The Annals of Statistics
1980, Vol. 8, No. 5, 1065–1070

LOCALLY ROBUST TESTS FOR SERIAL CORRELATION IN LEAST SQUARES REGRESSION[1]

BY TAKEAKI KARIYA

Hitotsubashi University

Kariya and Eaton and Kariya studied a robustness property of the usual tests for serial correlation against departure from normality. When the results were applied to a regression model $y = X\beta + u(X: nxk)$, it was assumed that the column space of X is spanned by some k latent vectors of the covariance matrix of error term u. In this paper we delete this assumption and in a much broader class of distributions derive a locally best invariant test for a one-sided problem and a locally best unbiased and invariant test for a two-sided problem. The null distributions of these tests are the same as those under normality.

1. Introduction. Usually tests for serial correlation in a linear regression model

$$y = X\beta + u, \quad \text{where} \quad X: n \times k \text{ and rank } (X) = k, \tag{1.1}$$

are developed under normality for u;

$$u \sim N_n(0, \Sigma(\sigma^2, \rho)), \text{ where } \Sigma(\sigma^2, \rho) = \sigma^2\Phi(\rho), \tag{1.2}$$

$$\Phi(\rho)^{-1} = I_n + \rho A, \text{ and } \rho \in \Lambda = \{\rho \in R | \Phi(\rho)^{-1} > 0\}. \tag{1.3}$$

Here A is an $n \times n$ known matrix and $\Phi(\rho)^{-1} > 0$ denotes the positive definiteness of $\Phi(\rho)^{-1}$ (see [2], [8], [9]). In this normal model, for testing $H: \rho = 0$ versus $K: \rho > 0$, the test which rejects for small values of

$$T = y'NANy/y'Ny, \text{ where } N = I - X(X'X)^{-1}X', \tag{1.4}$$

is a uniformly most powerful invariant (UMPI) test provided the following assumption holds;

$$L(X) \text{ is spanned by some } k \text{ latent vectors of } A, \tag{1.5}$$

where $L(X)$ denotes the column space of X. And for testing $H: \rho = 0$ versus $K: \rho \neq 0$, the test with c.r. (critical region) $T < c_1$ or $T > c_2$ is uniformly most powerful unbiased and invariant (UMPUI) under the assumption (1.5) (Anderson [2]). Under (1.5), Kariya and Eaton [9] and Kariya [8] showed that the UMPI and UMPUI properties of these tests can be extended to much broader classes of distributions and that the null distribution of T under any member of the classes is the same as that under normal distribution (1.2) with $\sigma^2 = 1$ and $\rho = 0$. It should be noted that under (1.5) the generalized least squares estimator is identically equal to the ordinary least squares estimator since X is Σ-invariant (see [11]). Hence if the purpose of testing $\rho = 0$ lies in checking the appropriateness of the use of the

Received September 1978; revised May 1979.

[1]This paper was prepared with support of Matsunaga Research Foundation.

AMS 1970 *subject classifications.* Primary 62G10; secondary 62F05, 62G35, 62J05.

Key words and phrases. Robustness, LBI test, LBUI test, serial correlation, invariance, Durbin-Watson test, least squares regression.

ordinary least squares estimator rather than detecting the existence of serial correlation, the assumption (1.5) is not interesting. Further, the assumption is not satisfied in many problems. The Anderson-Anderson test ([1]) is an example which satisfies it, while in the Durbin-Watson test it is not satisfied in general ([3], [4], [5]).

In this paper, without the assumption (1.5), the one-sided test is shown to be LBI (locally best invariant) in a much broader class of distributions, and for the two-sided problem, without the assumption, an LBUI (locally best unbiased and invariant) test is derived in the same class. The LBUI test is different from the two-sided test based on T stated above. Further the null distributions of these test statistics under any member of the class are the same as those under $N(0, I)$. Since the class of distributions we consider contains normal distribution, these tests are identically equal to the LBI and LBUI tests under normality (1.2). Durbin and Watson [5] has shown under normality (1.2) that the test with c.r.$T < c$ is LBI for testing $\rho = 0$ versus $\rho > 0$. In this sense, our result shows that the LBI property is robust against departure form normality. On the other hand, even under normality (1.2), it seems that the LBUI test has not been derived yet. Under the assumption (1.5) the LBUI is naturally reduced to the two-sided test based on T. To derive the distribution of a maximal invariant, a theorem due to Wijsman [13] is used although direct derivation is possible.

