Comparisons of Tests for the Presence of Random Walk Coefficients in a Simple Linear Model

JUKKA NYBLOM and TIMO MÄKELÄINEN*

The locally most powerful test is derived for the hypothesis that the regression coefficients are constant over time against the alternative that they vary according to the random walk process. When the regression equation contains the constant term only, comparisons are made with the tests suggested by LaMotte and McWhorter (1978). These are based on exact powers and on three different types of asymptotic efficiencies including the classical Pitman and Bahadur approaches and the new one due to Gregory (1980). The concept of the Bahadur efficiency is extended to cover also the random slopes. Suggestions are made for choosing the test.

KEY WORDS: Linear models; Time varying parameters; Pitman efficiency; Bahadur efficiency; Locally most powerful tests.

1. INTRODUCTION

It is common in time series regression work, in such fields as economics, that the statistical relationship under consideration may not remain constant over time, but appears to be subject to more or less gradual change. To deal with this and many other interesting situations, attention has been paid extensively in recent literature to the possibilities of stochastically modeling those changes that cannot be accounted for by the systematic part of the model. For a review see Rosenberg (1973). For further references see also LaMotte and McWhorter (1978).

The latter authors considered the particular Kalman filter or sequential parameter regression model

$$y_t = \mathbf{x}_t' \boldsymbol{\beta}_t + \epsilon_t, \qquad (1.1)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\delta}_t, \quad t = 1, 2, \ldots, T \qquad (1.2)$$

consisting of the regression part (1.1) with the observables y_t (random) and $\mathbf{x}_t$ ($p \times 1$, fixed) and the error terms ϵ_t, and of the $p \times 1$ vector parameter process (1.2), which is defined through the joint distribution of the disturbance vectors $\boldsymbol{\delta}_t$ and the fixed initial value $\boldsymbol{\beta}_0$. It is assumed here that $\epsilon_t \sim N(0, \sigma^2)$ and $\boldsymbol{\delta}_t \sim N(0, \tau^2 \mathbf{G})$ with $\mathbf{G}$ known but σ^2 and τ^2 unknown, and that these random variables are jointly independent.

The model is one of the simplest of the possible generalizations of the regression model in this direction. The constant parameter regression model is itself included as the special case $\tau^2 = 0$. LaMotte and McWhorter proposed a family of exact tests for the hypothesis $\rho = \tau^2/\sigma^2 = 0$ against $\rho > 0$. In view of our introductory remarks, this testing problem may be regarded as the diagnostic problem of establishing whether the simpler constant coefficients' regression model would in fact be adequate.

The purpose of the present article is to make comparisons between the tests of LaMotte and McWhorter and the locally most powerful invariant test, the LMPI test (to be obtained in Sec. 2), in the special case $p = 1$, $\mathbf{G} = 1$, $\mathbf{x}_t = 1$. The special model is one of a random walk observed with error. Prediction with this model has been extensively studied, and it is closely related to exponential smoothing (see Harrison and Stevens 1976). Note, however, that the process (1.2) arises in Harrison's and Stevens's work from a sequence of consecutive prior judgments by the forecaster while its role in our (and most other writers') context is simply that of a latent (unobservable) process.

Our comparisons of the tests are made on the basis of exact power calculations and also of asymptotic relative efficiency. In Section 2 the tests are introduced. Calculations of the critical points of the LMPI test and of the powers of the various tests are made in Section 3. These calculations show that near the null hypothesis the LMPI test is more powerful than any of the LaMotte and McWhorter tests. This is, of course, in accordance with the optimal character of the LMPI test. At more distant alternatives a more powerful test than the LMPI test can be found among the LaMotte and McWhorter tests. The asymptotic distributions of the test statistics are derived in Section 4. The optimal choices of the LaMotte and McWhorter tests are given in terms of the Pitman and the Bahadur asymptotic efficiencies in Section 5. Surprisingly these results are quite contradictory. Some approximate comparisons are made between the LMPI and the LaMotte and McWhorter tests based on the Pitman efficiencies. Tests are also compared by means of an asymptotic efficiency measure due to Gregory (1980). The picture from these comparisons is again different

* Jukka Nyblom is lecturer in Statistics, Section of Social and Regional Sciences, University of Joensuu, P.O. Box 111, SF-80101 Joensuu, Finland. Timo Mäkeläinen is Professor of Mathematical Statistics, Department of Mathematics, University of Helsinki, Hallituskatu 15, SF-00100 Helsinki, Finland. The research was completed when the first writer was a Junior Research Fellow in the Academy of Finland. The problem and the LMPI test considered in the article were suggested by the second writer. The study of the asymptotic distributions and the efficiencies and the calculation of the tables were carried out by the first writer.

© Journal of the American Statistical Association
December 1983, Volume 78, Number 384
Theory and Methods Section

from the previous ones. Appendix Sections A.1 and A.2 briefly discuss the notions of the Pitman and Bahadur asymptotic efficiencies. The Pitman efficiency is treated in a general setting allowing different asymptotic distributions for the competing tests. The definition of the Bahadur efficiency covers the case where the so-called Bahadur slope is not a constant but a nondegenerate random variable. In Appendix Section A.3 large deviation probabilities, needed in the asymptotic efficiency calculations, are given for linear combinations of χ^2 variables and ratios of such combinations.

2. TESTS FOR CONSTANCY OF THE PARAMETER PROCESS

We formulate here three types of tests for the hypothesis of the constancy of the regression coefficients. It will be obvious that the performance of any test must depend on the values of the explanatory variables, $\mathbf{x}_t$, in the data. Consequently, no progress appears possible without the study of specific choices of $\mathbf{x}_t$.

Our simple model

$$y_t = \beta_t + \epsilon_t,$$

$$\beta_t = \beta_{t-1} + \delta_t, \quad t = 1, 2, \ldots, T \tag{2.1}$$

(specified as in (1.1)-(1.2) and with $\mathbf{G} = 1$) has the advantage of admitting asymptotic analyses and also of having an interest of its own, as discussed in the Introduction. Despite the great simplicity of the model it need not lead to untypical results as far as more general models are concerned. Thus it turns out that our recommendations concerning the choice of the LaMotte and McWhorter tests are essentially the same as the ones given by LaMotte and McWhorter on the basis of their particular example of model (1.1)-(1.2).

