\pi/2 < \theta_i < \pi/2, i = 1, \cdots, n-2; -\pi < \theta_{n-1} < \pi)$. Further since $T = t(\theta)$, we change $(W, \theta_1, \cdots, \theta_{n-1})$ to $(W, \theta_1, \cdots, \theta_{n-2}, T)$ for each region of $\theta_{n-1} \in (-\pi, \pi/2)$, $\theta_{n-1} \in [-\pi/2, 0)$, $\theta_{n-1} \in [0, \pi/2)$ and $\theta_{n-1} \in [\pi/2, \pi)$. Then the pdf of $(W, \theta_1, \cdots, \theta_{n-1}, T)$ is

$$\gamma^{-n/2}|I + \lambda A|^{\frac{1}{2}} q([w + \lambda t w]/\gamma) w^{n/2-1} R(t, \theta_1, \cdots, \theta_{n-2})$$

where $R(t, \theta_1, \cdots, \theta_{n-2})$ does not depend on w. Thus the pdf of (T, W) is given by $g(t, w : \gamma, \lambda) r(t)$, with $r(t) = \int \cdots \int R(t, \theta_1, \cdots, \theta_{n-2})\, d\theta_1 \cdots d\theta_{n-2}$. Set $\lambda = 0$ to identify $r(t)$ with $f_1(t)$ in Lemma 2. This completes the proof.

We now suppose $h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$. Since h is of the form (3.2), the pdf of (T, W) is obtained from Theorem 2. Define

$$k(t : w, \gamma, \lambda) = \frac{g(t, w : \gamma, \lambda) f_1(t)}{\int g(t, w : \gamma, \lambda) f_1(t)\, dt} , \tag{3.4}$$

which is the conditional pdf of T given W. Let E_0 denote expectation under H_0.

LEMMA 3. *If ϕ is a test function in $\mathscr{D}_\alpha$, then*

$$E_0[\phi(T, W) \mid W] = \alpha \quad \text{a.e.} \quad (W) \tag{3.5}$$

$$E_0[T\phi(T, W) \mid W] = \alpha E_0 T \quad \text{a.e.} \quad (W) , \tag{3.6}$$

where $E_0[\cdot \mid W]$ denotes conditional expectation given W under H_0.

PROOF. Since $\phi \in \mathscr{D}_\alpha$, $E_h \phi \geqq \alpha$ for all $h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$ with $\lambda \neq 0$ and $E_h \phi \leqq \alpha$ for all $h \in \mathscr{F}_0(\gamma I)$. Hence $E_h \phi = \alpha$ for all $h \in \mathscr{F}_2(\gamma I) = \mathscr{F}_2(\gamma\Sigma(0))$ by a simple continuity argument. But for $h \in \mathscr{F}_2(\gamma I)$, $E_h \phi(T, W) = \int\int \phi(t, w) f_1(t) f_2(w)\, dt\, dw = E_0 E_0[\phi(T, W) \mid W]$. Thus $E_0 E_0[\phi(T, W) - \alpha \mid W] = 0$. Since W is complete for $\mathscr{F}_2(\gamma I)$, $\gamma > 0$, (3.5) follows. For (3.6), we first assume $X \sim N(0, \gamma\Sigma(\lambda))$, then by arguing as in Lehmann [13] Chapter 4, we obtain $E[T\phi(T, W) \mid W] = \alpha E[T \mid W]$ when $\lambda = 0$. Since the distribution of T under $N(0, \gamma I)$ is the same as that under $h \in \mathscr{F}_0(\gamma I)$ as shown in Lemma 2 and since T is independent of W under H_0, (3.6) follows.

Let $\tilde{\mathscr{D}}_\alpha$ be the set of level α test functions satisfying (3.5) and (3.6) where $0 < \alpha < 1$. Define the test function ϕ_0 by

$$\begin{aligned} \phi_0(t) &= 1 \quad \text{if} \quad t < c_1 \quad \text{or} \quad t > c_2 , \qquad \text{and} \\ \phi_0(t) &= 0 \quad \text{otherwise.} \end{aligned} \tag{3.7}$$

THEOREM 3. *If* $\phi \in \tilde{\mathscr{D}}_\alpha$ $(0 < \alpha < 1)$, *then*

$$E_h \phi_0 \geqq E_h \phi \quad \text{for all} \quad h \in \mathscr{F}_2(\gamma\Sigma(\lambda))\,, \quad \gamma > 0\,, \quad \lambda \neq 0\,. \tag{3.8}$$

Further, for any $h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$ *such that the* pdf *in* (3.4) *gives no mass to the set* $\{t \mid \phi_0(t, w) = \phi(t, w)\}$, *or simply* $\phi_0 \neq \phi$ a.e. $(T \mid W)$, *then the inequality in* (3.8) *is strict unless* $E_0 T = 0$. *Especially when the support of* $h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$ *is* R^n, *it holds strictly, provided* $\phi_0 \neq \phi$ a.e. *and* $E_0 T \neq 0$.

