A NU TEST FOR SERIAL CORRELATION OF REGRESSION MODEL RESIDUALS

Ray L. Marr and C. P. Quesenberry

A test is given for testing for lag h autocorrelation of residuals from a full-rank normal errors regression model. This test is based on the normalized uniform (NU) residuals obtained from conditional probability integral transformations (CPIT). The exact distribution theory of these values permits the evaluation of the significance level for this test for any regression model. The test can also be applied to sets of NU residuals obtained from different regression regimes such as for data sequences obtained from measurements from a manufacturing process subject to an automated compensating action. Some limited power comparisons are made with the Durbin-Watson (D-W) statistic for a one-sample simple linear regression model, and for this case the new NU test has power somewhat less than the exact D-W statistic, but much better than the D-W bounds test. Percentage points are given for tests for lags 1 through 6 and formulas are given that will treat all cases. A technique is suggested for distinguishing test significance due to one form of model mis-specification from significance due to serial correlation.

KEY WORDS: Regression residuals, serial correlation, CPIT transformations.

1. INTRODUCTION

Measurements taken on parts from a production process are usually ordered in time by the order of production. In many machining processes, measured variables can be related by a linear regression model to other variables, including the production sequence number. The regression model may be useful in controlling the process so as to produce parts with small mean squared errors about target value (cf. Quesenberry (1988)). The usual assumptions for a full-rank, normal-errors, regression model include the assumption that the true errors ("true residuals") are independent. However, with this type of time-ordered data it is desirable to test the assumption of independence against an alternative of serial correlation among the true errors.

The usual type of tool-wear process data has special features that complicate testing for autocorrelation: because of either manual or automated compensations to the tool during production, the dataset is not generated from a single regression model, but rather from a sequence of regression models with possibly different regression parameters and error variances. Then we require a test for autocorrelation of random errors for a specified lag h using all of the data. We shall develop a test for this type of problem in the present paper.

Although we are interested in tests for serial correlation in the more complex setting described above, we shall initially consider the simpler set-up based on a single model and define the following notation:

$y_j = (y_1, y_2, \cdots, y_j)'$ denotes a vector of dependent random variables;
for $j = 1, 2, \cdots, N$;

X_N denotes an $N \times m$ matrix of full rank;

x_j' denotes the j^{th} row of X_N for $j = 1, 2, \cdots, N$;

X_j denotes the first j rows of X_N for $j = 1, 2, \cdots, N$; (1.1)

$\beta = (\beta_1, \beta_2, \cdots, \beta_m)'$ is a vector of regression coefficients;

$e = (e_1, e_2, \cdots, e_N)'$

denotes the vector of true errors ("true residuals"), the deviation of the actual y_N from its expected value $X_N \beta$ in the model

$$y_N = X_N \beta + e; \rho_h = \text{correlation } (e_j, e_j + h), \text{ for } j = 1, 2, \cdots, N - h.$$

We are interested in testing hypotheses of the forms

(1)	(2)	(3)
H_0: $\{e_j\}_{j=1}^N$ independent	H_0: $\{e_j\}_{j=1}^N$ independent	H_0: $\{e_j\}_{j=1}^N$ independent
H_a: $\rho_h > 0$	H_a: $\rho_h < 0$	H_a: $\rho_h \neq 0$

A number of tests have been proposed for these testing problems for the case h = 1 (lag 1 autocorrelation). The best known and most widely used of these tests is that of Durbin and Watson (1950, 1951, 1971), the D-W test. This test is computed from the least squares residuals vector

$$\hat{e} = (\hat{e}_1, \hat{e}_2, \cdots, \hat{e}_N)' = \left[I_N - X_N (X_N' X_N)^{-1} X_N'\right] y \tag{1.3}$$

and the D-W test statistic, which is virtually equivalent to the von Neumann ratio computed from the least squares residuals, is defined by

$$d = \frac{\sum_{j=2}^{N} (\hat{e}_j - \hat{e}_{j-1})^2}{\sum_{j=1}^{N} \hat{e}_j^2}. \tag{1.4}$$

The distribution of d depends upon the design matrix X_N. However, Durbin and Watson (1951) gave upper and lower bounds for signficance points for the distributon of d under the null hypothesis of $\rho_1 = 0$ for $\alpha = 0.01$, 0.025, and 0.05 for up to five explanatory variables besides the intercept, so that conservative tests can be made using these points. These tables of bounds are now given in many widely available books such as Theil (1971) and Morrison (1983) and were reported to greater precision in Koerts

and Abrahamse (1969). Durbin and Watson also gave techniques for making approximately exact level tests by first making a bounds test and then, when the bounds test is inconclusive, making an appropriate test based upon an approximating distribution. We shall give results in this work that compare the power of the test proposed here with the power of the D-W test for some cases where the D-W is also appropriate.

Theil (1965, 1968) and Abrahamse and Koerts (1971) gave tests for serial correlation formed by computing essentially the von Neumann ratio from the BLUS residuals posed by Theil. The advantage of forming a test from the BLUS residuals is that, under the null hypothesis, the BLUS residuals (and any statistic computed from them) have a distribution that does not depend upon the X_N design matrix. However, we shall not here consider tests computed from BLUS residuals for the following reason. We find it difficult to interpret values of statistics computed from BLUS residuals in terms of values of ρ_h because the BLUS residuals are not individually identified with the regression data points (y_j, x_j') for $j = 1, 2, \cdots, N$.

A common problem in testing for autocorrelation is the appearance of a large spurious autocorrelation when an ill-fitting design matrix is specified. This occurs because the (deterministic) error in fitting the wrong model is often a continuous function. Hence consecutive deterministic errors tend to track one another. Since the autocorrelation test statistics usually fail to distinguish this deterministic "error autocorrelation" from the random error autocorrelation, the experimenter may be left with the disturbing question of whether the detected "autocorrelation" in the data is real or just the result of mis-specifying the model. After we have developed the NU test proposed here, we shall point out how this test can sometimes be used to decide when a significant autocorrelation test is actually due to mis-specifying the form of the regression function.

2. THE TRANSFORMATIONS AND NU RESIDUALS

Under the null hypothesis of independent errors in the models of (1.1), we obtain from results of O'Reilly and Quesenberry (1973) (see also Quesenberry (1986), pp. 257-259) the following results. Put

$$b_j = (X_j' X_j)^{-1} X_j' y_j \text{ and}$$

$$s_j^2 = y_j' \left[I - X_j (X_j' X_j)^{-1} X_j' \right] y_j \text{ for } j = m + 1, \cdots, N - 1;$$

$$B_j = \frac{\sqrt{j - m - 1}\,(y_j - x_j' b_{j-1})}{s_{j-1}\sqrt{1 + x_j'(X_{j-1}' X_{j-1})^{-1} x_j}} \text{ and} \tag{2.1}$$

$$u_{j-m-1} = G_{j-m-1}(B_j) \text{ for } j = m + 2, \cdots, N, \tag{2.2}$$

where $G_k(\cdot)$ is the cdf of Student's t statistic with k degrees of freedom.

Then $u = (u_1, u_2, \cdots, u_{N-m-1})'$ is a vector of iid uniform random variables on the unit interval (0,1), called uniform residuals. If we write n for the number of components in u from an uninterrupted sequence of N observations, then

$$n = \max (0, N - m - 1) . \tag{2.3}$$

Note that n is 0 unless the number of consecutive observations exceeds the total number of estimated parameters (σ^2 and the m components of β). Applying the standard normal inverse cdf Φ^{-1} componentwise to the vector u of iid uniform residuals, the vector $z = (z_1, z_2, \cdots, z_n)'$ of n iid N(0, 1) random variables is obtained, where

$$z_j = \Phi^{-1}(u_j) \text{ for } j = 1, 2, \cdots, n. \tag{2.4}$$

Observe that the factor $(y_j - x_j' b_{j-1})$ in the numerator of B_j is the j^{th} residual from a model fitted by least squares using the first $j - 1$ data points, and B_j is itself the Studentized version of this residual since the denominator is the usual root mean square error estimate of σ from this model.

The u_j's are called uniform residuals since they are obtained by transforming

these Studentized residuals by the corresponding distribution functions to obtain independent uniform random variables. By transforming the uniform residuals with the standard normal inverse distribution function, we obtain the Normalized Uniform ("NU") residuals $z_1, z_2, \cdots, z_n$ that are approximately in the scale of the standardized versions of the original model residuals, that is, of $\hat{e}_j/\sigma$. The basis for this assertion is that while the composition function $\Phi^{-1} \circ G_j(\cdot)$ for $j = 1, 2, \cdots, n$ is nonlinear, it is approximately linear due to the strong uniform convergence of the transforming Student's t-distribution functions to the normal distribution function (cf. O'Reilly and Quesenberry (1972)), see Figure 1. Also, by generating data from regression models with known autocorrelations among the residuals it can be seen that the residuals e (true), $\hat{e}$ (least squares) and z (NU) show similar patterns when plotted against the index. From these remarks it is evident that autocorrelation patterns among the true residuals $e_{m+2}, e_{m+3}, \cdots, e_N$ are reflected in similar autocorrelation patterns among the NU residuals $z_1, z_2, \cdots, z_n$. Thus we propose tests based on the NU residuals z because under the null hypothesis of independent residuals e the elements are independent standardized normal random variables. Thus any real-valued function of z will have a distribution that does not depend on β, σ^2, or the design matrix X_n.

3. THE NU TEST FOR SERIAL CORRELATION

In view of the discussion in the last section we now consider testing for autocorrelation among the NU residuals. For a sequence of NU residuals $z = (z_1, z_2, \cdots, z_n)'$, form the statistic

$$S_n^{(h)} = \sum_{k=1}^{n-h} z_k \, z_{k+h} \text{ for } n > h. \tag{3.1}$$

It will be shown in §4 that $S_n^{(h)}$ is asymptotically normal and has mean 0 and variance $\sigma_0^2(S_n^{(h)}) = n - h$ under the null hypothesis H_0: $\rho_h = 0$. If we have J independent sequences of NU residuals (with the same h^{th} order correlation ρ_h but with possibly different design matrices, parameters β, and variances σ^2), for the testing problems in (1.2)

we propose using the test statistic

$$S_{\underline{n}}^{(h)} = \sum_{j=1}^{J} S_{n_j}^{(h)}, \quad \text{where } \underline{n} = (n_1, n_2, \cdots, n_J)', \tag{3.2}$$

where the P-value for this "interrupted" case is calculated programmatically. If the desired P-value is not readily available, we propose using the test statistic $S_{n_+}^{(h)}$, which ignores the fact that observations may come from J different populations; since all the NU residuals are iid normally distributed under the null hypothesis, the distribution of this $S_{n_+}^{(h)}$ is exactly $F_{n_+}^{(h)}$, and the standardized value of this statistic can be compared with the tabled value if $h \leq 6$ and n_+ is small or with the approximated value if n_+ is large. The difference in the two statistics is that the $S_{n_+}^{(h)}$ incorporates additional cross-product terms that "bridge" the gaps between the populations at the cost of a little noise and power. For the three testing problems in (1.1) we reject H_0 for problems (1), (2), or (3) when the test statistic is too large, too small, or too far from 0, respectively.