As a departure from the main line of development let us briefly consider the situation from the point of view of likelihood ratio tests. The model (2.1) can be interpreted as the subfamily of the IMA (1, 1) processes having the MA parameter θ restricted by $0 < \theta \leq 1$ (Box and Jenkins 1970, p. 123). The value $\theta = 1$ corresponds to our null hypothesis. This value is on the boundary of the invertibility region $|\theta| < 1$, inside which the maximum likelihood estimator of θ is asymptotically $N(\theta, (1 - \theta^2)/T)$. Clearly special considerations would be called for if maximum likelihood methods were to be employed. Further evidence about the nonregular character of our problem is contained in later sections.

Since no one of our tests will appreciably simplify for model (2.1) we shall introduce them in terms of model (1.1)-(1.2). We also note that power optimal tests have, to our knowledge, not been studied previously for model (1.1)-(1.2).

From (1.1)-(1.2) we write

$$y_t = \mathbf{x}_t'\boldsymbol{\beta}_0 + \mathbf{x}_t'\boldsymbol{\delta}_1 + \cdots + \mathbf{x}_t'\boldsymbol{\delta}_t + \epsilon_t, \quad t = 1, \ldots, T.$$

Consequently $Ey_t = \mathbf{x}_t'\boldsymbol{\beta}_0$ and

$$\operatorname{cov}(y_s, y_t) = \tau^2 \cdot \min(s, t) \cdot \mathbf{x}_s'\mathbf{G}\mathbf{x}_t + \delta_{st}\sigma^2,$$

so that $\mathbf{y} = (y_1, \ldots, y_T)' \sim N(\mathbf{X}\boldsymbol{\beta}_0, \tau^2\mathbf{W}*(\mathbf{XGX}') + \sigma^2\mathbf{I})$ where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_T)'$ and $\mathbf{W} = [\min(s, t)]$, and an asterisk denotes the Hadamard (or elementwise) product of matrices. Writing $\rho = \tau^2/\sigma^2$ and $\mathbf{V} = \mathbf{W}*(\mathbf{XGX}')$ we have that

$$\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}_0, \sigma^2(\mathbf{I} + \rho\mathbf{V})).$$

We want to test the hypothesis H_0: $\rho = 0$ against H_1: $\rho > 0$. As noticed by LaMotte and McWhorter (1978) the problem is invariant in translations $\mathbf{y} \to \mathbf{y} + \mathbf{Xb}$, where $\mathbf{b}$ is a $p \times 1$ vector. A maximal invariant is $\mathbf{Z}'\mathbf{y}$, where the columns of $\mathbf{Z}$ form an orthonormal basis for the orthogonal complement of the column space of $\mathbf{X}$. We have

$$\mathbf{Z}'\mathbf{y} \sim N(\mathbf{0}, \sigma^2(\mathbf{I} + \rho\mathbf{Z}'\mathbf{VZ})).$$

Consider the spectral decomposition

$$\mathbf{Z}'\mathbf{VZ} = \lambda_1 \mathbf{P}_1 + \cdots + \lambda_n \mathbf{P}_n, \tag{2.2}$$

where $\lambda_1 > \lambda_2 > \cdots > \lambda_n \geq 0$. Here we have $\lambda_n > 0$ at least if no row of $\mathbf{X}$ is in the null space of $\mathbf{G}$, since the rank of $\mathbf{W}*\mathbf{XGX}'$ equals the number of positive diagonal elements of $\mathbf{XGX}'$ (see Lemma 3.4 in Styan 1973, e.g.). Now

$$\mathbf{y}'\mathbf{ZP}_k\mathbf{Z}'\mathbf{y} \sim \sigma^2(1 + \rho\lambda_k)\chi^2(r_k), \tag{2.3}$$

where $r_k = \operatorname{rank}(\mathbf{P}_k)$, $k = 1, \ldots, n$, these variables being independent. LaMotte and McWhorter suggested tests of the form

$$F_g = \frac{\sum_{k=1}^{g} \mathbf{y}'\mathbf{Z}\,\mathbf{P}_k\mathbf{Z}'\mathbf{y}/n_g}{\sum_{k=g+1}^{n} \mathbf{y}'\mathbf{Z}\,\mathbf{P}_k\mathbf{Z}'\mathbf{y}/m_g} > c, \tag{2.4}$$

where $n_g = r_1 + \cdots + r_g$ and $m_g = r_{g+1} + \cdots + r_n$. Under H_0, $F_g \sim F(n_g, m_g)$. LaMotte and McWhorter also give some guidance in how to choose the number g, primarily on the basis of a set of empirical results.

Because the problem is invariant not only in translations but also in scale transformations, we shall consider a further reduction by invariance. It can be shown that the most powerful invariant test against the alternative hypothesis $\rho = \rho_1$ has the critical region

$$\frac{\hat{\mathbf{e}}'\hat{\mathbf{e}}}{\tilde{\mathbf{e}}'(\mathbf{I} + \rho_1\mathbf{V})^{-1}\tilde{\mathbf{e}}} > c, \tag{2.5}$$

where $\hat{\mathbf{e}}$ contains the least squares residuals and $\tilde{\mathbf{e}}$ the generalized least squares residuals. The results may be found in King (1980). (Durbin and Watson (1971) gave an incorrect expression, the numerator of their statistic being $\tilde{\mathbf{e}}'\tilde{\mathbf{e}}$ instead of the correct $\hat{\mathbf{e}}'\hat{\mathbf{e}}$.) Because the test (2.5) depends on ρ_1, no UMPI test exists. The locally most powerful invariant (LMPI) test is therefore worth examining. From Durbin and Watson and King we obtain that the LMPI test rejects when

$$L = (\hat{\mathbf{e}}'\mathbf{V}\hat{\mathbf{e}})/(\hat{\mathbf{e}}'\hat{\mathbf{e}}) > c. \tag{2.6}$$

In contrast to (2.4) the distribution of (2.6) is cumbersome to deal with. Writing $\hat{\mathbf{e}} = \mathbf{ZZ}'\mathbf{y}$ and using the fact that $\mathbf{P}_1 + \cdots + \mathbf{P}_n = \mathbf{I}$ with $\mathbf{Z}, \mathbf{P}_1, \ldots, \mathbf{P}_n$ as above, (2.6) takes the form

$$L = \frac{\sum_{k=1}^{n} \lambda_k \mathbf{y}'\mathbf{Z}\,\mathbf{P}_k\mathbf{Z}'\mathbf{y}}{\sum_{k=1}^{n} \mathbf{y}'\mathbf{Z}\,\mathbf{P}_k\mathbf{Z}'\mathbf{y}} > c. \qquad (2.7)$$

Under H_0 the numerator of L is a linear combination of independent χ^2 variables and the denominator a χ^2 variable.