2. Problem and the results in [8], [9]. To state our problem in this paper, we define three classes of pdf's (probability density functions) with respect to Lebesgue measure on Euclidean n-space R^n. Let $\mathcal{F}$ be the class of all pdf's on R^n and with an $n \times n$ matrix $\Sigma > 0$, let

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

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

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

Clearly $\mathcal{F}_2(\Sigma) \subset \mathcal{F}_1(\Sigma) \subset \mathcal{F}_0(\Sigma)$. If $f(x) = |\Sigma|^{-\frac{1}{2}} q(x'\Sigma^{-1}x)$ belongs to $\mathcal{F}_2(\Sigma)$, then $g(x) = \int_0^\infty |\Sigma|^{-\frac{1}{2}} a^{-\frac{n}{2}} q(x'\Sigma^{-1}x/a)dG(a)$ also belongs to $\mathcal{F}_2(\Sigma)$ where G is a distribution function on $(0,\infty)$. Hence $\mathcal{F}_2(\Sigma)$ contains contaminated normal distribution, multivariate t-distribution, multivariate Cauchy distribution, etc., as well as $N(0, \Sigma)$.

Now suppose that h is a pdf of the error term u in (1.1). In the papers [8] and [9], under the assumption (1.5) the one-sided problem

(2.4) $$H_1 : h \in \mathcal{F}_1(\sigma^2 I), \sigma^2 > 0 \text{ versus } K_1 : h \in \mathcal{F}_1(\Sigma(\sigma^2, \rho)), \sigma^2 > 0, \rho > 0,$$

and the two-sided problem

(2.5) $$H_2 : h \in \mathcal{F}_2(\sigma^2 I), \sigma^2 > 0 \text{ versus } K_2 : h \in \mathcal{F}_2(\Sigma(\sigma^2, \rho)), \sigma^2 > 0, \rho \neq 0,$$

are considered where $\Sigma(\sigma^2, \rho)$ is given by (1.2) and (1.3). The results obtained there are summarized as

THEOREM 1. *Assume* (1.5) *and* $NAN \neq aN$ *for any* $a \in R$. *Then for the one-sided problem* (2.4), *the test with* c.r.$T < c$ *is* UMPI, *and for the two-sided problem* (2.5), *the test with* c.r.$T < c_1$ *or* $T > c_2$ *is* UMPUI. *Further the null distribution of* T *under* $h \in \mathcal{F}_i(\sigma^2 I)$ $(i = 1, 2)$ *is the same as that under* $N(0, I)$.

The group leaving the problems invariant is defined in the next section.

Without assuming (1.5), this paper treats the one-sided problem

(2.6)

$$H_0 : h \in \mathcal{F}_0(\sigma^2 I), \sigma^2 > 0 \text{ versus } K_0 : h \in \mathcal{F}_0(\Sigma(\sigma^2, \rho)), \qquad \sigma^2 > 0, \rho > 0,$$

and the two-sided problem

(2.7)

$$H_{00} : h \in \mathcal{F}_0(\sigma^2 I), \sigma^2 > 0 \text{ versus } K_{00} : h \in \mathcal{F}_0(\Sigma(\sigma^2, \rho)), \qquad \sigma^2 > 0, \rho \neq 0.$$

3. Main results. Let $R_+ = \{a \in R \mid a > 0\}$ and $G = R^k \times R_+$. Then the problems (2.6) and (2.7) are left invariant under the group G with the action;

$$y \to ay + Xg, \beta \to a\beta + g \text{ and } \sigma^2 \to a^2\sigma^2, \tag{3.1}$$

where $(g, a) \in G$, since under $h \in \mathcal{F}_0(\Sigma(\sigma^2, \rho))$, y has a pdf of the form

$$f(y \mid \beta, \sigma^2, \rho) = |\Sigma(\sigma^2, \rho)|^{-\frac{1}{2}} q\Big((y - X\beta)'\big[\Sigma(\sigma^2, \rho)\big]^{-1}(y - X\beta)\Big). \tag{3.2}$$

Choose a matrix $Z : n \times (n-k)$ such that $Z'Z = I_{n-k}$ and $ZZ' = N = I - X(X'X)^{-1}X'$, and define $v = (X'X)^{-\frac{1}{2}}X'y, \eta = (X'X)^{+\frac{1}{2}}\beta, w = Z'y$ and

$$H = \begin{pmatrix} (X'X)^{-\frac{1}{2}}X' \\ Z' \end{pmatrix}.$$

Then $w/\|w\|$ is clearly a maximal invariant under G, H is an orthogonal matrix and $(y - X\beta)'\Sigma(\sigma^2, \rho)^{-1}(y - X\beta) = (Hy - HX\beta)'[H\Sigma(\sigma^2, \rho)H']^{-1}(Hy - HX\beta)$, where $\|w\| = (w'w)^{\frac{1}{2}}$. From Kelker [10], the marginal pdf of w is now of the form