Critical values of the standardized statistic

$$\check{S}_n^{(h)} = S_n^{(h)} / \sigma_0\left(S_n^{(h)}\right)$$

have been tabled for select values of n and h to facilitate testing. An estimate of ρ_h is provided by

$$r_{\underline{n}}^{(h)} = S_{\underline{n}}^{(h)} / \sigma_0^2\left(S_{\underline{n}}^{(h)}\right).$$

Although the NU test may be thought to be computationally intensive, it should be recognized that the B_j are easily computed from update formulas; the NU transformation can be implemented by the approximation of §5; and the cdf of the $\check{S}_n^{(h)}$ is readily obtained by any quadrature algorithm suitable for Fourier integrals. Moreover, to test for nonzero autocorrelation ρ_h for different lags h, the NU residuals need only be computed once.

By comparison, to apply the Durbin-Watson test, one must compute not just a triangular matrix of coefficients but the whole N × N matrix; tables for the exact cdf do

not exist and a computer algorithm must first calculate the eigenvalues of the matrix $C\left[I_N - X_N(X_N'X_N)^{-1}X_N'\right]C'$, where C is the $(N-1) \times N$ upper triangular matrix with diagonal 1 and superdiagonal -1, before the quadrature algorithm may be applied. Alternatively, one could simply use the bounds test, but with a great loss of power.

4. DISTRIBUTION THEORY

In the propositions below, $\{Z_k\}$ will denote a set of independent standard normal random variables. The statistic

$$S_n^{(h)} = \sum_{k=1}^{n-h} Z_k Z_{k+h}$$

based on n NU residuals $(n > h)$ has cdf $F_n^{(h)}$ and the corresponding standardized statistic has cdf $\check{F}_n^{(h)}$. If the lag h is not specified, it is understood to be 1.

Lemma. (i) $E\left[S_n^{(h)}\right] = 0$; (ii) $\mathrm{Var}\left[S_n^{(h)}\right] = n - h.$ (4.1)

Proof: (i) Immediate. (ii) Because $E\left[Z_a Z_b Z_c Z_d\right] = 0$ unless the four subscripts are identical or consist of two pairs of identical subscripts,

$$\mathrm{Var}\left[\sum_{k=1}^{n-h} Z_k Z_{k+h}\right] = E\left[\sum_{k=1}^{n-h}\sum_{j=1}^{n-h} Z_k Z_{k+h} Z_j Z_{j+h}\right] = \sum_{k=1}^{n-h} E\left[Z_k^2 Z_{k+h}^2\right] = n - h. \quad \square$$

Lag h Reduction Lemma. Let $S_n^{(h)} = \sum_{k=1}^{n-h} Z_k Z_{k+h}$ have cdf $F_n^{(h)}$ and let $\bigstar$ denote the convolution operator. Then $F_n^{(h)} = (F_{s+1}^{(1)})^{\bigstar r} \bigstar (F_s^{(1)})^{\bigstar(h-r)}$, where integers r and s are defined by the division algorithm $n = hs + r$, $0 \le r < h < n$. (4.2)

Proof: $S_n^{(h)} = \sum_{k=1}^{n-h} Z_k Z_{k+h} = \sum_{k=1}^{h} \sum_{j=1}^{(n-h-k)/h} Z_{k+(j-1)h} Z_{k+jh}$,

where the inner summation sums over all subscripts in the same class modulo h. $\square$

Note that the subscripts and superscripts in the formula above appear additive.

Let $Z = (Z_1, Z_2, \cdots, Z_n)'$ and define $S_n = S_n^{(1)} = \sum_{k=1}^{n-1} Z_k Z_{k+1} = Z' A Z$, where

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \cdots & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \cdots & 0 \\ 0 & \frac{1}{2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{2} \\ 0 & 0 & \cdots & \frac{1}{2} & 0 \end{bmatrix}. \tag{4.3}$$

Then the moment generating function of S_n is

$$M_{S_n}(t) = (2\pi)^{-n/2} \int_{\mathfrak{R}^n} \exp(x'Ax\, t - x'x/2)\, dx = \| I - 2At\|^{-1/2}. \tag{4.4}$$

Now $I - 2At$ is positive definite for t sufficiently small. Moreover, since A is symmetric, there exists an orthogonal matrix P such that $P'AP = \Lambda_n$, where Λ_n is the diagonal matrix of eigenvalues $\{\cos \frac{k\pi}{n+1}\}_{k=1}^{n}$ in descending numerical order. One choice of P has entries

$$p_{j,k} = \frac{2 \sin(\frac{\pi jk}{n+1})}{\sqrt{2n + 1 - \frac{\sin(\pi k \frac{2n+1}{2n+2})}{\sin(\frac{\pi k}{n+1})}}} \tag{4.5}$$

Thus $M_{S_n}(t) = \| I - 2At\|^{-1/2} = \left[\|P'\| \; \| I - 2tA \| \; \|P\| \right]^{-1/2}$

$$= \| I - 2tP'AP \|^{-1/2} = \| I - 2t\Lambda_n\|^{-1/2}, \tag{4.6}$$

which is the moment generating function of a linear combination of independent χ_1^2-distributed random variables. Since the k^{th} diagonal entry of Λ_n is given by

$$\lambda_{k,n} = \cos \frac{k\pi}{n+1} = -\cos \frac{(n+1-k)\pi}{n+1} = -\lambda_{n+1-k,n}, \tag{4.7}$$

$$S_n \overset{d}{=} \sum_{k=1}^{n} \cos \frac{k\pi}{n+1}\, \chi_1^2 \overset{d}{=} \sum_{k=1}^{n/2} \cos \frac{k\pi}{n+1}\, (\chi_1^2)^{*} \overset{d}{=} \Big(\sum_{k=1}^{n/2} \cos \frac{k\pi}{n+1}\, \chi_1^2\Big)^{*}, \tag{4.8}$$

the symmetrized sum of $\left[\frac{n}{2} \right]$ independent weighted χ_1^2-distributed random variables, where $[\,\cdot\,]$ denotes the greatest integer function.

More specifically,

$$M_{S_n}(t) = \| I - 2\Lambda_n t\|^{-1/2} = \prod_{k=1}^{n/2} \left(1 - 4t^2 \cos^2 \frac{k\pi}{n+1}\right)^{-1/2} \text{ for } |t| < \cos \frac{[n/2]\pi}{n+1}. \tag{4.9}$$

Since the characteristic function is given by $\phi_{S_n}(t) \equiv M_{S_n}(it) \equiv \| I - 2iAt\|^{-1/2}$, we

shall find alternate expressions for the determinant

$$p_n(t) = \| I - 2iAt \| = \begin{Vmatrix} 1 & -it & 0 & \cdots & 0 \\ -it & 1 & -it & \cdots & 0 \\ 0 & -it & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -it \\ 0 & 0 & \cdots & -it & 1 \end{Vmatrix} . \qquad (4.10)$$

For $n = 1$, $p_1(t) \equiv 1$; for $n = 2$, $p_2(t) \equiv 1+t^2$. More generally, expanding the determinant by principal minors, we get the recursion

$$p_{n+2}(t) = 1 \cdot p_{n+1}(t) + it \cdot \begin{Vmatrix} -it & -it & 0 & \cdots & 0 \\ 0 & 1 & -it & \cdots & 0 \\ 0 & -it & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -it \\ 0 & 0 & \cdots & -it & 1 \end{Vmatrix} = p_{n+1}(t) + t^2 \, p_n(t). \qquad (4.11)$$

Solving the difference equation in n for each fixed t and incorporating the boundary conditions for $n = 1, 2$ gives the result

$$p_n(t) = \tfrac{1}{2} \left\{ \left(\tfrac{1}{2} + \sqrt{\tfrac{1}{4} + t^2} \right)^{n+1} - \left(\tfrac{1}{2} - \sqrt{\tfrac{1}{4} + t^2} \right)^{n+1} \right\} / \sqrt{\tfrac{1}{4} + t^2}. \qquad (4.12)$$

Alternately, using the recursion $p_{n+2}(t) = p_{n+1}(t) + t^2 \, p_n(t)$, it follows by induction on n that $p_n(t) \equiv \sum_{k=0}^{n-1} \binom{n-k}{k} t^{2k}$ for all n. Since $\binom{n-k}{k} = 0$ for integers $k > n - k$, we may write

$$p_n(t) \equiv \sum_{k=0}^{n/2} \binom{n-k}{k} t^{2k} \text{ and } \phi_{S_n}(t) \equiv \left\{ \sum_{k=0}^{n/2} \binom{n-k}{k} t^{2k} \right\}^{-1/2} \text{ for all n.} \qquad (4.13)$$

The equivalence of (4.12) and (4.13) is seen by applying the binomial theorem. Hence

$$M_{S_n}(t) = \phi_{S_n}(it) \equiv \left\{ \sum_{k=0}^{n/2} \binom{n-k}{k} (-1)^k \, t^{2k} \right\}^{-1/2}$$

on the interval of convergence $|t| < \cos \frac{[n/2]\pi}{n+1}$.

Since S_n is a convolution of absolutely continuous distributions, S_n is absolutely continuous. Applying the Lévy inversion theorem and using symmetry,

$$F_n(x) - F_n(0) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{1 - \exp(-ixt)}{it} \phi_{S_n}(t)\, dt$$

$$= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \left\{ \frac{\sin(xt)}{t} + i\,\frac{\cos(xt) - 1}{t} \right\} \phi_{S_n}(t)\, dt$$

$$= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{\sin(xt)}{t} \phi_{S_n}(t)\, dt = \frac{1}{\pi} \int_{0}^{\infty} \frac{\sin(xt)}{t} \phi_{S_n}(t)\, dt.$$

Hence $F_n(x) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin(xt)}{t} \phi_{S_n}(t)\, dt.$ (4.14)

Notation Let $n_1, n_2, \cdots, n_J > h$. Then $F_{\underset{\sim}{n}}^{(h)}(x) = F_{n_1, \cdots, n_J}^{(h)}(x)$ will denote convolution of the distribution functions $F_{n_1} \star \cdots \star F_{n_J}$.

