SIMULTANEOUS TESTS FOR TREND AND SERIAL CORRELATIONS
FOR GAUSSIAN MARKOV RESIDUALS
(Reprint from Econometrica, Vol. 34, No. 2, April 1966)

P. R. KRISHNAIAH
APPLIED MATHEMATICS RESEARCH LABORATORY
V. K. MURTHY
DOUGLAS AIRCRAFT CO., INC.
SANTA MONICA, CALIFORNIA

Distribution of this document is unlimited

OFFICE OF AEROSPACE RESEARCH
United States Air Force

ARCHIVE COPY
NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Qualified requesters may obtain copies of this report from the Defense Documentation Center, (DDC), Cameron Station, Alexandria, Virginia.

(Reproduction in whole or in part is permitted for any purpose of the U.S. Gov't)

Distribution of this document is unlimited

Copies of ARL Technical Documentary Reports should not be returned to Aerospace Research Laboratories unless return is required by security considerations, contractual obligations or notices on a specified document.
SIMULTANEOUS TESTS FOR TREND AND SERIAL CORRELATIONS FOR GAUSSIAN MARKOV RESIDUALS

BY P. R. KRISHNAIAH AND V. K. MURTHY

In this paper exact tests are proposed for testing the trend in the presence of autocorrelation, also for testing the trend and autocorrelation simultaneously in a first order Markov process. Also, the simultaneous confidence intervals associated with these tests are derived. These results are extended to a higher order Markov process.

SUMMARY

By extending a result of Ogawara [9], the problem of estimating the trend parameters when the residuals are serially correlated according to an hth order stationary Gaussian Markov process is reduced to the classical case where the residuals are uncorrelated and hence independent in view of normality. In this paper the problems of testing for the trend and other simultaneous hypotheses on the parameters are considered when the residuals are respectively a first and a general hth order Markov process of normal variates. Exact tests are obtained for testing the trend in the presence of autocorrelation and also for testing the trend and autocorrelation simultaneously in a first order Markov process. The simultaneous confidence bounds associated with these tests are also derived. These results are extended to hth order Markov process. For a general stationary Markov process, exact tests are proposed for testing the hypotheses of (i) no trend in the presence of serial correlations, (ii) independence among errors, and (iii) independence and no trend among errors. The critical values associated with these tests can be obtained by using the tables of the multivariate F distribution.

1. INTRODUCTION AND PRELIMINARY LEMMAS

Let n be a discrete time parameter. Let \( \{X_n\} \) be a discrete stochastic process of normal variates and suppose

\[
X_n = \mu_n + \varepsilon_n,
\]

where \( \mu_n \) is a non-random function of the parameter \( n \) and \( \{\varepsilon_n\} \) is a Gaussian stationary Markov process of order \( h \) with zero mean value. Equation (1.1) is called a model for fitting trend if \( \mu_n \) is a function of the discrete time parameter \( n \) only. On the other hand (1.1) is called a multiple regression model if

\[
\mu_n = \alpha + \beta_1 \xi_{1n} + \beta_2 \xi_{2n} + \ldots + \beta_p \xi_{pn},
\]

where the \( \xi \)'s are fixed variables, generally referred to as independent variables. The following two lemmas which we need in our further work are direct extensions to
the case of a general \( \mu_n \) of the corresponding lemmas given by Ogawara [9]. The proofs are omitted as they are essentially contained in reference [9].

**Lemma 1:** For \( \{X_n\} \) given by (1.1), let

\[
\begin{align*}
\mathbb{E}(e_n) &= 0, \\
\text{Var}(e_n) &= \sigma^2, \\
\text{Cov}(e_n, e_{n+1}) &= \sigma^2 \rho^k, \quad \rho < 1.
\end{align*}
\]

The \( e_n \)'s, stationary, Gaussian, and satisfying (1.3), are called a Gaussian stationary Markov process of order one. Then the conditional random variables \( \{X_K/X_{K-1}, X_{K+1}\} \). \( K = 1, 2, \ldots, m \) are independently normally distributed random variables with conditional expectation and variance given by

\[
E_c(X_{2K}) = \mu_{2K} - b \frac{\mu_{2K-1} + \mu_{2K+1}}{2} + b \frac{X_{2K-1} + X_{2K+1}}{2},
\]

and

\[
\text{Var}_c(X_{2K}) = \sigma_0^2,
\]

where

\[
b = \frac{2\rho}{1 + \rho^2},
\]

\[
\sigma_0^2 = \frac{\sigma^2(1 - \rho^2)}{1 + \rho^2}.
\]