PROOF. Fix $h(x) = \gamma^{-n/2}|I + \lambda A|^{\frac{1}{2}} q([x'x + \lambda x'Ax]/\gamma)$ where q is convex and nonincreasing on $[0, \infty)$, so $h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$. For a fixed value W, consider the problem of testing $H: \lambda = 0$ versus $K: \lambda = \lambda^* \neq 0$, where γ is arbitrarily fixed. Applying the generalized Neyman–Pearson Lemma ([13] page 83), the supremum of $E_K[\phi(T, W) \mid W]$ over the set $\tilde{\mathscr{D}}_\alpha$ is achieved by test functions of the form: $\phi_1 = 1$ if $k(t: w, \gamma, \lambda^*) > a_1 f_1(t) + a_2 t f_1(t)$; $\phi_1 = 0$ otherwise, where $k(t: w, \gamma, \lambda)$ is given in (3.4) and a_1 and a_2 are so chosen for ϕ_1 to satisfy (3.5) and (3.6). Since q is convex, $\phi_1 = 1$ either if $t < c$ or $t > c'$, or if $t < c$, or if $t > c'$. Here c and c' can be free from the fixed W since under H_0 the conditional pdf of T given W does not depend on W. We shall show that the last two c.r.'s $t < c$ $t > c'$ cannot satisfy (3.5) and (3.6). For example, suppose the c.r. $t < c$ satisfies these conditions. Then (3.6) together with (3.5) is $\int_{d_1}^{c} t f_1(t)\,dt = R(c) E_0 T$ where $R(c) = \int_{d_1}^{c} f_1(t)\,dt$. Since $\int_{d_1}^{c} t f_1(t)\,dt = cR(c) - \int_{d_1}^{c} R(t)\,dt$,

$$E_0 T = c - [R(c)]^{-1} \int_{d_1}^{c} R(t)\,dt\,. \tag{3.9}$$

Let $H(c)$ be the right-hand side of (3.9). Then $H(c)$ is a strictly increasing and continuous function of c $(d_1 \leqq c \leqq d_n)$ with $H'(c) > 0$. Hence (3.9) is impossible unless $c = d_n$ since $H(d_n) = E_0 T$. But $c = d_n$ contradicts $\alpha < 1$. Similarly the c.r. $t > c'$ cannot satisfy (3.5) and (3.6). Therefore ϕ_0 in (3.7) satisfies

$$E_K[\phi_0(T) \mid W] \geqq E_K[\phi(T, W) \mid W] \quad \text{a.e.} \quad (W) \qquad \text{for all} \quad \phi \in \tilde{\mathscr{D}}_\alpha\,. \tag{3.10}$$

Since ϕ_0 did not depend on the particular $h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$, (3.8) holds. Further, by applying the necessary part of the generalized Neyman–Pearson Lemma ([13], page 84 iv), the inequality in (3.10) is shown to be strict a.e. (W) for $\phi \neq \phi_0$ a.e. $(T \mid W)$. In fact, for $0 < \alpha < 1$, $(\alpha, \alpha E_0 T)$ is an interior point of the set

$$\{(\textstyle\int_{d_1}^{d_n} \phi(t, w) f_1(t)\,dt,\ \int_{d_1}^{d_n} \phi(t, w) t f_1(t)\,dt) \in R^2 \mid \phi \ \text{is a test on}\ [d_1, d_n] \times [0, \infty)\}\,,$$

which is nothing but the set $[0, 1] \times [0, E_0 T]$ if $E_0 T > 0$ or the set $[0, 1] \times [E_0 T, 0]$ if $E_0 T < 0$. Especially when the support of h is R^n for all $\gamma > 0$ and $\lambda \neq 0$, the pdf of W gives positive mass to any open set in $[0, \infty)$ and hence $\phi_0 \neq \phi$ a.e. implies $\phi_0 \neq \phi$ a.e. $(T \mid W)$, completing the proof.