3. EXACT DISTRIBUTIONS AND POWER COMPARISONS

In the model (2.1) we have $\mathbf{X} = \mathbf{1} = (1, \ldots, 1)'$ and $\mathbf{V} = \mathbf{W}$ with $\mathbf{X}, \mathbf{V}, \mathbf{W}$ as in Section 2. The test (2.6) can be written as

$$L = \frac{\sum_{st} \min(s, t)(y_s - \bar{y})(y_t - \bar{y})}{\sum_t (y_t - \bar{y})^2} = \frac{\sum_{t=1}^{T} \left[\sum_{s=t}^{T} (y_s - \bar{y})\right]^2}{\sum_t (y_t - \bar{y})^2} > c. \qquad (3.1)$$

In order to derive its distribution we must determine the eigenvalues $\lambda_{kT} = \lambda_k$ occuring in the representations (2.2) and (2.7). Here $\mathbf{Z}$ is $T \times (T - 1)$, $\mathbf{Z}'\mathbf{1} = 0$, and $\mathbf{Z}'\mathbf{Z} = \mathbf{I}$. We may begin with a matrix $\mathbf{A}$ that transforms $\mathbf{y}$ to the successive differences

$$z_t = y_{t+1} - y_t = \delta_{t+1} + \epsilon_{t+1} - \epsilon_t. \qquad (3.2)$$

This $\mathbf{A}$ satisfies $\mathbf{A1} = 0$, is $(T - 1) \times T$, and is of rank $T - 1$. There exists a nonsingular $(T - 1) \times (T - 1)$ matrix $\mathbf{B}$ such that $\mathbf{B}'\mathbf{AA}'\mathbf{B} = \mathbf{I}$. Hence we may choose $\mathbf{Z} = \mathbf{A}'\mathbf{B}$. From (3.2) we see that the component in $\operatorname{cov}(\mathbf{Ay})$ involving ρ is $\sigma^2\rho\mathbf{I}$. On the other hand this component is $\sigma^2\rho\mathbf{AWA}'$ so that $\mathbf{AWA}' = \mathbf{I}$. It follows then that $\mathbf{Z}'\mathbf{WZ} = \mathbf{B}'\mathbf{B}$. Because the eigenvalues of $\mathbf{B}'\mathbf{B}$ and $\mathbf{BB}' = (\mathbf{AA}')^{-1}$ are identical, the eigenvalues of $\mathbf{Z}'\mathbf{WZ}$ are obtainable as the inverse values of those of

$$\mathbf{AA}' = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 2 \end{bmatrix}$$

From Anderson (1971, Theorem 6.5.5) we find the latter to be

$$2(1 - \cos(\pi k/T)), \quad k = 1, \ldots, T - 1. \qquad (3.3)$$

The corresponding normalized eigenvectors are

$$\sqrt{(2/T)}\,(\sin(\pi k/T), \sin(2\pi k/T), \ldots, \sin((T - 1)\pi k/T)), \; k = 1, \ldots, T - 1. \qquad (3.4)$$

Together with (2.3) and (2.4), (3.3) now shows, for the L in (3.1), that

$$L \sim \frac{\sum_{k=1}^{T-1} \lambda_{kT}(1 + \rho\lambda_{kT})u_k^2}{\sum_{k=1}^{T-1} (1 + \rho\lambda_{kT})u_k^2}, \qquad (3.5)$$

where the u_k are iid $N(0, 1)$ variables and $\lambda_{kT}^{-1} = 2(1 - \cos\pi k/T)$, $k = 1, 2, \ldots, T - 1$. Similarly

$$F_g \sim \frac{\sum_{k=1}^{g} (1 + \rho\lambda_{kT})u_k^2/g}{\sum_{k=g+1}^{T-1} (1 + \rho\lambda_{kT})u_k^2/(T - g - 1)}. \qquad (3.6)$$

F_g can be computed from the formulas

$$F_g = \frac{S_g/g}{(\mathrm{SSE} - S_g)/(T - g - 1)}, \qquad (3.7)$$

and

$$S_g = \frac{2}{T} \sum_{k=1}^{g} \lambda_{kT} \left(\sum_{t=1}^{T-1} (y_{t+1} - y_t)\sin\frac{t\pi k}{T} \right)^2,$$

$$\mathrm{SSE} = \sum_{t=1}^{T-1} (y_t - \bar{y})^2. \qquad (3.8)$$

When one uses F_g as a test criterion, only tables of the F distribution are needed. In order to produce a table for the L criterion of (3.1) we have solved c_α from

$$\alpha = P(L/(T - 1) > c_\alpha) = P\left(\sum_{k=1}^{T-1} (\lambda_{kT}/(T - 1) - c_\alpha)u_k^2 > 0 \right). \qquad (3.9)$$

The numerical computations have been performed using Imhof's (1961) technique of inversion of the characteristic function. We refer the interested reader to this paper. (A minor departure from Imhof's procedure is that the integration intervals are *halved* at each step.) Except for final round-off errors the results have guaranteed accuracy better than .01 ($\geq$ Imhof's ϵ) in all calculations.

Table 1 shows critical points of the LMPI statistic. In Table 2 we give powers of the LMPI and the various LaMotte and McWhorter tests, and the power envelope (largest power attainable). The significance level is .05 and $T - 1 = 20$ and 50. We observe that there is no uniformly most powerful test among the LMPI and LaMotte and McWhorter tests. We can roughly say that the LMPI test is the best in the range of its power (0, .4) F_2 in the range (.4, .5), F_3 in (.5, .7), F_4 in (.7, .8), and

Table 1. Critical Points of the LMPI Statistic L/(T − 1)

T − 1	α .10	.05	.025	.01
10	.401	.504	.598	.707
20	.375	.485	.594	.730
30	.366	.478	.590	.736
40	.361	.474	.588	.739
50	.358	.471	.587	.740
60	.357	.470	.586	.741
80	.354	.468	.585	.742
100	.353	.467	.584	.742
∞	.347	.461	.584	.743

NOTE: The last line is adapted from Anderson and Darling (1952) with the kind permission of the authors and the Institute of Mathematical Statistics.