$$\bar{f}(w \mid \sigma^2, \rho) = |Z'\Sigma(\sigma^2, \rho)Z|^{-\frac{1}{2}} \bar{q}\Big(w'\big[Z'\Sigma(\sigma^2, \rho)Z\big]^{-1}w\Big), \tag{3.3}$$

since the pdf of $\tilde{y} = Hy$ is from (3.2)

$$|H\Sigma(\sigma^2, \rho)H'|^{-\frac{1}{2}} q\Big((\tilde{y} - \tilde{\eta})'\big[H\Sigma(\sigma^2, \rho)H'\big]^{-1}(\tilde{y} - \tilde{\eta})\Big) \quad \text{where} \quad \tilde{\eta} = \binom{\eta}{0}.$$

Here $\bar{q}$ depends on q and (n, k) but not on β, σ^2, ρ, and A. In terms of w, $w/\|w\|$ is a maximal invariant under group R_+ with the action;

$$w \to aw \text{ and } \sigma^2 \to a^2\sigma^2 \text{ for } a \in R_+, \tag{3.4}$$

and an invariant test is a test based on $w/\|w\|$ only. Without loss of generality, we assume $\sigma^2 = 1$. The next lemma is an application of a theorem due to Wijsman [13].

LEMMA 1. *Let $t(w)$ be a maximal invariant under the transformation* (3.4). *Let P_ρ^T be the distribution induced by $T = t(w)$ under ρ. Then the* pdf *of T with respect to P_0^T evaluated at $T = t(w)$ is given by*

$$\text{(3.5)} \qquad \frac{dP_\rho^T}{dP_0^T}(t(w)) = f_T(t(w)|\rho) = \frac{\int_0^\infty a^{n-k-1}\bar{f}(aw|1,\rho)da}{\int_0^\infty a^{n-k-1}\bar{f}(aw|1,0)da},$$

where $\bar{f}$ is the pdf *in* (3.3). *Naturally $f_T(t(w)|0) = 1$.*

PROOF. In Theorem 4 of Wijsman [13], let $G = \{aI_{n-k}|a \in R_+\}, \mu_G(dg) = da/a$, and $X = R^{n-k}$. Then the result follows when R^{n-k} is shown to be a linear Cartan G-space. Since $P(w = 0) = 0$, it is sufficient to show that $X^0 = R^{n-k} - \{0\}$ is a linear Cartan G-space. For $x \in X^0$, let $V(x) = \{z \in R^{n-k}|\|z - x\| < \|x\|/2\}$. Then the closure of the set $\{aI \in G|aIV(x) \cap V(x) \neq \phi\}$ is compact, which completes the proof.

Now from (3.3), the numerator of (3.5) is evaluated as

$$\text{(3.6)} \qquad |Z'\Phi(\rho)Z|^{-\frac{1}{2}}\left[w'(Z'\Phi(\rho)Z)^{-1}w\right]^{-(n-k)/2}\int_0^\infty a^{n-k-1}\bar{q}(a^2)da.$$

Here the integral in (3.6) is finite since $\bar{q}(w'w)$ is a pdf on R^{n-k} (see Kelker [10]). Further it is noted that (3.6) holds even when $\bar{q}$ vanishes in some region on $[0,\infty)$. The denominator of (3.5) is simply obtained by setting $\rho = 0$ in (3.6) and so (3.5) is finally evaluated as

$$\text{(3.7)} \qquad f_T(t(w)|\rho) = |Z'\Phi(\rho)Z|^{-\frac{1}{2}}\left\{w'[Z'\Phi(\rho)Z]^{-1}w/w'w\right\}^{-(n-k)/2}.$$

Since this is the pdf of a maximal invariant T with respect to P_0^T with $T = t(w)$, based on this pdf, we test $H_0 : \rho = 0$ versus $K_0 : \rho > 0$, and $H_{00} : \rho = 0$ versus $K_{00} : \rho \neq 0$. From Ferguson [7] pages 235–238, for testing H_0 versus K_0, a locally best invariant test, say ϕ_1, is given by the c.r.

$$\text{(3.8)} \qquad \left[\partial f_T(t(w)|\rho)/\partial\rho\right]|_{\rho=0} > cf_T(t(w)|0) = c$$

and for testing H_{00} versus K_{00}, a locally best unbiased and invariant test, say ϕ_2, is given by the c.r.

$$\text{(3.9)} \quad \left[\partial^2 f_T(t(w)|\rho)/\partial\rho^2\right]|_{\rho=0} > c_1 f_T(t(w)|0) + c_2\left[\partial f_T(t(w)|\rho)/\partial\rho\right]|_{\rho=0}.$$