THEOREM 4. *The test* ϕ_0 *in* (3.7) *is* UMPU *for testing* $H_0: h \in \mathscr{F}_0(\gamma I)$, $\gamma > 0$ *versus* $K_2: h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$, $\gamma > 0$, $\lambda \neq 0$.

PROOF. Since $\mathscr{D}_\alpha \subset \tilde{\mathscr{D}}_\alpha$, the result follows immediately from Theorem 3.

Next we state conditions for which the above results still hold for reparametrization of γ and λ. Let $\tau(\theta_1, \theta_2) = (\gamma(\theta_1, \theta_2), \lambda(\theta_1, \theta_2))$ denote a reparametrization of γ and λ by θ_1 and θ_2, where $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$ and Θ_i's are open intervals in R^1. Suppose τ satisfies the conditions: (1) τ is a continuous function from $\Theta_1 \times \Theta_2$ into R^2 such that the image $\tau(\Theta_1 \times \Theta_2)$ contains an open set in $(0, \infty) \times \Lambda$ and $\gamma(\theta_1, \theta_2) > 0$ for all $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$; (2) there exists a unique point $\theta_2^* \in \Theta_2$ such that $\lambda(\theta_1, \theta_2^*) = 0$ for all $\theta_1 \in \Theta_1$, and such that $\tau(\theta_1, \theta_2^*)$ is an interior point of the image for each θ_1, and (3) $\Sigma(\lambda(\theta_1, \theta_2))^{-1} = I + \lambda(\theta_1, \theta_2)A > 0$ for all $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$. Then almost clearly the above results hold in terms of (θ_1, θ_2). (See [13] Chapter 4 for the completeness result in Lemma 1.) Such reparametrizations are found in the examples below.

REMARK 1. The moments of X may not exist for some $h \in \mathscr{F}_i(\Psi)$ ($\Psi > 0$, $i = 0, 1, 2$). But if $E|X| < \infty$, then $E(X) = 0$, and if $E(X'X) < \infty$, Ψ is the covariance matrix of X.

REMARK 2. The null hypotheses which we are usually more interested in will be H_1: $h \in \mathscr{F}_1(\gamma\Sigma(0)) = \mathscr{F}_1(\gamma I)$ for the alternative K_1, and H_2: $h \in \mathscr{F}_2(\gamma\Sigma(0)) = \mathscr{F}_2(\gamma I)$ for the alternative K_2. Of course, the above results hold when H_0 is replaced by H_1 or H_2 for K_1 or K_2 respectively.

EXAMPLE 1. *Circular serial correlation* ([1]). Let $X = (X_1, \cdots, X_n)'$ be generated by $X_j = \rho X_{j-1} + u_j$, $X_0 = X_n$ $(j = 1, \cdots, n)$. Here the pdf of $U = (u_1, \cdots, u_n)'$ is assumed to belong to $\mathscr{F}_i(\gamma I)$ ($i = 1$ or 2), which contains $N(0, \gamma I)$. Since $X = \rho BX + U$ with

$$B = \begin{bmatrix} 0 & & & & 1 \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & 0 \end{bmatrix},$$

the pdf of X, say h, belongs to $\mathscr{F}_i(\gamma\Sigma(\lambda))$, where

$$\gamma = \gamma(\sigma^2, \rho) = \sigma^2/(1 + \rho^2), \qquad \lambda = \lambda(\sigma^2, \rho) = \rho/(1 + \rho^2), \tag{3.11}$$

and $\Sigma(\lambda)^{-1}$ is given by (1.2) with $A = -(B + B')$. By Theorem 1, for testing H_1: $h \in \mathscr{F}_1(\sigma^2 I)$, $\sigma^2 > 0$ versus K_1: $h \in \mathscr{F}_1(\gamma\Sigma(\lambda))$, $\sigma^2 > 0$, $\rho > 0$, the test $T \equiv X'AX/X'X < c$ is UMP. By Theorem 4, for testing H_2: $h \in \mathscr{F}_2(\sigma^2 I)$, $\sigma^2 > 0$ versus K_2: $h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$, $\sigma^2 > 0$, $\rho \neq 0$, the test $T < c_1$ or $T > c_2$ is UMPU.