F_5 in (.8, .9). When the power exceeds .9 the number of observations becomes more important. These remarks are supported by the asymptotic results (compare the Pitman efficiency in Sec. 4).

4. ASYMPTOTIC DISTRIBUTIONS AND EFFICIENCY

We next derive the asymptotic distributions for the LMPI test statistic L.

Theorem 1. Under H_0: $\rho = 0$

$$L/T \xrightarrow{d} \pi^{-2} \sum_{k=1}^{\infty} k^{-2} u_k^2, \tag{4.1}$$

where $\xrightarrow{d}$ denotes the convergence in distribution and $u_1, u_2, \ldots$ are iid $N(0, 1)$ variables. Under H_1: $\rho > 0$

$$L/T^2 \xrightarrow{d} \pi^{-2} \frac{\sum_{k=1}^{\infty} k^{-4} u_k^2}{\sum_{k=1}^{\infty} k^{-2} u_k^2} \tag{4.2}$$

with $u_1, u_2, \ldots$ as above.

Note. The parameter ρ does not appear in (4.2). Hence asymptotically the power of the LMPI test (3.1) is independent of the alternative.

Proof. Assume first that $\rho = 0$, and let $u_1, u_2, \ldots$ be as in the theorem. Then by (3.5)

$$L/T \sim \frac{T^{-2} \sum_{k=1}^{T-1} \lambda_{kT} u_k^2}{T^{-1} \sum_{k=1}^{T-1} u_k^2}.$$

Table 2. Powers of the LMPI and the F_g Tests at the Level .05

ρ	LMPI	g 1	2	3	4	5	10	15	Power Envelope
				a. T − 1 = 20					
.05	**.250**	.243	.238	.214	.193	.176	.114	.076	.255
.10	**.365**	.350	.363	.344	.320	.296	.186	.104	.384
.15	.437	.416	**.448**	.437	.415	.388	.254	.134	.473
.20	.486	.460	**.509**	.505	.486	.461	.316	.163	.538
.25	.522	.492	.554	**.557**	.542	.519	.370	.192	.589
.30	.551	.517	.589	**.598**	.587	.566	.414	.218	.629
.35	.573	.536	.617	**.631**	.623	.605	.496	.244	.661
.40	.592	.552	.640	**.658**	.653	.637	.491	.268	.689
.45	.607	.566	.658	**.680**	.678	.665	.522	.290	.712
.50	.621	.577	.674	**.699**	**.699**	.688	.551	.311	.732
.60	.642	.594	.700	.730	**.734**	.726	.598	.350	.765
.80	.672	.619	.734	.771	**.781**	.778	.670	.416	.811
1.00	.692	.635	.757	.798	**.812**	**.812**	.720	.493	.841
1.50	.722	.658	.789	.837	.855	**.861**	.796	.558	.885
2.00	.739	.671	.807	.857	.878	**.886**	.838	.617	.909
2.50	.750	.679	.818	.870	.892	**.901**	.864	.659	.924
3.00	.754	.685	.825	.878	.901	**.911**	.882	.689	.934
4.00	.767	.692	.835	.888	.912	**.924**	.904	.731	.946
5.00	.774	.696	.840	.895	.919	**.931**	.918	.758	.953
				b. T − 1 = 50					
.01	**.301**	.289	.291	.276	.259	.243	.230	.186	.311
.02	.436	.411	**.442**	.436	.421	.403	.387	.326	.468
.03	.518	.482	.538	**.541**	.532	.517	.502	.438	.570
.04	.578	.531	.602	**.615**	.610	.599	.586	.525	.644
.05	.618	.566	.650	**.669**	**.669**	.661	.650	.595	.698
.10	.737	.659	.771	.807	.819	**.822**	.819	.790	.845
.15	.793	.700	.822	.863	.880	.887	**.888**	.875	.907
.20	.827	.725	.851	.894	.912	.920	**.924**	.918	.937
.25	.849	.741	.869	.912	.931	.940	**.944**	.943	.957
.30	.865	.752	.881	.924	.943	.952	.956	**.958**	.969
.35	.877	.761	.890	.933	.952	.960	.965	**.968**	.976

NOTE: The largest powers are bold faced for each ρ.

Because the denominator converges to 1 in probability, it is sufficient to show that the numerator converges to (4.1) in probability.

Let us recall that $\lambda_{kT} = (2(1 - \cos\pi k/T))^{-1} = (4 \sin^2\pi k/2T)^{-1}$. It follows from this that the difference $T^{-2}\lambda_{kT} - (\pi k)^{-2}$ is positive and increasing with k. Together with the Markov inequality this implies that for every $\epsilon > 0$

$$P\left[\sum_{k=1}^{T-1} | T^{-2}\lambda_{kT} - (\pi k)^{-2} | u_k^2 > \epsilon\right]$$

$$< \frac{T-1}{\epsilon}(T^{-2}\lambda_{T-1,T} - (\pi(T-1))^{-2}).$$

Since $\lambda_{T-1,T} \to \frac{1}{4}$, (4.1) follows. The result (4.2) is proved similarly.

Notice that the distribution of the right side of (4.1) is the same as the limiting distribution of the Cramer-von Mises goodness-of-fit test statistic (Anderson and Darling 1952, see Table 1.)

In the same way as Theorem 1 one can prove the next theorem.

Theorem 2. Under the sequence of alternatives $\rho_T = \rho T^{-2} + o(T^{-2})$

$$L/T \xrightarrow{d} \pi^{-2} \sum_{k=1}^{\infty} (k^{-2} + \rho\pi^{-2}k^{-4})u_k^2, \quad (4.3)$$

where $u_1, u_2, \ldots$ are iid $N(0, 1)$ variables.

The corresponding results for the LaMotte and McWhorter tests F_g are collected in the next theorem.