$\check{F}_{\underset{\sim}{n}}^{(h)}(x) = \check{F}_{n_1, \cdots, n_J}^{(h)}(x)$ will denote the cdf of the standardized convolution

$$\check{F}_{\underset{\sim}{n}}^{(h)}(x) = F_{\underset{\sim}{n}}^{(h)}\left(x\sqrt{n_1 + \cdots + n_J - Jh}\right). \tag{4.15}$$

Again, if the lag h is not specified, it is understood to be 1. Since the statistic $S_{\underset{\sim}{n}}^{(h)}$ is a weighted sum of symmetrized independent χ_1^2-distributed random variables, $S_{\underset{\sim}{n}}^{(h)}$ is symmetric about 0. Moreover, $S_{\underset{\sim}{n}}^{(h)}$ has finite (absolute) moments of all positive orders, being a sum of random variables with finite absolute moments of all positive orders. So, in particular, all odd-ordered moments are 0.

Note that $\mathrm{Var}\, S_{\underset{\sim}{n}}^{(h)} = \sum_{j=1}^{J} (n_j - h)$ equals the number of summands $Z_j Z_k$. This is true for any quadratic form $Z'AZ$ in which the diagonal entries are 0 and the non-diagonal entries are either 0 or $\frac{1}{2}$, provided A is chosen as a symmetric matrix.

Theorem: (Rate of convergence to normality) $\| \check{F}_{\underset{\sim}{n}}(x) - \Phi(x) \| \le \frac{128C}{3\pi^2} (n_+ - J)^{-1/2}$, where $n_+ - J = \sum_{j=1}^{J} (n_j - 1)$ is the variance of S_n, C is the constant in the Berry-Esséen theorem, and $\|\cdot\|$ denotes the sup norm over $x \in \Re$. (Van Beeck (1972) establishes a value $C \approx 0.7975$; Zahl (1966) establishes $C \approx 0.651$ under more restrictive conditions applicable in this case.)

Proof: By applying the Berry-Esséen theorem to the summands of

$$S_n \overset{d}{=} \sum_{j=1}^{J} \sum_{k=1}^{n_j/2} \cos \frac{k\pi}{n_{j+1}} (\chi_1^2)^*,$$

where the $\left(\chi_1^2\right)^{\cdot}$ represent independent symmetrized χ_1^2-distributed random variables,

$$\| \check{F}_n(x) - \Phi(x) \| \leq C \left(\text{Var } S_n \right)^{-3/2} \sum_{j=1}^{J} \sum_{k=1}^{\frac{1}{2}n_j} \cos^3 \frac{k\pi}{n_j+1} \, E \, | \, (\chi_1^2)^{\cdot} |^3 .$$

Now $\text{Var } S_n = n_+ - J$, where $n_+ = \sum_{j=1}^{J} n_j$ and J is the dimension of the vector n, and

$$E \, | \, (\chi_1^2)^{\cdot} |^3 = E \, | \, Z_1^2 - Z_2^2 \, |^3 = E \left[\, | \, Z_1 - Z_2 |^3 \cdot | \, Z_1 + Z_2 |^3 \, \right] = 8 \, E^2 | \, Z |^3$$

$$= 8 \, E^2 (\chi_1^2)^{3/2} = \tfrac{64}{\pi},$$

where the Z represent generic iid standard normal random variables. Let $[\![\cdot]\!]$ denote the greatest integer function; this will be suppressed from the upper limit of summation when there is no chance of confusion.

For any positive integer n, $\sum_{k=1}^{\frac{1}{2}n} \cos^3 \frac{k\pi}{n+1} = \sum_{k=1}^{n/2} \left(\frac{1}{4} \cos \frac{3k\pi}{n+1} + \frac{3}{4} \cos \frac{k\pi}{n+1} \right)$

$$= \frac{1}{4} \left\{ \frac{\cos \frac{3\pi [\![\frac{n}{2}]\!]}{n+1} - \cos \frac{3\pi [\![\frac{n+2}{2}]\!]}{n+1}}{2 \left(1 - \cos \frac{3\pi}{n+1}\right)} - \frac{1}{2} \right\} + \frac{3}{4} \left\{ \frac{\cos \frac{\pi [\![\frac{n}{2}]\!]}{n+1} - \cos \frac{\pi [\![\frac{n+2}{2}]\!]}{n+1}}{2 \left(1 - \cos \frac{\pi}{n+1}\right)} - \frac{1}{2} \right\},$$

which reduces to $\frac{1}{8} \left[- \frac{\sin \frac{3\pi}{n+1}}{1 - \cos \frac{3\pi}{n+1}} + 3 \frac{\sin \frac{\pi}{n+1}}{1 - \cos \frac{\pi}{n+1}} - 4 \right]$ if n is odd,

and to $\frac{1}{8} \left[- \frac{2 \sin \frac{3\pi}{2n+2}}{1 - \cos \frac{3\pi}{n+1}} + \frac{6 \sin \frac{\pi}{2n+2}}{1 - \cos \frac{\pi}{n+1}} - 4 \right]$ if n is even.

Since the sequences $\frac{1}{8}[\cdot]/(n-1)$ both converge to their upper bound $\frac{2}{3\pi}$ as n approaches infinity, we have $\frac{1}{8}[\cdot] \leq \frac{2}{3\pi} \, (n-1)$.

Therefore $\| \check{F}_{\underset{\sim}{n}}(x) - \Phi(x) \| \leq C \left(\text{Var } S_{\underset{\sim}{n}} \right)^{-3/2} \frac{2}{3\pi} \sum_{j=1}^{J} (n_j - 1) \cdot \frac{64}{\pi}$

$$= \frac{128C}{3\pi^2} \left(n_+ - J \right)^{-\frac{1}{2}}. \qquad \square$$

Corollary $\| \check{F}_{\underset{\sim}{n}}^{(h)}(x) - \Phi(x) \| \leq \frac{128C}{3\pi^2} / \sigma_0\left(S_{\underset{\sim}{n}}^{(h)}\right)$, where $\sigma_0\left(S_{\underset{\sim}{n}}^{(h)}\right) = \left(n_+ - Jh \right)^{1/2}$.

Proof: The proof follows immediately from the previous theorem after applying the Lag h Reduction Lemma to $S_{\underset{\sim}{n}}^{(h)}$. $\square$

In particular, $\| \check{F}_{\underset{\sim}{n}}^{(h)}(x) - \Phi(x) \| \leq 2.815 / \sigma_0\left(S_{\underset{\sim}{n}}^{(h)}\right) = 2.815 \left(n_+ - Jh \right)^{-\frac{1}{2}}$.

It should be noted that this rate of convergence holds irrespective of whether the maximum number of uninterrupted data values approaches infinity. It is likely that a more practical (faster) rate of convergence theorem can be established directly, as the rate in the Berry-Esséen theorem must hold for sums of arbitrary independent random variables and here we are using weighted sums of symmetrized χ_1^2-distributed random variables (equivalently, weighted sums of pairwise products of independent standard normals).

5. NUMERICAL COMPUTATION OF THE NU TRANSFORMATION

Under the null hypothesis of independence of the e_j's, large CPIT residuals will seldom occur and the NU transformation can be performed by applying an algorithm for Φ^{-1} to the cdf of Student's t. However for simulation studies where $\rho \neq 0$, large CPIT residuals are more common. Hence it is desirable to have an approximation to the NU transformation for extreme values, where the composition of available algorithms may be unsatisfactory.

For any fixed k, let $\nu(t) = \Phi^{-1}(G_k(t))$, where $G_k(t)$ is the cdf of Student's t-distribution with k degrees of freedom. Since k is fixed in this section, we temporarily suppress the dependence of $\nu(t)$ on k . Clearly $\nu(t) = -\nu(-t)$ since Φ and G_k are both symmetric cdfs.

For $t > 0$ let $Q_k(t) = 2G_k(t) - 1$ and $\text{erf}(t) = 2\,\Phi\,(t\sqrt{2}) - 1$; let inverf$(\cdot)$ be the functional inverse of erf$(\cdot)$. Then $\nu(t) = \sqrt{2}\ \text{sgn}(t)\ \text{inverf}\ Q_k(|t|)$, where sgn$(\cdot)$ is the signum function. This alternative form of $\nu(t)$ is numerically preferable if inverf is implemented via the algorithm of Strecok (1968) and $Q_k(t)$ is computed by the finite expansion given in Abramowitz and Stegun (1970), p. 948, §§ 26.7.3 - 26.7.4.

It can be shown that $\nu(t) \sim \sqrt{2k \ln t} \sim k \ln\left(1 + \frac{t^2}{k}\right)$ as $t \to \infty$. The second form was adopted to avoid definitional problems for $0 \leq t < 1$ and to more faithfully mirror

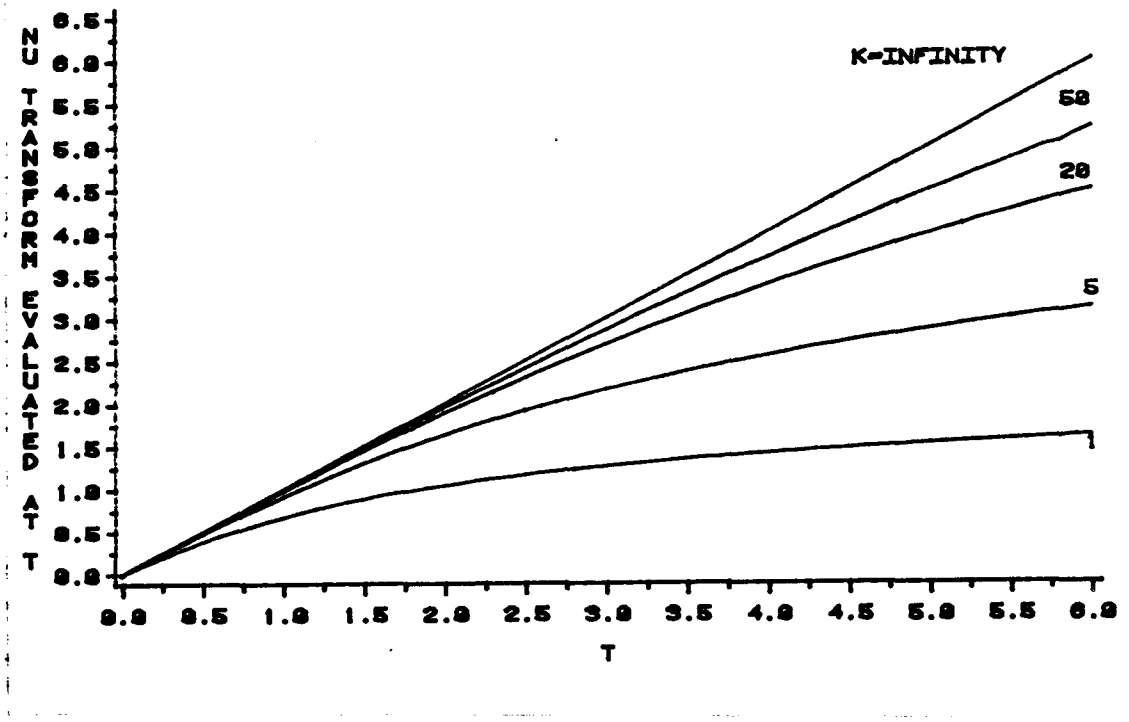

Figure 1: $\nu_k(t)$ vs. t for k=1,2,5,20,50,∞

the form of Student's t density. Moreover, empirically the second form appears to better approximate $\nu(t)$. Still better approximations to $\nu(t)$ can be derived from the asymptotic approximation