Evaluating (3.8) yields the c.r. of the form $T < c$, and evaluating (3.9) yields the c.r. of the form

$$\text{(3.10)} \qquad T^2 + \frac{4}{n-k+2}\frac{y'NAMANy}{y'Ny} > c_3 T + c_4,$$

where T is defined by (1.4) and $M = X(X'X)^{-1}X'$. Here c_3 and c_4 are chosen such that for a significance level $\alpha(0 < \alpha < 1), \phi_2$ satisfies the size condition $E_0[\phi_2] =$ and the condition for local unbiasedness, $[\partial E_\rho(\phi_2)/\partial\rho]|_{\rho=0} = 0$. Since $E_\rho(\phi_2) = \int\phi_2(t(w))f_T(t(w)|\rho)dP_0^T$, the latter condition is equivalent to $E_0(T\phi_2) = \alpha E_0(T)$ or

$$\text{(3.11)} \qquad E_0[T\phi_2] = \alpha \mathrm{tr}NA/(n-k).$$

Thus we obtain

THEOREM 2. *Assume* $NAN \neq aN$ *for any* $a \in R$. *Let* h *be the* pdf *of error term* u *in* (1.1). *Then for the one-sided testing problem in* (2.6), *an* LBI *test* ϕ_1 *is given by the* c.r. $T < c$, *and for the two-sided testing problem in* (2.7) *an* LBUI *test* ϕ_2 *is given by the* c.r. (3.10) *where* c_3 *and* c_4 *are determined from* $E_0(\phi_2) = \alpha$ *and* (3.10). *Further the null distributions of these test statistics under any* $h \in \mathcal{F}_0(\sigma^2 I_n)$ *are the same as those under* $N(0, I)$.

The latter part of this theorem is clear from [9].

As Theorem 2 shows, without the assumption (1.5) the LBUI test ϕ_2 is not the same as the UMPUI test under (1.5), while the LBI test ϕ_1 is the same as the UMPI test under (1.5). When (1.5) holds, $NAM = 0$ and the c.r. (3.10) is reduced to

$$T^2 - c_3 T - c_4 > 0. \tag{3.12}$$

In order for (3.12) to satisfy $E_0(\phi_2) = \alpha$ for $0 < \alpha < 1, c_3^2 + 4c_4 > 0$ is necessary and sufficient, in which case the c.r. $T > c'$ or $T < c$ is derived. Hence the LBI test coincides with the two-sided test based on T if and only if (1.5) holds. That is, without (1.5), the usual two-sided test is not LBUI. Secondly, it is noted that the pdf of a maximal invariant $T = t(w)$ does not depend on $h \in \mathcal{F}_0(\Sigma(\sigma^2, \rho))$ since P_0^T is a uniform distribution on the sphere $\{x \in R^{n-k} \mid \|x\| = 1\}$ and is independent of h. Hence since any $h \in \mathcal{F}_0(\Sigma(\sigma^2, \rho))$ leads to the same pdf in (3.7), even if h is the pdf of $N(0, \Sigma(\sigma^2, \rho))$, nothing new comes out. When $u \sim N(0, \Sigma(\sigma^2, \rho))$, Durbin and Watson [5] have shown that the test with c.r. $T < c$ is LBI for testing $\rho = 0$ versus $\rho > 0$. But they did not derive the LBUI test. Finally we remark that the pdf of a maximal invariant in (3.7) can be directly derived by transforming w into the polar coordinates in R^{n-k}, as the referee points out.

Acknowledgment. The author appreciates the referees for valuable comments and informing him that the result on the one-sided testing problem is also contained in a paper by M. L. King submitted to *The Annals of Statistics*.

REFERENCES

[1] 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.

[2] ANDERSON, T. W. (1948). On the theory of testing serial correlation. *Skand. Aktuar.*, **31** 88–116.

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

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

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

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

[7] FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, New York.

[8] KARIYA, T. (1977). A robustness property of the tests for serial correlation. *Ann. Statist.* **5** 1212–1220.

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

[10] KELKER, D. (1970). Distribution theory of spherical distribution and a location scale parameter generalization. *Sankhyā A* **32** 419–430.

[11] KRUSKAL, W. (1961). The coordinate-free approach to Gauss-Markov estimation and its application to missing and extra observations. *Fourth Berk. Symp. Math. Statist. Probability* **1** 435–451. Univ. California, Berkley.

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

[13] WIJSMANN, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Fifth Berk. Symp. Math. Statist. Probability,* **1** 389–400. Univ. California, Berkley.

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