Theorem 3. Let F_g be as in (3.8). Then

(i) $gF_g \xrightarrow{d} \chi^2(g)$, under H_0

(ii) $gF_g/T \xrightarrow{d} \dfrac{\sum_{k=1}^{g} k^{-2}u_k^2}{\sum_{k=g+1}^{\infty} k^{-2}u_k^2}$, under H_1

(iii) $gF_g \xrightarrow{d} \sum_{k=1}^{g} (1 + \rho(\pi k)^{-2})u_k^2$, under the sequence $\rho T^{-2} + o(T^{-2})$.

Theorems 2 and 3 are the basis for calculating the Pitman asymptotic relative efficiencies. Because the asymptotic distributions are not the standard ones we have included in Appendix A.1 a short note on the Pitman efficiency. It follows from Theorem A.1 and Theorem 3 (i), (iii) that the Pitman efficiency of a LaMotte and McWhorter test relative to another such test is the square root of the ratio of the corresponding numbers ρ_g obtained from the equations

$$P\left[\sum_{k=1}^{g} (1 + \rho_g(\pi k)^{-2})u_k^2 > \chi_\alpha^2(g)\right] = \gamma_0 \quad (4.4)$$

with $\chi_\alpha^2(g)$ the upper α critical point of the $\chi^2(g)$ distribution, $g = 1, 2, \ldots$. This efficiency measure depends on α and γ_0. Table 3, in which some numerical values are given, suggests that in general the departure from optimality is not severe when $3 \le g \le 8$, that the dependence on α for fixed γ_0 is not strong, and that the dependence on γ_0 for fixed α is more important. As for the efficiency of the LMPI test with respect to the LaMotte and McWhorter tests, its computation seems to be a formidable task, since we must solve ρ_L (Theorems A.1, 1 and 2) from

$$P\left[\sum_{k=1}^{\infty} ((\pi k)^{-2} + \rho_L(\pi k)^{-4})u_k^2 > c_\alpha\right] = \gamma_0, \quad (4.5)$$

where c_α is the asymptotic critical point of the LMPI test. Therefore we have, instead, solved ρ_L from the equation

$$P\left[a\sum_{k=1}^{10} ((\pi k)^{-2} + \rho_L(\pi k)^{-4})x_k + b > c_\alpha\right] = \gamma_0, \quad (4.6)$$

where $x_1, \ldots, x_{10}$ are iid $\chi^2(\nu)$ variables. The constants a, b, ν are found by fitting the first three moments of the random variables in (4.6) to those of the variables in (4.5). The accuracy of this approximation is checked by comparing the probabilities (4.5) and (4.6) when $\rho_L = 0$ and $\alpha = .01, .05, .10$ (.10), .90, .95, .99. The approximation differed from the correct probabilitiy by not more than

Table 3. The Pitman Asymptotic Relative Efficiencies of the LMPI and F_g Tests

	$\alpha = .05, \gamma_0 = .6$		$\alpha = .05, \gamma_0 = .8$	
g	ρ_g	$(\rho_3/\rho_g)^{1/2}$	ρ_g	$(\rho_5/\rho_g)^{1/2}$
1	128.00	.807	580.8	.561
2	87.33	.977	234.6	.883
3	83.33	1.000	194.5	.970
4	84.35	.994	184.8	.995
5	86.73	.980	183.1	1.000
6	89.62	.964	185.0	.995
7	92.55	.949	188.2	.986
8	95.57	.934	192.0	.977
9	98.61	.919	196.1	.966
10	101.58	.906	200.1	.957
L	95.30	.935	259.0	.841
	$\alpha = .01, \gamma_0 = .8$		$\alpha = .0001, \gamma_0 = .8$	
g	ρ_g	$(\rho_5/\rho_g)^{1/2}$	ρ_g	$(\rho_6/\rho_g)^{1/2}$
1	1010.4	.520	2317.7	.469
2	372.9	.856	768.8	.663
3	299.1	.956	589.1	.930
4	279.3	.989	535.7	.975
5	273.4	1.000	515.3	.994
6	274.2	.999	509.4	1.000
7	277.2	.993	509.5	1.000
8	281.4	.986	512.8	.997
L	445.1	.784		

NOTE: For every combination of α and γ_0 the first column contains the solution of (4.4) and (4.5) and the second the efficiencies with respect to the optimal choice.

four digits in the fourth decimal. Numerical values from (4.6) are given in Table 3.

Although the Pitman efficiency does not seem to depend strongly on the usual significance levels for fixed γ_0, the limiting Pitman efficiency, as $\alpha \to 0$, gives quite a different picture of the situation. Keeping γ_0 fixed we deduce from (4.4) that

$$\lim_{\alpha\to 0} \frac{\chi_\alpha^2(g)}{\rho_g(\alpha, \gamma_0)} = d_g$$

and from (4.5) that

$$\lim_{\alpha\to 0} \frac{c_\alpha}{\rho_L(\alpha, \gamma_0)} = d_L,$$

where d_g and d_L satisfy

$$P\left[\sum_{k=1}^{g} (\pi k)^{-2} u_k^2 > d_g\right] = \gamma_0$$

and

$$P\left[\sum_{k=1}^{\infty} (\pi k)^{-4} u_k^2 > d_L\right] = \gamma_0.$$

Using Lemma A.3.1(i) we obtain that the limiting Pitman efficiency of F_g with respect to $F_{g'}$ is $[d_g/d_{g'}]^{1/2}$ and that of the LMPI test with respect to F_g is $\pi[d_L/d_g]^{1/2}$. Because d_g is an increasing function of g, there is no optimum choice of g. We also note that the LMPI test is more efficient than F_1 in the limit.

This result suggests (as pointed out by an associate editor) also considering tests where $g = g_T \to \infty$ as $T \to \infty$. However, it turns out that while there exists an asymptotic distribution for F_{g_T} (appropriately normed) for certain sequences of alternatives ρ_T the appropriate rate of convergence in $\rho_T \to 0$ is of lower order than T^{-2}. Consequently, the Pitman efficiency of F_{g_T} relative to F_g or L is zero.

Because the Pitman efficiency depends on α and γ_0, Gregory (1980) has suggested an efficiency measure that is the limiting ratio of the asymptotic powers themselves when $\alpha \to 0$. Contrary to the Pitman efficiency this does not have the interpretation of being a limiting ratio of numbers of observations. The following theorem shows that in the sense of this measure the LMPI test is optimal among the tests considered here.