$$\nu^{(0)}(t) = \sqrt{k \ln\left(1+\frac{t^2}{k}\right)} \ .$$

$$\nu^{(1)}(t) = \sqrt{(k-1) \ln\left(1 + \frac{t^2}{k}\right) + 2 \ln t - 2 \ln C - 2 \ln \nu^{(0)}(t)} \ ,$$

$$\text{where } C = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k/2}\ \Gamma\left(\frac{k}{2}\right)} \ .$$

$$\nu^{(2)}(t) = -1 + \sqrt{1 + C + (k+1) \ln\left(1 + \frac{t^2}{k}\right) + 2\,\nu^{(1)}(t) + 2 \ln \frac{d\nu^{(1)}(t)}{dt}} \ .$$

$$\nu^{(3)}(t) = -\tfrac{1}{2} + \sqrt{\tfrac{1}{2}\left[\nu^{(2)}(t) + 1\right]^2 + \ln \frac{d\nu^{(2)}(t)}{dt} - \frac{1}{4} - \ln C + \frac{k+1}{2} \ln\left(1+\frac{t^2}{k}\right)} \ .$$

Thus starting with an asymptotic approximation $\nu^{(0)}(t)$ whose ratio to $\nu(t)$ approaches 1 as $t\to\infty$, we have found approximations $\nu^{(1)}(t)$, $\nu^{(2)}(t)$, and $\nu^{(3)}(t)$ whose difference from $\nu(t)$ approaches 0 as $t\to\infty$. While the approximation $\nu^{(3)}(t)$ works well, it is only the latest variant of a series of approximations and there is no reason to think more rapidly converging iterates cannot be found.

Five algorithms for $\nu(t)$ were compared over a broad range of values of $\nu(t)$:

(1) $\nu(t) = \sqrt{2}\ \mathrm{sgn}(t)\ \mathrm{inverf}(Q_k(\ |t|\ ))$, where sgn is the usual signum function, inverf is the IMSL implementation DERFI of Strecok's algorithm, $Q_k(\cdot)$ is the finite expansion of the cdf of the absolute value of Student's t calculated as in Abramowitz and Stegun (1970, corrected version) §§26.7.3-26.7.4, and $|\cdot|$ is the absolute value operator;

(2) $\nu(t) = \Phi^{-1} \circ G_k(t)$, where Φ^{-1} is implemented by the IMSL routine DNORIN and G_k is implemented by the IMSL routine DTDF;

(3) $\nu(t) = \Phi^{-1} \circ G_k(t)$, where Φ^{-1} is implemented by the GAUINV routine (see Kennedy and Gentle (1980)) and G_k is implemented by the finite series from integration by parts given in Abramowitz and Stegun §§26.7.3-26.7.4;

(4) $\nu(t)$ calculated in double precision from the algorithms derived above for $\nu^{(3)}$;

(5) $\nu(t) = \sqrt{2 \ln\left(\frac{\nu(t+\Delta t) - \nu(t-\Delta t)}{2\,\Delta t}\right) + (k+1)\ln(1+\frac{t^2}{k}) + 2\ln C}$,

where the $\nu(\cdot)$ on the righthand side are approximated by the algorithm of (4) and $C = \ln k + 2\ln\Gamma(k/2) - 2\ln\Gamma(\frac{k+1}{2})$.

In general, for t such that $|\nu(t)| < 6$, (1) is slightly more accurate than (3), which is in turn slightly better than (2). For any practical work, (3) is excellent for t such that $|\nu(t)| < 7$ (error on the order of 10^{-6}), though any of (1)-(3) is acceptable. Either (4) or (5) should be used if $|\nu(t)| > 7$. As might be

expected, (5) is better than (4) for all t. Based on these facts, either (1) or (3) should be used for t for which $|\nu(t)| < 7$ and (4) or (5) should be used otherwise, the choice depending on availability of algorithms and execution speed.

6. CRITICAL VALUES OF THE STANDARDIZED STATISTIC $\check{S}_{\underline{n}}^{(h)}$

Critical values were computed by numerical evaluation of the Lévy inversion integral for the cdf via the IMSL double precision quadrature subroutines DQDAWO (near the removable singularity at 0) and DQDAWF (elsewhere) and inverting the cdf via a FORTRAN implementation of the Illinois method of *regula falsi* (see for example Kennedy and Gentle (1980)) using felicitous starting values. The IMSL quadrature programs are proprietary versions of the QUADPACK programs QAWO and QAWF of Piessens et al. (1983). All calculations were run on the IBM 3081 at Triangle Universities Computing Center, Research Triangle Park, NC.

Selected critical values are tabulated in Tables 2.1-2.6 for lags up to 6 for the case of uninterrupted stretches of data for $2 \leq n \leq 100$ and a few select higher values, where $\underline{n} = (n)$. Approximation formulas are provided in §7 for $n \geq 25$ and up to lag 6. An exact test can also be performed for any $\underline{n}$ for lags up to 6 using $S_{n_+}^{(h)}$ as the test statistic; a somewhat less powerful exact test could be performed for any $\underline{n}$ and any lag using additional cross-product terms to bridge the gaps. Recall that $n_+ = \sum n_j$, where $\underline{n} = (n_1, n_2, \cdots, n_J)$ is the run sequence of NU residuals.

The following relations should be observed. In general, by the symmetry of $\check{S}_{\underline{n}}$ we have

$$\check{F}_{\underline{n}}^{-1}(1 - p) = -\check{F}_{\underline{n}}^{-1}(p) \text{ for any } p \text{ in } [0, 1] \text{ and any } \underline{n}\text{, and hence}$$

$$\check{F}_{\underline{n}}^{-1}(0.5) = 0 \text{ for any } \underline{n}.$$

Also, since $\check{S}_2 \stackrel{d}{=} \check{S}_3$, it follows that $\check{F}_{2,\cdots,2}(t) = \check{F}_{3,\cdots,3}(t)$.

Finally, note that $\check{F}^{-1}_{\underset{\sim}{n},1}(p) = \check{F}^{-1}_{\underset{\sim}{n}}(p)$ if $0 \leq p \leq 1$.

In applications, the test statistic may be taken so that n_+ is the total number of NU residuals from all runs.

Table 2.1: Lag 1 Critical Values

n	.40	.30	.20	.10	.05	.025	.01	.005
2	.08873	.24939	.51688	1.03438	1.59510	2.18195	2.98381	3.60421
3	.08873	.24939	.51688	1.03438	1.59510	2.18195	2.98381	3.60421
4	.13816	.32173	.59087	1.07791	1.59492	2.13590	2.87807	3.45422
5	.15025	.34584	.62512	1.11225	1.61121	2.12126	2.81086	3.34270
6	.16780	.37171	.65310	1.13306	1.61706	2.10616	2.76048	3.26102
7	.17759	.38923	.67476	1.15036	1.62161	2.09303	2.71894	3.19511
8	.18663	.40382	.69167	1.16343	1.62474	2.08211	2.68504	3.14138
9	.19336	.41550	.70562	1.17407	1.62694	2.07263	2.65665	3.09667
10	.19908	.42521	.71712	1.18283	1.62861	2.06445	2.63250	3.05880
11	.20375	.43336	.72685	1.19019	1.62990	2.05731	2.61171	3.02628
12	.20773	.44029	.73516	1.19648	1.63092	2.05103	2.59361	2.99802
13	.21112	.44627	.74236	1.20192	1.63176	2.04547	2.57770	2.97321
14	.21405	.45147	.74865	1.20667	1.63246	2.04052	2.56360	2.95125
15	.21661	.45603	.75419	1.21087	1.63305	2.03608	2.55102	2.93167
16	.21885	.46006	.75911	1.21461	1.63356	2.03209	2.53972	2.91408
17	.22084	.46365	.76351	1.21796	1.63401	2.02848	2.52950	2.89819
18	.22262	.46687	.76747	1.22098	1.63442	2.02519	2.52023	2.88376
19	.22421	.46976	.77105	1.22372	1.63478	2.02220	2.51177	2.87171
20	.22565	.47238	.77430	1.22621	1.63510	2.01945	2.50403	2.85940
21	.22695	.47477	.77726	1.22850	1.63540	2.01693	2.49690	2.84812
22	.22814	.47694	.77998	1.23060	1.63568	2.01461	2.49033	2.83718
23	.22922	.47894	.78248	1.23254	1.63593	2.01246	2.48424	2.82770
24	.23022	.48077	.78478	1.23433	1.63617	2.01047	2.47859	2.81889
25	.23113	.48247	.78692	1.23599	1.63639	2.00862	2.47339	2.81069

Table 2.2: Lag 2 Critical Values

n	.40	.30	.20	.10	.05	.025	.01	.005
3	.08873	.24939	.51688	1.03438	1.59510	2.18195	2.98381	3.60421
4	.15779	.36121	.64792	1.13804	1.62817	2.11830	2.76622	3.25635
5	.15462	.35472	.63823	1.12685	1.62041	2.11881	2.78477	3.29367
6	.15779	.36121	.64792	1.13804	1.62817	2.11830	2.76622	3.25635
7	.17634	.38727	.67301	1.15055	1.62405	2.09687	2.72238	3.19627
8	.18774	.40485	.69148	1.16168	1.62336	2.08245	2.68816	3.14612
9	.19309	.41488	.70460	1.17293	1.62643	2.07336	2.65949	3.10120
10	.19800	.42382	.71587	1.18206	1.62848	2.06527	2.63510	3.06311
11	.20315	.43230	.72557	1.18917	1.62963	2.05815	2.61450	3.03083
12	.20733	.43942	.73392	1.19538	1.63060	2.05188	2.59641	3.00254
13	.21066	.44540	.74121	1.20091	1.63145	2.04625	2.58036	2.97756
14	.21357	.45062	.74758	1.20575	1.63216	2.04124	2.56611	2.95538
15	.21618	.45525	.75319	1.21001	1.63279	2.03677	2.55338	2.93555
16	.21847	.45934	.75817	1.21380	1.63332	2.03274	2.54193	2.91772
17	.22049	.46298	.76264	1.21721	1.63379	2.02909	2.53159	2.90161
18	.22228	.46624	.76666	1.22028	1.63421	2.02577	2.52218	2.88696
19	.22390	.46918	.77029	1.22307	1.63459	2.02274	2.51361	2.87360
20	.22536	.47184	.77359	1.22561	1.63493	2.01997	2.50575	2.86226
21	.22668	.47426	.77661	1.22793	1.63524	2.01742	2.49852	2.85080
22	.22789	.47647	.77937	1.23007	1.63553	2.01507	2.49185	2.84020
23	.22899	.47850	.78190	1.23204	1.63579	2.01289	2.48568	2.83005
24	.23000	.48036	.78424	1.23386	1.63604	2.01088	2.47995	2.82111
25	.23093	.48208	.78641	1.23556	1.63627	2.00900	2.47461	2.81278