4. Applications to a linear model. The results are applied to a regression model $y = X\beta + u$, where X: $n \times k$, rank $(X) = k$ and X is fixed. Let h be the pdf of error term u. Here $h \in \mathscr{F}_1(\gamma\Sigma(\lambda))$ or $h \in \mathscr{F}_2(\gamma\Sigma(\lambda))$ is assumed where $\Sigma(\lambda)^{-1}$ is given by (1.2). This assumption seems quite reasonable since it contains such heavy-tailed distributions as the multivariate t-(Cauchy) distribution, the contaminated normal distribution, etc. The crucial assumption is that the column space of X, denoted by $L(X)$, is spanned by some k latent vectors of A

in (1.2), say e_j's ($j = 1, \cdots, k$). Under this assumption, there exists a $Q \in \mathcal{O}(n)$ such that with $z = Qy$ and $v = Qu$

$$z = \binom{\beta^*}{0} + v \quad \text{and} \quad Q'AQ = \operatorname{diag}\{\tilde{d}_1, \cdots, \tilde{d}_n\} = D \quad (\beta^* : k \times 1) \tag{4.1}$$

where the first k $\tilde{d}_j$'s are the roots of A corresponding to e_j's. Let $z = (z_1', z_2')'$ where $z_1 : k \times 1$ and $z_2 : (n-k) \times 1$. Suppose that $h(u) = |\gamma\Sigma(\lambda)|^{-\frac{1}{2}} q(u'\Sigma(\lambda)^{-1}u/\gamma)$ belongs to $\mathcal{F}_i(\gamma\Sigma(\lambda))$ ($i = 1$ or 2). Then the pdf of $v = Qu$ is given by

$$\gamma^{-n/2}|I + \lambda D|^{\frac{1}{2}} q(v'[I + \lambda D]v/\gamma) . \tag{4.2}$$

Consider the problem of testing $H_1 : h \in \mathcal{F}_1(\gamma\Sigma(0))$, $\gamma > 0$ versus $K_1 : h \in \mathcal{F}_1(\gamma\Sigma(\lambda))$, $\gamma > 0$, $\lambda > 0$. Clearly the problem is invariant under the translation $z_1 \to z_1 + b$ ($b \in R^k$) in the canonical form (4.1). By invariance we only consider the class of level α tests based on z_2 alone, denoted by $\mathcal{D}_\alpha^I$. Since v has the pdf (4.2), as in Kelker [12], the marginal pdf of z_2 is given by

$$\bar{f}(z_2) = \gamma^{-(n-k)/2}|I_{n-k} + \lambda D_2|^{\frac{1}{2}} \bar{q}(z_2'[I + \lambda D_2]z_2/\gamma) \tag{4.3}$$

where $D_2 = \operatorname{diag}\{\tilde{d}_{n-k+1}, \cdots, \tilde{d}_n\}$. Here $\bar{q}$ depends only on the form of q and the integers k and n. For $h \in \mathcal{F}_1(\gamma\Sigma(\lambda))$, $\bar{q}$ is easily shown to be nonincreasing. Therefore the problem is reduced to that of testing $H_1 : \bar{f} \in \mathcal{F}_1(\gamma\Phi(0))$, $\gamma > 0$ versus $K_1 : \bar{f} \in \mathcal{F}_1(\gamma\Phi(\lambda))$, $\gamma > 0$, $\lambda > 0$ with $\Phi(\lambda)^{-1} = I_{n-k} + \lambda D_2$. Now Theorem 1 is applicable and the test with c.r. $T = z_2'D_2z_2/z_2'z_2 < c$ is UMP in the class $\mathcal{D}_\alpha^I$. In the same way, the test with c.r. $T < c_1$ or $T > c_2$ is shown to be UMPU in the class $\mathcal{D}_\alpha^I$ for testing $H_2 : h \in \mathcal{F}_2(\gamma\Sigma(0))$, $\gamma > 0$ versus $K_2 : h \in \mathcal{F}_2(\gamma\Sigma(\lambda))$, $\gamma > 0$, $\lambda \neq 0$. Here $\bar{q}$ is shown to be nonincreasing and convex. It should be remarked that if $D_2 = 0$ or $D_2 = aI_{n-k}$ for some $a > 0$, the problem cannot be tested through invariance. These results are summarized in the original terms as