Theorem 4. If $\chi_\alpha^2(g)$ and c_α are as before then

$$\lim_{\alpha\to 0} \frac{P\left[\sum_{k=1}^{\infty} ((\pi k)^{-2} + \rho(\pi k)^{-4}) u_k^2 > c_\alpha\right]}{P\left[\sum_{k=1}^{g} (1 + \rho(\pi k)^{-2}) u_k^2 > \chi_\alpha^2(g)\right]}$$

$$= \infty \quad \text{for } g > 1,$$

$$= c > 1 \quad \text{for } g = 1.$$

Proof. From a theorem of Zolotarev (1961) we obtain that

$$\lim_{\alpha\to 0} \frac{P\left[\sum_{k=1}^{\infty} ((\pi k)^{-2} + \rho(\pi k)^{-4}) u_k^2 > c_\alpha\right]}{P[(\pi^{-2} + \rho\pi^{-4}) u_1^2 > c_\alpha]} = C_L$$

and

$$\lim_{\alpha\to 0} \frac{P\left[\sum_{k=1}^{g} (1 + \rho(\pi k)^{-2}) u_k^2 > \chi_\alpha^2(g)\right]}{P[(1 + \rho\pi^{-2}) u_1^2 > \chi_\alpha^2(g)]} = C_g$$

exist. Using l'Hospital's rule and Lemma A.3.1 we further obtain that

$$\lim_{\alpha\to 0} \frac{P[(\pi^{-2} + \rho\pi^{-4}) u_1^2 > c_\alpha]}{P[(1 + \rho\pi^{-2}) u_1^2 > \chi_\alpha^2(g)]}$$

$$= \lim_{\alpha\to 0} \left(\frac{[c_\alpha/(\pi^{-2} + \rho\pi^{-4})]^{-1/2}}{[\chi_\alpha^2(g)/(1 + \rho\pi^{-2})]^{-1/2}}\right)$$

$$\times \left(\frac{\exp[-c_\alpha/(2(\pi^{-2} + \rho\pi^{-4}))]}{\exp[-\chi_\alpha^2(g)/(2(1 + \rho\pi^{-2}))]}\right)$$

$$= \infty, \quad \text{if } g > 1$$

$$= \prod_{k=2}^{\infty} (1 - k^{-2})^{(1-\theta)/2}, \quad \text{if } g = 1,$$

where $\theta = \rho/(\pi^2 + \rho)$. The case $g > 1$ is thus proved. Because

$$C_L = \prod_{k=2}^{\infty} \left(1 - \frac{(\pi k)^{-2} + \rho(\pi k)^{-4}}{\pi^{-2} + \rho\pi^{-4}}\right)^{-1/2}$$

$$= \prod_{k=2}^{\infty} (1 - k^{-2})^{-1/2} (1 + \theta k^{-2})^{-1/2},$$

the limit in the case $g = 1$ has the value

$$\prod_{k=2}^{\infty} (1 - k^{-2})^{-\theta/2} (1 + \theta k^{-2})^{-1/2}$$

$$= \left(\frac{\pi 2^\theta (1 + \theta)\sqrt{\theta}}{\sinh \pi\sqrt{\theta}}\right)^{1/2}. \quad (4.7)$$

It is easy to see that the general factor in the product (4.7) is an increasing function of θ and for $\theta = 0$ it equals one. Hence the proof is complete.

From the proof of the theorem it is clear that the choice $g = 1$ is the most efficient in the Gregory sense among the LaMotte and McWhorter tests. However, we cannot infer that at small significance levels the LMPI or F_1 test would have any optimum properties in the Pitman sense, because if α converges to zero then ρ increases without bound for fixed γ_0 in (4.4) and (4.5).

Another common efficiency measure is due to Bahadur. This involves calculation of the so-called exact slope of the test statistic (Bahadur 1971). In departure from the general theory there exists in our problem no nonrandom exact slope. In Appendix A.2 we extend the notion of an exact slope to include slopes that are nondegenerate random variables. Therefore we have added a note also on the Bahadur efficiency in Appendix A.2. From Theorem A.3.1 and the proof Theorem 1 we obtain that under H_0

$$\lim_{T\to\infty} \frac{\log P(T^{-2}L > b_T)}{T} = \frac{1}{2}\log(1 - \pi^2 b),$$

if $b_T \to b$, $0 < b < \pi^{-2}$. Then Theorem 2 and (A.2.3) imply that the LMPI test has the efficacy $-\frac{1}{2}\log(1 - \pi^2 b_L)$ where b_L is determined by

$$P\left[\pi^{-2}\frac{\sum_{k=1}^{\infty} k^{-4}u_k^2}{\sum_{k=1}^{\infty} k^{-2}u_k^2} > b_L\right] = \gamma_0. \qquad (4.8)$$

Similarly the efficacies for the LaMotte and McWhorter tests are $-\frac{1}{2}\log(1 - b_g)$, $g = 1, 2, \ldots$, where each b_g is the solution of

$$P\left[\frac{\sum_{k=1}^{g} k^{-2}u_k^2}{\sum_{k=1}^{\infty} k^{-2}u_k^2} > b_g\right] = \gamma_0. \qquad (4.9)$$

The asymptotic relative efficiencies are the ratios of these efficacies. We have made no (exact or approximate) numerical calculations. However, we see from (4.8) and (4.9) that the LMPI test is more efficient than the LaMotte and McWhorter test F_1. From (4.9) we find that with g b_g increases to 1 and consequently $-\frac{1}{2}\log(1 - b_g)$ increases without bound. Hence the Bahadur efficiency and the limiting Pitman efficiency, although not the same, behave similarly and do not offer a practical means to choose among the tests F_g.

5. CONCLUSIONS

We have compared the LMPI test and the tests suggested by LaMotte and McWhorter (1978) on the basis of exact powers and three different asymptotic efficiency measures due to Pitman, Bahadur, and Gregory, the last one dating from 1980. Only the Pitman efficiency, which depends on the significance level α and the required power γ_0, seems to agree reasonably with the exact powers. The Bahadur efficiency and also the limiting Pitman efficiency (as $\alpha \to 0$) provide reliable comparisons in situations with finite numbers of observations only when extremely small significance levels are used. These conclusions conform with findings made in other contexts (Groeneboom and Oosterhoff 1981). The Gregory efficiency seems to reflect the fact that one can make the range of the parameter values where the LMPI test is superior to the LaMotte and McWhorter tests as large as one may wish by sufficiently decreasing the significance level. Contrary to the Pitman and the Bahadur efficiencies this limiting process is accomplished without any reference to the power attained.