Table 2.3: Lag 3 Critical Values

	.60	.70	.80	.90	.95	.975	.99	.995
4	.08873	.24939	.51688	1.03438	1.59510	2.18195	2.98381	3.60421
5	.15779	.36121	.64792	1.13804	1.62817	2.11830	2.76622	3.25635
6	.18837	.40973	.70260	1.17724	1.63409	2.08041	2.65993	3.09274
7	.18524	.40359	.69383	1.16772	1.62818	2.08227	2.67828	3.12816
8	.18592	.40503	.69610	1.17068	1.63069	2.08295	2.67401	3.11801
9	.18837	.40973	.70260	1.17724	1.63409	2.08041	2.65993	3.09274
10	.19701	.42247	.71513	1.18343	1.63150	2.06848	2.63579	3.05973
11	.20314	.43220	.72541	1.18937	1.63047	2.05941	2.61529	3.03024
12	.20781	.43996	.73404	1.19489	1.63023	2.05210	2.59745	3.00377
13	.21072	.44538	.74097	1.20051	1.63129	2.04660	2.58146	2.97907
14	.21339	.45029	.74715	1.20541	1.63212	2.04165	2.56725	2.95706
15	.21582	.45472	.75268	1.20971	1.63280	2.03719	2.55452	2.93733
16	.21818	.45886	.75763	1.21341	1.63328	2.03318	2.54319	2.91970
17	.22025	.46254	.76209	1.21677	1.63371	2.02953	2.53290	2.90368
18	.22208	.46583	.76611	1.21981	1.63412	2.02621	2.52351	2.88907
19	.22369	.46877	.76976	1.22261	1.63449	2.02316	2.51491	2.87568
20	.22515	.47144	.77308	1.22517	1.63482	2.02037	2.50701	2.86339
21	.22647	.47387	.77611	1.22751	1.63514	2.01780	2.49975	2.85282
22	.22769	.47610	.77889	1.22966	1.63543	2.01543	2.49303	2.84160
23	.22880	.47814	.78144	1.23165	1.63570	2.01325	2.48682	2.83190
24	.22982	.48002	.78380	1.23349	1.63595	2.01122	2.48104	2.82289
25	.23076	.48176	.78599	1.23520	1.63618	2.00933	2.47566	2.81448

Table 2.4: Lag 4 Critical Values

n	.60	.70	.80	.90	.95	.975	.99	.995
5	.08873	.24939	.51688	1.03438	1.59510	1.18195	2.98381	3.60421
6	.15779	.36121	.64792	1.13804	1.62817	1.11830	2.76622	3.25635
7	.18837	.40973	.70260	1.17724	1.63409	1.08041	2.65993	3.09274
8	.20468	.43653	.73310	1.19864	1.63591	1.05650	2.59591	2.99512
9	.20205	.43143	.72595	1.19115	1.63162	1.05873	2.61171	3.02490
10	.20185	.43112	.72569	1.19132	1.63234	1.05985	2.61271	3.02506
11	.20293	.43322	.72867	1.19448	1.63416	1.05891	2.60603	3.01251
12	.20468	.43653	.73310	1.19864	1.63591	1.05650	2.59591	2.99512
13	.20941	.44388	.74059	1.20247	1.63432	1.04900	2.58056	2.97405
14	.21308	.44989	.74708	1.20623	1.63353	1.04290	2.56694	2.95454
15	.21606	.45495	.75276	1.20983	1.63320	1.03777	2.55469	2.93654
16	.21855	.45929	.75779	1.21320	1.63314	1.03334	2.54360	2.91993
17	.22036	.46265	.76207	1.21661	1.63366	1.02972	2.53337	2.90417
18	.22203	.46574	.76598	1.21968	1.63411	1.02642	2.52403	2.88974
19	.22356	.46858	.76955	1.22248	1.63450	1.02339	2.51548	2.87650
20	.22498	.47118	.77282	1.22503	1.63485	1.02061	2.50762	2.86431
21	.22633	.47362	.77582	1.22732	1.63514	1.01806	2.50043	2.85312
22	.22755	.47586	.77859	1.22944	1.63541	1.01570	2.49376	2.84273
23	.22867	.47791	.78114	1.23141	1.63566	1.01351	2.48757	2.83308
24	.22970	.47979	.78350	1.23324	1.63590	1.01148	2.48181	2.82409
25	.23064	.48153	.78569	1.23495	1.63613	1.00958	2.47643	2.81569

Table 2.5: Lag 5 Critical Values

n	.60	.70	.80	.90	.95	.975	.99	.995
6	.08873	.24939	.51688	1.03438	1.59510	2.18195	2.98381	3.60421
7	.15779	.36121	.64792	1.13804	1.62817	2.11830	2.76622	3.25635
8	.18837	.40973	.70260	1.17724	1.63409	2.08041	2.65993	3.09274
9	.20468	.43653	.73310	1.19864	1.63591	2.05650	2.59591	2.99512
10	.21460	.45340	.75265	1.21239	1.63673	2.04020	2.55281	2.92954
11	.21244	.44923	.74686	1.20646	1.63354	2.04241	2.56612	2.95420
12	.21193	.44828	.74567	1.20559	1.63356	2.04366	2.56917	2.95853
13	.21235	.44914	.74694	1.20705	1.63455	2.04349	2.56630	2.95271
14	.21333	.45100	.74948	1.20953	1.63570	2.04220	2.56031	2.94214
15	.21460	.45340	.75265	1.21239	1.63673	2.04020	2.55281	2.92954
16	.21750	.45810	.75762	1.21505	1.63572	2.03509	2.54218	2.91483
17	.21989	.46213	.76206	1.21768	1.63516	2.03076	2.53247	2.90094
18	.22192	.46563	.76606	1.22023	1.63488	2.02700	2.52356	2.88789
19	.22368	.46873	.76968	1.22266	1.63479	2.02367	2.51534	2.87564
20	.22522	.47149	.77297	1.22496	1.63481	2.02070	2.50774	2.86417
21	.22644	.47376	.77587	1.22725	1.63512	2.01817	2.50062	2.85319
22	.22758	.47588	.77856	1.22938	1.63541	2.01582	2.49401	2.84297
23	.22864	.47784	.78106	1.23134	1.63568	2.01365	2.48787	2.83393
24	.22963	.47967	.78338	1.23317	1.63592	2.01162	2.48214	2.82455
25	.23054	.48138	.78554	1.23487	1.63615	2.00974	2.47679	2.81623

Table 2.6: Lag 6 Critical Values

n	.60	.70	.80	.90	.95	.975	.99	.995
7	.08873	.24939	.51688	1.03438	1.59510	2.18195	2.98381	3.60421
8	.15779	.36121	.64792	1.13804	1.62817	2.11830	2.76622	3.25635
9	.18837	.40973	.70260	1.17724	1.63409	2.08041	2.65993	3.09274
10	.20468	.43653	.73310	1.19864	1.63591	2.05650	2.59591	2.99512
11	.21460	.45340	.75265	1.21239	1.63673	2.04020	2.55281	2.92954
12	.22122	.46493	.76627	1.22209	1.63724	2.02842	2.52170	2.88217
13	.21944	.46150	.76152	1.21731	1.63480	2.03049	2.53292	2.90272
14	.21882	.46034	.76002	1.21604	1.63453	2.03174	2.53673	2.90868
15	.21891	.46055	.76039	1.21660	1.63506	2.03194	2.53593	2.90660
16	.21945	.46160	.76184	1.21808	1.63582	2.03130	2.53244	2.90024
17	.22025	.46313	.76390	1.21999	1.63657	2.03005	2.52745	2.89170
18	.22122	.46493	.76627	1.22209	1.63724	2.02842	2.52170	2.88217
19	.22314	.46817	.76979	1.22406	1.63658	2.02474	2.51390	2.87128
20	.22481	.47102	.77301	1.22601	1.63619	2.02153	2.50664	2.86087
21	.22626	.47357	.77597	1.22792	1.63598	2.01868	2.49987	2.85096
22	.22755	.47587	.77868	1.22976	1.63589	2.01612	2.49355	2.84156
23	.22871	.47796	.78119	1.23152	1.63589	2.01379	2.48764	2.83266
24	.22976	.47986	.78351	1.23319	1.63594	2.01167	2.48209	2.82424
25	.23064	.48150	.78561	1.23485	1.63616	2.00980	2.47684	2.81641

7. APPROXIMATE CRITICAL POINTS

Critical values were calculated for $n = 2^k n_0$, $k = 0\ (1)\ 6$, where $n_0 = 25$ for $h = 1, 2, 4, 5$ and $n_0 = 24$ for $h = 3, 6$. The different choice of sequences for lags 3 and 6 cut computation time by half since the lag reduction algorithm entails multiplication of different characteristic functions rather than a simple exponentiation as when h divides n. Because $\check{F}_n(t) \xrightarrow{d} \Phi(t)$, it is reasonable to approximate the fractional points by the formula

$$\check{F}_n^{-1}(1-\alpha) \approx \Phi^{-1}(1-\alpha) \cdot \{1 + \gamma_1/n + \gamma_2/n^2 + \gamma_3/n^3 + \gamma_4/n^4\}. \tag{7.1}$$

Aitken's method of accelerating the rate of convergence was applied to these sequences of 7 observations $n\left[\check{F}_n^{-1}(1-\alpha) - \Phi^{-1}(1-\alpha)\right]$ for each lag $h = 1$ to 6 at each of the relevant significance levels to estimate the large-sample coefficients γ_1 of $1/n$.