**Lemma 2:** For \( \{X_n\} \) given by (1.1), let \( \{e_n\} \) be a stationary Gaussian Markov process of order \( h \), i.e., the autocorrelation function \( \rho_K \) satisfies the finite difference equation

\[
\rho_K + a_1 \rho_{K-1} + \ldots + a_K \rho_{K-h} = 0, \quad K = 1, 2, \ldots, a_h \neq 0
\]

and the \( a \)'s are such that the roots of the equation

\[
z^h + a_{h-1}z^{h-1} + \ldots + a_1 z + a_0 = 0
\]

all lie within the unit circle. Then the conditional random variables \( \{X_{K(K+1)}/X_{K(K+1)} - p, X_{K(K+1)} + p, p = 1, 2, \ldots, h\} \). \( K = 1, 2, \ldots, m \), are independently normally distributed with conditional expectation and variance given by

\[
E_c(X_{K(K+1)}) = \mu_{2K} - \sum_{p=1}^{h} b_p \frac{\mu_{K(K+1)} - p + \mu_{K(K+1)} + p}{2}
\]

\[
+ \sum_{p=1}^{h} b_p \frac{X_{K(K+1)} - p + X_{K(K+1)} + p}{2},
\]
and

\[ \text{Var}_c(X_{K+1}) = \sigma_0^2, \]

where \( \{h_p\}, p = 1, 2, \ldots, h, \) are given by

\[
\begin{bmatrix}
    b_h \\
    b_{h-1} \\
    \vdots \\
    b_1 \\
    b_0
\end{bmatrix}
= 2
\begin{bmatrix}
    1 & \rho_{h-1} & \rho_{h+1} & \cdots & \rho_{h+h} \\
    \rho_{h-1} & 1 & \rho_{h+1} & \cdots & \rho_{h+2} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    \rho_{h-1} & \rho_{h+1} & \cdots & 1 & \rho_{h+2} \\
    \rho_{h} & \rho_{h+1} & \cdots & \rho_{h+2} & 1
\end{bmatrix}^{-1}
\begin{bmatrix}
    \rho_0 \\
    \rho_1 \\
    \vdots \\
    \rho_h \\
    \rho_{h+1}
\end{bmatrix},
\]

and

\[ \sigma_0^2 = \frac{1 + a_1 \rho_1 + \cdots + a_h \rho_h}{1 + a_1^2 + \cdots + a_h^2}. \]

2. Exact Tests for Linear Trend and Serial Correlation in First Order Markov Process

Consider

\[ X_n = \mu_n + \varepsilon_n, \]

where

\[ \mu_n = \alpha + \beta n, \]

and the stationary process \( \{\varepsilon_n\} \) of normal variates satisfies (1.3). (2.1) and (2.2) constitute the model for linear trend when the residuals \( \{\varepsilon_n\} \) are no longer independent but form a stationary Gaussian Markov process of order one. Substituting (2.2) in (2.4) we obtain for the conditional expectation and variance of \( X_{2K} \) given \( X_{2K-1} \) and \( X_{2K+1} \) the following:

\[ E_c(X_{2K}) = \beta_1 + \beta_2 K + \beta_3 X_K', \quad K = 1, 2, \ldots, m; \]

where

\[
\begin{align*}
\beta_1 &= \alpha (1 - b), \\
\beta_2 &= 2\beta(1 - b), \\
\beta_3 &= \frac{2\rho}{1 + \rho^2}, \\
X_k' &= \frac{X_{2K-1} + X_{2K+1}}{2},
\end{align*}
\]
and

\begin{equation}
\text{Var}_c(X_{i,k}) = \frac{\sigma^2 (1 - \rho^2)}{1 + \rho^2}.
\end{equation}

Equations (2.3)-(2.5) constitute the usual regression model for independent
residuals with \( \beta_1 \) and \( \beta_2 \) corresponding to fitting a linear trend and \( \beta_3 \), which
vanishes if and only if \( \rho = 0 \), corresponding to fitting the fixed variate \( X_k \).

Let us now consider the problem of testing for trend in the original model given
by (2.1) and (2.2). Let \( H_0 \) denote the hypothesis that there is no linear trend in the
model (2.1), i.e.,

\begin{equation}
H_0: \quad \alpha = 0, \quad \beta = 0.
\end{equation}

Let

\begin{equation}
H_{01}: \quad \alpha = 0
\end{equation}

and

\begin{equation}
H_{02}: \quad \beta = 0.
\end{equation}

Then \( H_0 = H_{01} \cap H_{02} \). Let \( H_{03} \) denote the hypothesis that the stationary Gaussian
process \( \{e_n\} \) is a process of independent variates. In other words \( H_{03}: \quad \rho = 0 \).

Using the symbol \( \Rightarrow \) to denote “implies and is implied by” it is clear from (2.4) that

\begin{equation}
\begin{align*}
\alpha = 0 & \Rightarrow \beta_1 = 0, \\
\beta = 0 & \Rightarrow \beta_2 = 0, \\
\rho = 0 & \Rightarrow \beta_3 = 0,
\end{align*}
\end{equation}

since \( b \neq 1 \). So, the hypothesis \( H_0 \) of “no linear trend” in the given model (2.1) is
equivalent to the hypothesis that \( \beta_1 = 0 \) and \( \beta_2 = 0 \) in the conditional model (2.3)
which is a classical regression model with independent residuals. Now, let \( \hat{\beta}_i \) be the
least squares estimate of \( \beta_i \) under the model (2.3).