Primarily on the basis of the exact powers and the Pitman efficiency we make the following rough recommendations for choosing a test in practice. If the alternatives very near the null hypothesis are important then use the LMPI test. If attention is focused on alternatives where the best test has power between .6 and .8, then choose one of the LaMotte and McWhorter tests F_3, F_4, or F_5. If the number of observations is large and very remote alternatives are of concern, then use F_g for some $g > 5$.

The application of the LaMotte and McWhorter tests almost always needs a computer. On the other hand the LMPI test needs a separate table or a computer program for calculating the significance probability. The usual tables of the F distribution suffice for the LaMotte and McWhorter tests.

APPENDIX

A.1 Pitman Efficiency

The usual method of computing the Pitman (asymptotic relative) efficiency of two tests is inapplicable in our problems, because the limiting distributions of the competing test statistics are of different types. As a rule the evaluation of Pitman efficiencies takes for granted that through appropriate norming the asymptotic distributions are normal or χ^2 distributions. However, it turns out that the Pitman efficiency can be computed directly from the asymptotic powers and without reference to the method of finding the latter.

Let θ be a real parameter indexing a family of distributions and consider the hypotheses

$$H_0: \theta = \theta_0, \quad H_1: \theta > \theta_0.$$

We want to compare, at the level of significance α, two tests based on statistics T_{1n} and T_{2n}, having power functions $\gamma_{1n}(\theta)$ and $\gamma_{2n}(\theta)$, respectively, where n is the number of observations. Let us denote

$$\theta_n = \theta_0 + (\delta_n/n^r),$$

where $\delta_n > 0$, $r > 0$, and $\delta_n \to \delta > 0$. We assume that the limits

$$\lim \gamma_{in}(\theta_n) = \gamma_i(\delta), \quad i = 1, 2$$

(depending only on i and δ) exist for all such sequences. If there exists a sequence of natural numbers $N_n \to \infty$ such that

$$\lim \gamma_{2N_n}(\theta_n) = \gamma_0 = \gamma_1(\delta)$$

then the limit

$$\lim \frac{1/n}{1/N_n} = e_P = e_P(\gamma_0, \alpha)$$

is called the Pitman efficiency of T_{1n} with respect to T_{2n}.

The efficiency measure e_P is approximately the ratio of the number of observations needed to obtain the power γ_0 at the level α under the same sequence of alternatives. Generally e_P depends on γ_0 and α.

Theorem A.1. If δ_1 and δ_2 are such numbers that $\gamma_1(\delta_1) = \gamma_2(\delta_2) = \gamma_0$ and γ_1, γ_2 are strictly increasing functions then

$$e_P = [\delta_2/\delta_1]^{1/r}.$$

Proof. This follows from a computation that may be found in Rao (1973, p. 469). Let $\theta_n = \theta_0 + \xi_n/n^r$, where $\xi_n \to \delta_1$. Then

$$\gamma_0 = \gamma_2(\delta_2) = \lim \gamma_{2N_n}(\theta_n)$$

$$= \lim \gamma_{2N_n}\left(\theta_0 + \left[\frac{N_n}{n}\right]^r \cdot \frac{\xi_n}{\xi_{N_n}} \cdot \frac{\xi_{N_n}}{N_n{}^r}\right)$$

$$= \gamma_2(e_P{}^r\delta_1)$$

for $N_n/n \to e_P$ and $\xi_{N_n} \to \delta_1$. Hence $\delta_2 = e_P{}^r\delta_1$.

As a corollary it is easy to see that if $\gamma_i(\delta) = h(c_i\delta + b_\alpha)$, where c_i is positive and h strictly increasing and only b_α depends on α, we have $e_P = (c_1/c_2)^{1/r}$, this quantity being independent of α and γ_0. With h the standard normal distribution function this example is the one occurring most frequently in applications. In our problems it does not. Neither does the independence of α and γ_0 obtain.

A.2 Bahadur Efficiency

Let us now consider the testing problem of Appendix A.1 from the point of view of the Bahadur efficiency. The idea of the Bahadur efficiency is to keep the alternative fixed and let the size of the test decrease at the rate needed to obtain the required power. This is formalized in the following.

Definition. Let T_{in}, $i = 1, 2$ be two sequences of test statistics such that large values of each lead to rejection of the hypothesis $\theta = \theta_0$ in favor of the alternative $\theta > \theta_0$. Let $\gamma_0 \in (0, 1)$ be given and suppose that $\alpha_\nu \to 0$ is a sequence of significance levels and $n_{1\nu}$, $n_{2\nu}$ two sequences of numbers of observations (depending on θ) such that

$$\lim_{\nu\to\infty} P_\theta(T_{in_{i\nu}} > c_{i\nu}) \to \gamma_0, \quad i = 1, 2,$$

where the $c_{i\nu}$ are the critical points at the level α_ν. Then the limit

$$\lim_{\nu\to\infty} \frac{1/n_{1\nu}}{1/n_{2\nu}} = e_B = e_B(\theta, \gamma_0),$$

if it exists and is independent of α_ν, is called the Bahadur efficiency of T_{1n} with respect to T_{2n}.

Bahadur has shown that this definiton is applicable in a large class of testing problems (see Bahadur 1971, e.g.). In fact, the concept of efficiency is only taken up in his work when the class in question has already been defined and then in order to elucidate the notion of an exact slope. The problems considered in this article, however, fall outside the class considered by Bahadur.

We shall suppose that, for some r,

$$T_{in}/n^r \to B_i(\theta), \quad i = 1, 2 \qquad \text{(A.2.1)}$$

in distribution. Bahadur (1971, Theorem 7.2) assumes that the $B_i(\theta)$ are constants (i.e., nonrandom) while we also admit nondegenerate random variables. (The reader may note that T_{in} occurs only through T_{in}/n^r so that the choice of (T_{in}, r) is just a matter of convenience. Bahadur (1971) always has $r = \frac{1}{2}$.)