Additional critical values for the standardized statistic $\check{S}_n^{(h)}$ were calculated for lags $h = 1$ to 6, for upper significance levels $\alpha = 0.4, 0.3, 0.2, 0.1, 0.05, 0.025, 0.01, 0.005$, based on $n = h + 1\ (1)\ 100\ (5)\ 120\ (10)\ 200\ (25)\ 350$. Using those critical values for which $n > 24$ and the critical values of the sequence used to perform the acceleration, linear regression coefficients $\gamma_2, \gamma_3, \gamma_4$ were estimated for $1/n^2$, $1/n^3$, $1/n^4$ from the errors of approximation in using only the accelerated coefficients of $1/n$. The rationale for this schizophrenic approach is that γ_1 should be chosen as $\lim_{n\to\infty} n\left[\check{F}_n^{-1}(1-\alpha) - \Phi^{-1}(1-\alpha)\right]$ to perform well for large n (since higher order terms in $1/n$ are negligible for large n); choosing the remaining coefficients by least squares is then the best we can do for small n.

An appropriate criterion for goodness of approximation is the relative error in the significance level. In practice a relative error in significance of 4% seems tolerable; for example, using an approximate 0.005 critical value whose actual level is between 0.0048 and 0.0052. For $n \geq 25$ the approximations below did <u>much</u> better than this in all cases except for approximations at the 0.005 level of significance for $70 \leq n \leq 79$; but even in these cases, the 4% relative error criterion was met.

Table 3 shows upper bounds within which the approximation's relative errors in significance (ϵ/α) typically lie for the eight upper significance levels (α) treated.

Table 3: Bounds for Approximation Errors

α:	0.4	0.3	0.2	0.1	0.05	0.025	0.01	0.005
ϵ/α:	$4x10^{-9}$	$4x10^{-8}$	$4x10^{-7}$	$4x10^{-6}$	$4x10^{-5}$	$4x10^{-4}$	$4x10^{-3}$	$4x10^{-2}$

The following tables give the recommended coefficients for use in the approximation (7.1) for common significance levels and lags up to 6 in the uninterrupted case; since the numbers are intended for automated computers, all known digits are listed.

Table 4.1: Lag 1 Coefficients

α	γ_1	γ_2	γ_3	γ_4
.40	-2.2018609079067952	.02483247	5.56491955	-.22031284
.30	-2.0439005481334844	.78742505	8.22258051	1.94856091
.20	-1.7186855465224651	2.05015448	10.53609512	-80.75271870
.10	-1.0182364890230256	3.77856794	-11.83118531	-42.85325259
.05	-.2196386902033893	3.67304351	-41.87959508	171.64333
.025	.6298870507074382	1.78134365	-77.30203422	674.12052
.010	1.7840115715504092	-2.98269798	-102.58336	1242.29727
.005	2.7141006702279862	-31.28489656	1181.08626	-17130.05392

Table 4.2: Lag 2 Coefficients

α	γ_1	γ_2	γ_3	γ_4
.40	-2.2018581315533850	.71008553	11.56017480	-7.80268431
.30	-2.0440186898650741	.12199870	13.37007106	2.19411157
.20	-1.7186782305021462	1.47581621	15.99012927	-94.11589238
.10	-1.0181797850996761	3.43360150	-7.75974765	-63.47489761
.05	-.2196921456332459	3.61257587	-41.23167017	166.09539
.025	.6307435303877188	1.89097057	-75.83876556	633.25675
.010	1.7859785374745305	-2.50075083	-108.63472	1274.49982
.005	2.7147613920075995	-32.93569570	1339.59160	-19779.26161

Table 4.3: Lag 3 Coefficients

α	γ_1	γ_2	γ_3	γ_4
.40	-2.2018575666916276	-1.44542890	19.15504990	.04138097
.30	-2.0435633167117739	-.62467675	23.41911638	-28.36569361
.20	-1.7186767113438366	.90131398	22.68636753	96.50830012
.10	-1.0181200942297766	3.08353258	-2.54332788	-86.69856452
.05	-.2210600309476833	3.73691584	-48.32353182	266.90143
.025	.6309096833245742	3.08713322	-78.22946587	632.97737
.010	1.8373288485475277	-8.87869485	171.00196	-2407.60781
.005	2.7251078190057194	-35.30502626	1505.52761	-22380.39014

Table 4.4: Lag 4 Coefficients

α	γ_1	γ_2	γ_3	γ_4
.40	-2.2018563202461787	-2.16934009	27.24573400	60.69992507
.30	-2.0439869938329420	-1.24404699	29.30598728	45.82082416
.20	-1.7186719036997094	.32533953	30.65317514	-75.97173879
.10	-1.0180818808376466	2.74034453	2.84096708	-84.91993372
.05	-.2210044324757457	3.66542685	-47.27045697	272.50639
.025	.6310939600739425	2.28268333	-81.23131684	633.42523
.010	1.7905246764717355	-1.62414677	-121.64846	1284.72265
.005	2.7257994177485485	-34.07814800	1468.19528	-22029.82427

Table 4.5: Lag 5 Coefficients

α	γ_1	γ_2	γ_3	γ_4
.40	-2.2018533349497773	-3.02159907	48.58135915	-73.45782908
.30	-2.0439983335145693	-1.98039687	44.44474461	20.94451687
.20	-1.7147782847511177	-.78712448	61.97234757	-299.35216
.10	-1.0180603806482367	2.42919737	5.94211564	4.1765832
.05	-.2206549288331952	3.56341017	-45.49481651	290.70388
.025	.6311398265245110	2.49197871	-85.08359918	629.49069
.010	1.7923375236878381	-1.41825626	-112.33476	949.19392
.005	2.7253570485782885	-32.83323820	1430.55714	-21766.6278

Table 4.6: Lag 6 Coefficients

α	γ_1	γ_2	γ_3	γ_4
.40	-2.2018442560015819	-3.33812426	21.80037537	903.63192
.30	-2.0439669717652733	-2.42230919	33.20595563	650.28942
.20	-1.7217435937478880	-.34224500	26.55923169	453.95547
.10	-1.0181089207042184	2.02908328	18.18284135	-48.44684922
.05	-.2205013028683279	3.42857580	-39.28322150	222.53210
.025	.6320449869795524	2.59663812	-85.78973639	578.48392
.010	1.8093001853119177	-3.18255493	-27.12092207	-360.88842
.005	2.7260068895936032	-30.34289271	1310.37656	-20298.18760

9. SIMULATION RESULTS

Single precision standard normal variates were generated by the Kinderman-Monahan (1977) algorithm (see Knuth (1977)). The NU transformation $\Phi^{-1}(G_k(t))$ was computed in double precision by composition of the formulas §§26.7.3–26.7.4 of Abramowitz & Stegun (1972) and the algorithm GAUINV (Kennedy and Gentle (1980), p. 95) for arguments t for which $\Phi(-7) < G_k(t) < \Phi(7)$, and by the asymptotic approximation $\nu^{(3)}(t)$ otherwise.

"Exact" critical values for the Durbin-Watson test were calculated by estimating the Lévy inversion integral (4.14) for the cdf via the IMSL double precision quadrature subroutines DQDAGI (DQDAWF can be used alternatively if DQDAWO is used near the removable singularity at 0) and inverting the cdf via a FORTRAN implementation of the Illinois method using Henshaw's (1966) approximation as starting values. Though Henshaw's approximation is good, it was inadequate for the simulation studies as it tends to underestimate the tail probabilities, leading to critical values that reject H_0: $\rho = 0$ as much as 40% more often in the upper tails than the nominal when H_0 is in fact true.

From the simulation results with 20 observations from a linear trend and a stationary AR(1) error structure, the proposed NU test always rejected more often than the Durbin-Watson bounds test. The exact test based on the Durbin-Watson statistic (whose distribution

depends on the design matrix) generally performed somewhat better than the NU test in the cases under simulation.

The general form of the AR(h) model we consider in the simulation assumes that

$$y_N = X_N \beta + e, \tag{8.1}$$

where $e = (e_1, e_2, \cdots, e_N)'$ is a vector of random disturbances satisfying

$$e_k = \rho_h e_{k-h} + v_k, \text{ and} \tag{8.2}$$

$$v_k \text{ are iid } N(0, \sigma^2). \tag{8.3}$$

Actually there are two natural families satisfying these assumptions:

(1) the stationary errors family, for which (8.2) is postulated for <u>all</u> integers k. This is the infinite memory case,

(2) the initially error-free family, for which (8.2) is postulated only for positive integers k; errors e_k for $k \leq 0$ are taken to be zero.

This is the finite memory case.

Then
$$e_k = \begin{cases} \sum_{j=0}^{\infty} \rho_h^j v_{k-jh}, & \text{if } |\rho_h| < 1 \text{ in the stationary case;} \\ \sum_{j=0}^{(k-1)/h} \rho_h^j v_{k-jh}, & \text{in the initially error-free case.} \end{cases} \tag{8.4}$$

If $h = 1$, $\text{Var}(y) = E\, e\, e' = \Sigma_1$, where the n x n matrix Σ_1 is defined by

$$\sigma^2 \rho^{|j-k|}/(1-\rho^2) \text{ in the stationary case}$$

$$(j, k)- \text{ entry of } \Sigma_1 = \sigma^2 \left(\rho^{|j-k|} - \rho^{j+k}\right)/(1-\rho^2) \text{ in the error-free case if } \rho \neq 1 \tag{8.5}$$

$$\sigma^2 \min(j, k) \text{ in the error-free case if } \rho = 1.$$

Since the AR(h) model in (8.3) above consists of h independent, interspersed AR(1) models with a dilation of time, it follows that

Table 5.1: $\beta_0 = 1, \beta_1 = 0.02$

ρ	Significance Level α	% DW Rej.	% NU Rej.	% Rej. DW Bds.	% Inconclusive DW Bds.Test
0	5.0	4.9	5.0	1.3	3.6
	2.5	2.5	2.5	0.6	1.8
	1.0	0.9	1.1	0.2	0.7
	0.5	0.5	0.5		
.1	5.0	10.2	9.7	3.7	6.6
	2.5	5.6	5.4	1.6	4.0
	1.0	2.5	2.3	0.7	1.9
	0.5	1.3	1.3		
.2	5.0	18.1	16.5	7.3	10.9
	2.5	10.9	9.6	3.8	7.0
	1.0	5.4	4.6	1.4	4.0
	0.5	3.2	2.5		
.4	5.0	43.0	38.7	24.4	18.7
	2.5	31.4	26.8	15.7	15.8
	1.0	20.0	15.8	8.4	11.7
	0.5	14.0	10.3		
.8	5.0	84.6	80.7	72.5	12.2
	2.5	77.8	71.9	63.6	14.2
	1.0	68.5	60.1	51.8	16.7
	0.5	61.3	51.0		

$$\text{Var(y)} = \begin{bmatrix} \sigma_{11}\mathrm{I_h} & \sigma_{12}\mathrm{I_h} & \cdots & \sigma_{1n}\mathrm{I} \\ \sigma_{12}\mathrm{I_h} & \sigma_{22}\mathrm{I_h} & \cdots & \sigma_{2n}\mathrm{I} \\ \vdots & \vdots & & \vdots \\ \sigma_{1n}\mathrm{I} & \sigma_{2n}\mathrm{I} & \cdots & \sigma_{nn}\mathrm{I} \end{bmatrix},$$

where the last row- and column-blocks are trimmed to make Var(y) an N x N matrix. If h divides N, then this reduces to

$$\text{Var(y)} = \Sigma_1 \otimes \mathrm{I_h}, \tag{8.6}$$

where $\otimes$ denotes the right Kronecker product of the matrices.