THEOREM 5. *Let $y = X\beta + u$ be a linear model where $X : n \times k$ is fixed with* rank $(X) = k$. *Let h be the* pdf *of the error term u. Assume the column space is spanned by some k latent vectors of A in* (1.2). *Then for testing $H_1 : h \in \mathcal{F}_1(\gamma I)$, $\gamma > 0$ versus $K_1 : h \in \mathcal{F}_1(\gamma\Sigma(\lambda))$, $\gamma > 0$, $\lambda > 0$, the test which rejects H_1 for small values of*

$$T = y'MAMy/y'My \quad \text{where} \quad M = I - X(X'X)^{-1}X' \tag{4.4}$$

is UMPI *unless $MAM = 0$ or aM for some $a > 0$, and for testing $H_2 : h \in \mathcal{F}_2(\gamma I)$, $\gamma > 0$, versus $K_2 : h \in \mathcal{F}_2(\gamma\Sigma(\lambda))$, $\gamma > 0$, $\lambda \neq 0$, the* c.r. *$T < c_1$ or $T > c_2$ is* UMPUI (UMPU *in the class of invariant tests*) *unless $MAM = 0$ or aM for some $a > 0$.*

EXAMPLE 2. The model in Example 1 is generalized as

$$y_j = \mu + u_j , \qquad u_j = \rho u_{j-1} + v_j , \qquad v_0 = v_n , \tag{4.5}$$

where the pdf of $v = (v_1, \cdots, v_n)'$ belongs to $\mathcal{F}_i(\gamma I)$ ($i = 1$ or 2). Then the pdf of $u = (u_1, \cdots, u_n)'$, say h, belongs to $\mathcal{F}_i(\gamma\Sigma(\lambda))$, where γ and λ are given by (3.11) and $A = -(B + B')$. Let $y = (y_1, \cdots, y_n)'$ and $e = (1, \cdots, 1)'$ ($\in R^n$). Since $y = \mu e + u$ and e is a latent vector of A, Theorem 5 is applicable and the corresponding tests are UMPI or UMPUI. Here T in (4.4) is written as

$$\sum_{j=1}^{n} (y_j - \bar{y})(y_{j-1} - \bar{y}) / \sum_{j=1}^{n} (y_j - \bar{y})^2 , \quad y_0 = y_n$$

which is known as R. L. Anderson's circular serial correlation ([1]). The model is further generalized by Anderson–Anderson [2] to the case

$$E(y_j) = \sum \alpha_i \cos [2\pi j(i-1)/n] + \sum \beta_i \sin [2\pi j(i-1)/n] ,$$

in which the assumption in Theorem 5 is satisfied (see [1] for the latent vectors).

EXAMPLE 3. Let the error term u_j be generated by an autoregressive process of the type

$$u_j = \rho u_{j-1} + v_j , \quad |\rho| < 1 . \tag{4.6}$$

First assume $v = (v_1, \cdots, v_n)' \sim N(0, \sigma^2 I_n)$. Then the covariance matrix of $u = (u_1, \cdots, u_n)'$ is given by $\sigma^2 R$ where the (i, j) element of R is $\rho^{|i-j|}/(1-\rho^2)$. Let $C = (c_{ij})$ be the $n \times n$ matrix with $c_{ij} = 0$ except $c_{11} = c_{nn} = 1$, and let $B = (b_{ij})$ be the $n \times n$ matrix with $b_{ij} = 0$ except $b_{k,k-1} = b_{l,l+1} = 1$ ($k = 2, \cdots, n$; $l = 1, \cdots, n-1$). Then $R^{-1} = (1+\rho^2)I - \rho B - \rho^2 C$. The matrix R^{-1} is often approximated by Anderson's matrix Φ ([3]) where Φ differs from R^{-1} in having $(1 + \rho^2 - \rho)$ instead of 1 in the upper left and lower right corners ([3], [10]). The approximated model is written as $y = X\beta + u$, $u \sim N(0, \gamma\Sigma(\lambda))$ where $\Sigma(\lambda)^{-1} = I + \lambda A$, $\lambda = \sigma^2/(1-\rho^2)$, $\lambda = \rho/(1-\rho^2)$ and $A = -(B + C)$.