From (A.2.1) and our definiton it is deduced that if $B_i(\theta)$ is nonrandom, then

$$(c_{i\nu}/n_{i\nu}{}^r) \to B_i(\theta).$$

If $B_i(\theta)$ is nondegenerate and if there exists a unique number $b_i(\theta)$ such that $P_\theta(B_i(\theta) > b_i(\theta)) = \gamma_0$, then

$$(c_{i\nu}/n_{i\nu}{}^r) \to b_i(\theta).$$

Analogously to the theory of Bahadur it is true in our problems that

$$\lim_{\nu\to\infty} \frac{\log P_{\theta_0}(n_{i\nu}{}^{-r}\, T_{in_{i\nu}} > b_\nu)}{n_{i\nu}} = -g_i(b) \qquad \text{(A.2.2)}$$

for some (continuous) g_i and any sequence b_ν converging to b. We then have that

$$\lim_{\nu\to\infty} \frac{\log\alpha_\nu}{n_{i\nu}} = \lim_{\nu\to\infty} \frac{\log P_{\theta_0}(T_{in_{i\nu}}/n_{i\nu}{}^r > c_{i\nu}/n_{i\nu}{}^r)}{n_{i\nu}}$$

$$= -g_i(b_i(\theta)), \quad i = 1, 2. \qquad \text{(A.2.3)}$$

It follows that the Bahadur efficiency of T_{1n} with respect to T_{2n} is

$$e_B = [g_1(b_1(\theta))/g_2(b_2(\theta))].$$

A.3 Tail Probabilities for Linear Combinations of χ^2 Variables and Their Ratios

Let $X_1, X_2, \ldots$ be a sequence of independent χ^2 variables with $r_1, r_2, \ldots$ degrees of freedom, respectively, and let $\lambda_1, \lambda_2, \ldots$ be a strictly decreasing sequence of positive numbers. We denote by G_m the $\chi^2(m)$ distribution function. The following lemma is a consequence of a theorem by Zolotarev (1961) (see also Hoeffding 1964, and Gregory 1980).

Lemma A.3.1. Let F be a distribution function of the random variable $\sum \lambda_k X_k$ (it exists if $\sum r_k\lambda_k$ converges), and further let $x_\alpha = F^{-1}(1 - \alpha)$ and $y_\alpha = G_m{}^{-1}(1 - \alpha)$. Then

(i) $\lim_{\alpha\to 0} \dfrac{x_\alpha}{y_\alpha} = \lambda_1$

(ii) $\lim_{\alpha\to 0} y_\alpha - \dfrac{x_\alpha}{\lambda_1}$

$$= -\infty, \qquad \text{if } m < r_1$$

$$= \log\left[\prod_{k=2}^{\infty}\left(1 - \frac{\lambda_k}{\lambda_1}\right)^{r_k}\right], \qquad \text{if } m = r_1$$

$$= \infty, \qquad \text{if } m > r_1$$

The next result will be published elsewhere and is given here without proof.

Theorem A.3.1. Let $\lambda_{1n}, \ldots, \lambda_{nn}$, $n = 1, 2, \ldots$ be sequences of numbers such that

(1) $\max_{1 \le k \le n} | \lambda_{kn} - \lambda_k | \to 0$, when $n \to \infty$,

(2) $\lambda_1, \lambda_2, \ldots$ is a strictly decreasing sequence of numbers converging to 0.

Then

$$\lim_{n\to\infty} \frac{1}{N_n} \log P \left[\frac{\sum_{k=1}^{n} \lambda_{kn} X_k}{\sum_{k=1}^{n} X_k} > b_n \right] = \frac{1}{2} \log \left(1 - \frac{b}{\lambda_1} \right), \quad 0 < b < \lambda_1,$$

where $N_n = r_1 + \cdots + r_n$, and $b_n \to b$.

[*Received January 1982. Revised January 1983.*]

REFERENCES

ANDERSON, T.W. (1971), *The Statistical Analysis of Time Series*, New York: John Wiley.

ANDERSON, T.W., and DARLING, D.A. (1952), "Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Processes," *The Annals of Mathematical Statistics*, 23, 193–212.

BAHADUR, R.R. (1971), *Some Limit Theorems in Statistics*, Philadelphia: Society for Industrial and Applied Mathematics.

BOX, G.E.P., and JENKINS, G.M. (1970), *Time Series Analysis, Forecasting and Control*, San Francisco: Holden-Day.

DURBIN, J., and WATSON, G.S. (1971), "Testing for Serial Correlation in Least Squares Regression. III," *Biometrika*, 58, 1–19.

GREGORY, GAVIN G. (1980), "On Efficiency and Optimality of Quadratic Tests," *The Annals of Statistics*, 8, 116–131.

GROENEBOOM, P., and OOSTERHOFF, J. (1981), "Bahadur Efficiency and Small-Sample Efficiency," *International Statistical Review*, 49, 127–141.

HARRISON, P.J., and STEVENS, C.F. (1976), "Bayesian Forecasting," *Journal of the Royal Statistical Society*, Ser. B, 38, 205–247.

HOEFFDING, W. (1964), "On a Theorem of V.M. Zolotarev," *Theory of Probability and Its Applications*, 9, 89–91.

IMHOF, J.P. (1961), "Computing the Distribution of Quadratic Forms in Normal Variables," *Biometrika*, 48, 419–426.

KING, M.L. (1980), "Robust Tests for Spherical Symmetry and Their Applications to Least Squares Regression," *The Annals of Statistics*, 8, 1265–1271.

LaMOTTE, LYNN ROY, and McWHORTER, ARCHER, JR. (1978), "An Exact Test for the Presence of Random Walk Coefficients in a Linear Regression Model," *Journal of the American Statistical Association*, 73, 816–820.

RAO, C.R. (1973), *Linear Statistical Inference and Its Applications* (2nd ed.), New York: John Wiley.

ROSENBERG, BARR (1973), "A Survey of Stochastic Parameter Regression," *Annals of Economic and Social Measurements*, 2, 381–397.

STYAN, GEORGE P.H. (1973), "Hadamard Products and Multivariate Statistical Analysis," *Linear Algebra and Its Applications*, 6, 217–240.

ZOLOTAREV, V.M. (1961), "Concerning a Certain Probability Problem," *Theory of Probability and Its Applications*, 6, 201–204.