Although the simulations were run for both the initially error-free model and the

Table 5.2: $\beta_0 = 1, \beta_1 = 0.002$

ρ	Significance Level α	% DW Rej.	% NU Rej.	% Rej. DW Bds.	% Inconclusive DW Bds.Test
	5.0	5.0	5.0	1.4	3.6
0	2.5	2.4	2.5	0.6	1.8
	1.0	1.0	1.0	0.2	0.8
	0.5	0.5	0.5		
	5.0	9.5	9.3	3.2	6.4
.1	2.5	5.1	4.9	1.4	3.7
	1.0	2.1	2.0	0.5	1.6
	0.5	1.2	1.0		
	5.0	18.4	16.9	7.8	10.7
.2	2.5	11.3	10.1	3.9	7.4
	1.0	5.8	5.0	1.7	4.1
	0.5	3.3	2.8		
	5.0	43.0	24.2	38.6	18.9
.4	2.5	31.2	26.8	15.6	15.6
	1.0	20.1	15.9	8.4	11.8
	0.5	13.9	10.4		
	5.0	84.7	80.6	72.7	12.0
.8	2.5	78.2	72.2	63.3	14.9
	1.0	68.6	59.8	51.6	17.1
	0.5	61.0	50.8		

stationary model, only the results for the stationary model are reported here since the results are very similar for $|\rho| < 1$. The simulation results are summarized in Tables 5.1 - 5.3. The simple linear regression model with correlated error structure for these tables is

$$Y_k = \beta_0 + \beta_1 k + e_k, \text{ where } e_k = \rho e_{k-1} + v_k, \{v_k\}_{k=1}^{20} \text{ are iid Normal(0,1)}$$

and 27,000 iterations were run for all cases. The parameters for this study were taken to be values similar to ones that we have observed in actual machining processes.

10. EXAMPLES

Example 1. The data in Table 6 (see Figures 2 and 3 in addition to the data listing) are part diameter measurements obtained from five different settings of a machining process. Possibly different regression regimes can be identified by the change in the tool setting. Part Diameter

Table 5.3: $\beta_0 = 40$, $\beta_1 = 1$

ρ	Significance Level α (%)	% DW Rej.	% NU Rej.	% Rej. DW Bds.	% Inconclusive DW Bds.Test
	5.0	5.1	5.0	0.2	8.9
0	2.5	2.5	2.5	0.1	4.9
	1.0	0.9	1.0	0.0	2.1
	0.5	0.5	0.5		
	5.0	13.7	10.4	1.2	20.0
.1	2.5	7.8	6.0	0.5	12.9
	1.0	3.8	2.7	0.1	6.9
	0.5	2.1	1.5		
	5.0	27.9	19.0	4.0	34.8
.2	2.5	18.4	11.9	1.9	25.6
	1.0	10.3	5.9	.7	16.4
	0.5	6.6	3.5		
	5.0	66.7	46.1	24.7	51.4
.4	2.5	55.7	34.4	16.0	50.2
	1.0	41.5	22.0	8.8	44.7
	0.5	32.5	15.1		
	5.0	98.2	89.8	86.6	12.4
.8	2.5	96.5	84.2	80.1	18.0
	1.0	93.6	76.0	70.3	25.9
	0.5	90.5	69.3		

was regressed against Serial Number, allowing different coefficients for linear tool-wear, intercept, and diameter variance for each setup (setting); NU residuals were computed and the corresponding NU statistics calculated for lags h = 1(1)7. The null hypothesis of part diameter independence is to be tested against the alternative of positive serial correlation.

Recall that $S_{\underset{\sim}{n}}^{(h)} = \sum_{j=1}^{J} S_{n_j}^{(h)}$, where

$$S_{n_j}^{(h)} = \sum_{k=1}^{n_j} Z_{kj} Z_{k+hj},$$

Z_{kj} represents the k^{th} NU residual from the j^{th} regression setup,

$\underset{\sim}{n} = (n_1, n_2, \cdots, n_J)$ is the vector whose j^{th} component is the number of NU residuals in the jth regression setup,

h is the lag to be tested,

J is the number of regression setups considered, and

$\text{Var } S_{\underline{n}}^{(h)}$, the variance of the unnormalized statistic, is equal to

the total number of cross-product terms in the sum $S_{\underline{n}}^{(h)}$.

Form the test statistic $\check{S}_{\underline{n}}^{(h)} = S_{\underline{n}}^{(h)} / \sqrt{\text{Var } S_{\underline{n}}^{(h)}}$ and calculate the probability of a more extreme value of the test statistic. In the absence of a table of critical values for the grouping structure n = (7, 6, 7, 6, 8), or a computer algorithm to calculate the P-value, a simpler (exact under the null hypothesis, though less powerful) procedure is available, which will be illustrated for lag 1 below.

Since under the null hypothesis of independent errors the NU residuals are all iid standard normal random variables, incorporating the additional NU residuals cross-product terms $\sum_{j=1}^{J} Z_{n_j,j} Z_{1,j+1}$ to bridge the gaps between setups (or regression regimes) will give rise to the same one-dimensional family of distributions under consideration at the cost of a little random noise.

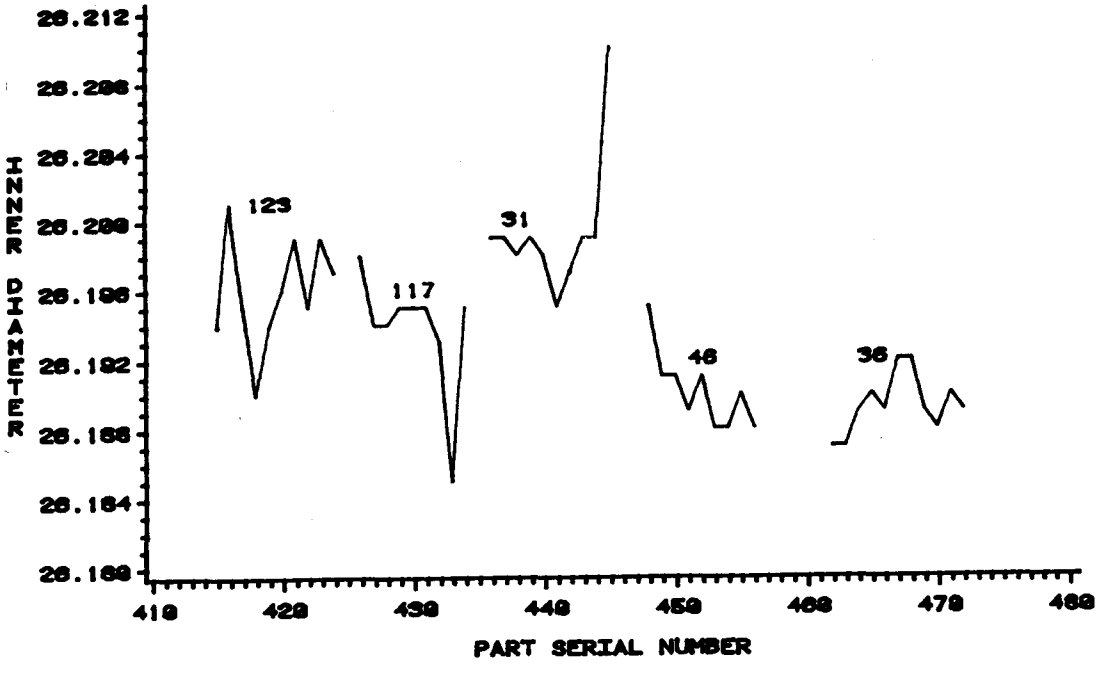

Figure 2: Part Diameter vs. Serial Number, with Tool Setting labeled

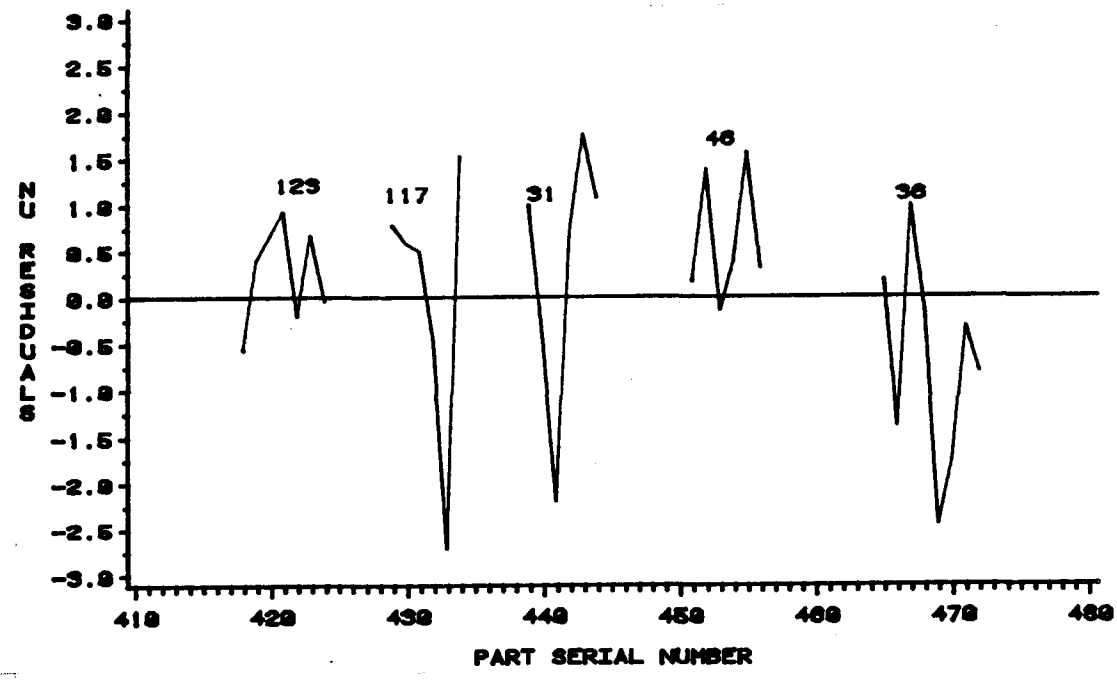

Figure 3: NU residuals of the same data plotted against Serial Number

For clarity when dealing with the Part Diameter data, a single subscript will be used to refer to the part serial number, which uniquely determines the regression setup and its run order within that setup.