Suppose the distribution of u may not be normal but is just known to be (or well approximated by) a member of $\mathscr{F}_i(\gamma\Sigma(\lambda))$ ($i = 1$ or 2). Then if X satisfies the assumption in Theorem 5, the test based on T in (4.4) has the optimal properties stated in Theorem 5. The latent roots of A are $d_i = \cos [\pi(i-1)/n]$ and the latent vectors are $(\cos [\pi(i-1)/2n], \cos [3\pi(i-1)/2n], \cdots, \cos [(2n-1)\pi(i-1)/2n])'$ ($i = 1, \cdots, n$). The statistic $T + 2 = y'M[2I + A]My/y'My$ is well known as the Durbin–Watson test statistic ([5], [6]) and when $k = 1$ and $X = e$, the test statistic $n(T+2)/(n-1)$ is also well known as the von Neumann ratio, which is usually written as $\delta^2/s^2 = n \sum_{i=2}^{n} (y_i - y_{i-1})^2/(n-1) \sum_{i=1}^{n} (y_i - \bar{y})^2$. Since e is a latent vector of A, the test based on δ^2/s^2 has the optimality property ([16], [17]).

EXAMPLE 4. Let $\Psi = \sigma^2[(1-\rho)I + \rho ee']$ (intra-class covariance structure). Then $\Psi^{-1} = [I - \rho ee'/(n\rho - \rho + 1)]/\sigma^2(1-\rho)$. Hence with $\gamma = \sigma^2(1-\rho)$ and $\lambda = \rho/(n\rho - \rho + 1)$, $\Psi^{-1} = [I + \lambda A]/\gamma$ where $A = -ee'$. However if the regression equation contains a constant term or the column space $L(X)$ contains the vector e, the problem stated in Theorem 5 cannot be tested since $MAM = 0$. If $L(X)$ does not contain e and if each column of X is orthogonal to e, then Theorem 5 is clearly applicable and $(n-k)(y'e)^2/ny'My = (n-k)T/n$ is F-distributed with degrees of freedom 1 and $n - k$ under H_i ($i = 1$ or 2) ([11]).

REMARK 3. Even under normality, the assumption concerning $L(X)$ in Theorem 5 is necessary for obtaining the corresponding results. On the other hand, Durbin and Watson [7] have shown that under normality their test in Example 3 is locally most powerful invariant for the one-sided problem even if $L(X)$ is not spanned by any k latent vectors of A. The result is applicable to the tests based on T in (4.4) under normality. However, without normality, or in our setting such a result seems difficult to obtain.

Acknowledgment. The author wishes to thank the referees for their valuable comments.

REFERENCES

[1] ANDERSON, R. L. (1942). Distribution of the serial correlation coefficient. *Ann. Math. Statist.* **13** 1-13.

[2] ANDERSON, R. L. and ANDERSON, T. W. (1950). Distribution of the circular serial correlation for residuals from a fitted Fourier series. *Ann. Math. Statist.* **21** 59-81.

[3] ANDERSON, T. W. (1948). On the theory of testing serial correlation. *Skand. Aktuarietidskr.* **31** 88-116.

[4] BASU, D. (1955). On statistics independent of a complete sufficient statistic. *Sankhyā* **15** 377-380.

[5] DURBIN, J. and WATSON, G. S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika* **37** 409-428.

[6] DURBIN, J. and WATSON, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika* **38** 159-178.

[7] DURBIN, J. and WATSON, G. S. (1971). Testing for serial correlation in least squares regression. III. *Biometrika* **58** 1-19.

[8] EATON, M. L. (1972). *Multivariate Statistical Analysis*. Univ. of Copenhagen.

[9] JOHNSON, N. L. and KOTZ, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*. Wiley, New York.

[10] KADIYALA, K. R. (1970). Testing for the independence of regression disturbance. *Econometrica* **38** 97-117.

[11] KARIYA, T. and EATON, M. L. (1977). Robust tests for spherical symmetry. *Ann. Statist.* **5** 206-215.

[12] KELKER, D. (1970). Distribution theory of spherical distributions and a location scale parameter generation. *Sankhyā* **A32** 419-430.

[13] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[14] PRESS, S. J. (1969). On serial correlation. *Ann. Math. Statist.* **40** 186-196.

[15] TILLMAN, J. A. (1975). The power of the Durbin-Watson test. *Econometrica* **43** 959-974.

[16] VON NEUMANN, J. (1941). Distribution of mean square successive difference to the variance. *Ann. Math. Statist.* **12** 367-395.

[17] VON NEUMANN, J. (1942). A further remark concerning the distribution of the ratio of the mean square successive difference to the variance. *Ann. Math. Statist.* **13** 86-88.

INSTITUTE OF ECONOMIC RESEARCH
HITOTSUBASHI UNIVERSITY
KUNITACHI, TOKYO 186,
JAPAN