The lags $h = 6, 7$ have so few degrees of freedom (the four terms $z_{418}\, z_{424} + z_{439}\, z_{445} + z_{465}\, z_{471} + z_{466}\, z_{472}$ form the statistic $S^{(6)}$) that the power to detect nonzero autocorrelation is low. Also, it should be observed that the test statistics $S_{\underline{n}}^{(h)}$ for $h = 1, 2, \cdots$ are <u>not</u> independent, and the evaluations of test significance given here is with regard to the marginal distributions of the individual statistics for different lags. We do not give here a method for combining these tests into an overall trest. The statistics were all calculated here because it requires little additional labor after computation of the NU residuals.

To illustrate the simplified procedure for this data:

$$S_{n+}^{(1)} = S_{\underline{n}}^{(1)} + Z_{424}\, Z_{429} + Z_{434}\, Z_{439} + Z_{445}\, Z_{451} + Z_{456}\, Z_{465}$$

$$= 8.8684 + 1.2622 = 10.1306;$$

$$\text{Var } S_{n_+}^{(1)} = \text{Var } S_{\underline{n}}^{(1)} + 4 = 29 + 4 = 33;$$

$$\check{S}_{n_+}^{(1)} = S_{n_+}^{(1)} / \sqrt{\text{Var } S_{n_+}^{(1)}} = 1.7635 > 1.63777 \text{ (from the tables, } \alpha = 0.05).$$

The hypothesis of independent errors is rejected at the .05 level of significance in favor of the alternative that there is positive autocorrelation.

Note further that the observed value 1.6481 of the lag 1 NU test statistic has a P-value of 0.0492, as compared with the approximating normal distribution's P-value of 0.0498. Examination of the percentage points in the tables of section 6 shows that for one-sided tests with α between 0.05 and 0.50, using the standard normal tables is conservative for all lags h = 1, $\cdots$, 6; however, for one-sided α less than or equal to .025, the opposite is true – using the standard normal tables to evaluate P-values gives a slightly small value. In both cases of course, the bias decreases with increasing n. □

To determine whether apparent serial correlation is actually the result of model mis-specification, a data set may be split into two or more data segments, for each of which separate NU residuals are calculated anew. Since these residuals are all iid standard normal under the null hypothesis of independence, they may be combined in the usual way to form a statistic $\check{S}_{\underline{n}}^{(h)}$ or $\check{S}_{n_+}^{(h)}$; if the independence hypothesis does not remain rejectable under this segmentation of the data set, there is evidence of model mis-specification. Although such segmentation of datasets is time-honored in regression modeling, using the NU statistic procedure has the advantage that the information from the segments can be re-assembled losing only the degrees of freedom used to estimate the additional parameters and generating a statistic in the same family as the one based on the unsegmented dataset. We illustrate this procedure below with a second example.

Example 2 The values of y listed in Table 7 below were generated from the model $y = 1 + 0.05\, x^2 + v$, where the variates v are iid normal$(0, \sigma^2)$ random disturbances and x is the sequence of integers from 1 to 20. However, y was fitted to the model $y = \beta_0 + \beta_1 x +$ error using the NU transformation on the 20 observations; the resulting NU residuals z are listed in

Table 6: Data for Example 1

Part Diameter	Tool Setting	Serial No.	NU Residuals
26.194	123	415	...
26.201	123	416	...
26.195	123	417	...
26.190	123	418	-.5624
26.194	123	419	.4001
26.196	123	420	.6668
26.199	123	421	.9330
26.195	123	422	-.2109
26.199	123	423	.6760
26.197	123	424	.0499
26.198	117	426	...
26.194	117	427	...
26.194	117	428	...
26.195	117	429	.7813
26.195	117	430	.5814
26.195	117	431	.4858
26.193	117	432	-.4787
26.185	117	433	-2.7149
26.195	117	434	.5225
26.199	31	436	...
26.199	31	437	...
26.198	31	438	...
26.199	31	439	.9852
26.198	31	440	.4587
26.195	31	441	-2.1995
26.197	31	442	.7071
26.199	31	443	1.7663
26.199	31	444	1.0696
26.210	31	445	3.7083
26.195	46	448	...
26.191	46	449	...
26.191	46	450	...
26.189	46	451	.1764
26.191	46	452	.3854
26.188	46	453	-.1449
26.188	46	454	.3875
26.190	46	455	1.5609
26.188	46	456	.3080
26.187	36	462	...
26.187	36	463	...
26.189	36	464	...
26.190	36	465	.1764
26.189	36	466	-1.3854
26.192	36	467	.9897
26.192	36	468	-.1496
26.189	36	469	-2.4605
26.188	36	470	-1.7534
26.190	36	471	-.3158
26.189	36	472	-.8095

Table 7: Statistics for Example 1

	h	$S^{(h)}$	Variance	$\check{S}y^{(h)}$	Est. rho	
1	8.868352		29	1.64681	.306 *	(P=.0492)
2	0.292127		24	.05963	.012	(P=.4745)
3	4.610235		19	1.05766	.243	(P=.1474)
4	-5.281053		14	-1.41142	-.377	(P=.9240)
5	-0.476594		9	-.15886	-.053	(P=.5681)
6	4.747348		4	2.37367	1.187	(P=.0146)
7	-0.142831		1	-.14283	-.143	(P=.6395)

the third column ($\check{S}_{17} = 2.60485$, $P < 0.009$). On partitioning the dataset into two data segments of length 10 and calculating NU residuals (z', column 4) for each of these data segments, an appropriate NU statistic (either $\check{S}_{7,7} \stackrel{d}{=} \check{S}_{14}^{(2)}$, or $\check{S}_{14}^{(1)}$ depending on the availability of tables) can be calculated from the normalized uniform residuals ($\check{S}_{7,7} = 0.54683$, $P = 0.26433$). Since the NU statistic based on the partitioned dataset shows no evidence of autocorrelation, while the unpartitioned dataset does, we infer that the detected "serial correlation" is an artifact of forcing the data to an inadequate model.

Table 7: Data for Example 2

x	y	z	z'
1	4.10585	...	...
2	1.84976	...	...
3	2.27203	...	...
4	3.63638	.83766	.83766
5	6.88702	1.33556	1.33556
6	6.72663	.26211	.26211
7	6.28006	-.38554	-.38554
8	6.31325	-.57817	-.57817
9	6.07366	-.82606	-.82606
10	8.66134	.58678	.58678
11	8.02143	-.32480	...
12	11.98955	1.77307	...
13	13.32778	1.50438	...
14	13.71418	.86521	-.84304
15	13.99398	.36368	-.90446
16	14.79788	.26601	-.67361
17	18.82935	1.98346	1.06124
18	18.94198	1.04947	-.06025
19	19.95746	.87303	-.25784
20	23.36098	1.94325	1.15477

BIBLIOGRAPHY

Abrahamse, A. P. J., and Koerts, J. (1969), "A Comparison Between the Power of the Durbin-Watson Test and the Power of the BLUS Test," Journal of the American Statistical Association 64: 938–948.

Abrahamse, A. P. J., and Koerts, J. (1971), "New Estimators of Disturbances in Regression Analysis," Journal of the American Statistical Association 66: 71–74.

Abramowitz, M., and Stegun, I. A. (1972), Handbook of Mathematical Functions, 10th (corrected) printing, National Bureau of Standards Applied Mathematics Series 55, Washington DC: U.S. Government Printing Office.

De Boor, C. (1971), "CADRE: An Algorithm for Numerical Quadrature," Mathematical Software, (ed. by John R. Rice), New York: Academic Press, Chapter 7.

Durbin, J., and Watson, G. S. (1950), "Testing for Serial Correlation in Least Squares Regression. I," Biometrika 37: 409–428.

Durbin, J., and Watson, G. S. (1951), "Testing for Serial Correlation in Least Squares Regression. II," Biometrika 38: 159–178.

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

Henshaw, R. C. (1966), "Testing Single-Equation Least Squares Regression Models for Autocorrelated Disturbances," Econometrica 34:646–660.

IMSL, Version 10.0, (MATH/LIBRARY 1.0, SFUN/LIBRARY 2.0), April 1987, Houston, Texas.

Kennedy, W. J., and Gentle, J. E. (1980), Statistical Computing, New York: Marcel Dekker.

Kinderman, A. J., and Monahan, J. F. (1977), "Computer Generation of Random Variables Using the Ratio of Uniform Deviates," ACM Transactions On Mathematical Software, 3(3): 257–260.

Knuth, D. E. (1977), The Art of Computer Programming, vol. 2: Algorithms (2nd ed.), Reading, Massachusetts: Addison-Wesley.

Koerts, J., and Abrahamse, A. P. J. (1969), On the Theory and Application of the General Linear Model, Rotterdam: Rotterdam University Press.

Morrison, D. F. (1983), Applied Linear Statistical Methods, Englewood Cliffs: Prentice-Hall.

O'Reilly, F., and Quesenberry, C. P. (1972), "Uniform Strong Consistency of Rao-Blackwell Distribution Function Estimators," Annals of Mathematical Statistics 43(5):1678–1679.

O'Reilly, F., and Quesenberry, C. P. (1973), "The Conditional Probability Integral Test and Applications to Obtain Composite Chi-Square Goodness of Fit Tests," Annals of Statistics 1: 74–83.

Piessens, R., Doncker-Kapenga, E., Überhuber, C., Kahaner, D. K. (1983), QUADPACK: A Subroutine Package for Automatic Integration, New York: Springer-Verlag.

Quesenberry, C. P. (1986), "Some Transformation Methods in Goodness-of-Fit," Handbook of Goodness of Fit, (ed. by R. B. D'Agostino and M. A. Stephens), Chapter 6, New York: Marcel Dekker.

Quesenberry, C. P. (1988), "An SPC Approach to Compensating a Tool-Wear Process," Journal of Quality Technology 20: 220–229.

Strecok, A. J. (1968), "On Calculation of the Inverse of the Error Function," Mathematics of Computation 22: 144–158.

Theil, H. (1965), "The Analysis of Disturbances in Regression Analysis," Journal of the American Statistical Association 60: 1067–1079.

Theil, H. (1968), "A Simplification of the BLUS Procedures for Analysing Regression Disturbances," Journal of the American Statistical Association 63: 242–251 .

Theil, H. (1971), Principles of Econometrics, New York: John Wiley and Sons, Inc.

Van Beeck, P. (1972), "An Application of Fourier Methods to the Problem of Sharpening the Berry-Esséen Inequality," Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 23: 187–196.

Zahl, S. (1966), "Bounds for the Central Limit Theorem Error," SIAM Journal of Applied Mathematics 14(6): 1225–1245.