It is known from normal regression theory that the joint distribution of \( \hat{\beta}_1, \hat{\beta}_2, \) and \( \hat{\beta}_3 \) is trivariate normal with means \( \beta_1, \beta_2, \) and \( \beta_3 \) and the covariance matrix
\( \Sigma = \sigma^2(c_{ij}) \) where \( \sigma^2 \) is given in (1.7) and \( (c_{ij}) \) denotes the inverse of the matrix of
the coefficients of normal equations used in estimating \( \beta_1, \beta_2, \) and \( \beta_3 \). Now, let
\( H_{0i}: \beta_i = 0 \) for \( i = 1, 2, 3 \) and let \( S^2_e \) be the sum of squares due to deviations from
regression (2.6). In addition, let

\begin{equation}
F_i = \frac{(m-3)\hat{\beta}_i^2}{c_{ii}S^2_e}, \quad i = 1, 2, 3.
\end{equation}

When \( H_0 \) is true, \( \hat{\beta}_1^2/c_{11}\sigma^2 \) and \( \hat{\beta}_2^2/c_{22}\sigma^2 \) are jointly distributed as a bivariate chi-
square distribution with 1 degree of freedom. Also, \( S^2_e/\sigma^2 \) is another chi-square
variate with \( m - 3 \) degrees of freedom distributed independently of \( \hat{\beta}_1^2 \) and \( \hat{\beta}_2^2 \).
When \( H_0 \) is true, we know from [4,5] that the joint distribution of \( F_1 \) and \( F_2 \) (holding \( X_{2:k-1} \) and \( X_{2:k+1} \) fixed) is given by

\[
f_2(F_1, F_2) = \frac{(m-3)^{m-3n/2}(1-\rho_1^2)^{m-2n/2}}{\sqrt{\Pi 1'(m-3)/2}}
\]

(2.7)

\[ \sum_{i=0}^{\infty} \frac{(p_{12})^{2i}}{i! \Gamma(i+1)\left[(m-3)(1-\rho_1^2)+F_1+F_2\right]^{(i-1)/2} + 2i} \]

where \( \rho_{12} = \frac{c_{12}(c_{12})^4} {c_{42}(c_{22})^4} \). Here we note that \( \rho_{12} \) depends upon \( K \) and the fixed variables \( X_{2:k-1} \) and \( X_{2:k+1} \) only. The rule for the testing \( H_{01}, H_{02}, \) and \( H_0 \) simultaneously is described below.

Accept or reject \( H_{0i} (i = 1, 2) \) according as \( F_i \leq F_i \) where \( F_i \) is chosen such that

\[
P(F_i \leq F_i; \ i = 1, 2 | H_0) = (1 - \alpha) \quad (2.8)
\]

The total hypothesis \( H_0 \) is accepted if and only if \( H_{01} \) and \( H_{02} \) are accepted. The critical values \( F_i \) can be obtained by using the tables of the bivariate \( F \) distribution with \( (1, n) \) degrees of freedom. The simultaneous confidence intervals associated with the above test are given by

\[
P(|\beta_i - \beta_i| \leq \sqrt{F_i c_i S_i^2/(m-3)}; \ i = 1, 2) = (1 - \alpha) \quad (2.9)
\]

where \( F_i \) is given by (2.8). The above simultaneous confidence intervals based upon the distribution of (2.7) are valid since our test procedure is associated with the conditional model (2.3). They can be derived by using the fact that

\[
P(F_i \leq F_i; \ i = 1, 2 | \bar{A}) = P(F_i \leq F_i; \ i = 1, 2 | H_0)
\]

where

\[
F_i^* = \frac{(\hat{\beta}_i - \beta_0)^2 (m-3)}{c_i S_i^2} \quad A_i; \ \beta_i \neq 0 \text{ and } A = \bigcap_{i=1}^{2} A_i.
\]

If we are interested in testing \( H_{01}, H_{02}, \) and \( H_{03} \) simultaneously, the procedure is to accept or reject \( H_{0i} (i = 1, 2, 3) \) according as \( F_i \leq F_i \) where \( F_i \) is chosen such that

\[
P(F_i \leq F_i; \ i = 1, 2, 3 | \bigcap_{i=1}^{3} H_{0i}) = (1 - \alpha) \quad (2.10)
\]

The total hypothesis \( \bigcap_{i=1}^{3} H_{0i} \), that is, the hypothesis of no trend and no serial correlation, is accepted if and only if the individual hypotheses \( H_{01}, H_{02}, \) and \( H_{03} \) are accepted. The joint distribution of \( F_1, F_2, \) and \( F_3, \) when \( \bigcap_{i=1}^{3} H_{0i} \) is true, is the trivariate central \( F \) distribution with \( (1, m-3) \) degrees of freedom. For a detailed discussion of the multivariate \( F \) distribution, the reader is referred to [6,7]. Krishi-
naiah and Major Armitage are constructing tables for the percentage points of the multivariate \( F \) distribution with \((1, n)\) degrees of freedom. The simultaneous confidence bounds associated with this test are given by

\[
P_i \left[ \hat{\theta}_i - \theta_i \right] \leq \sqrt{n} \sum_{i=1}^{m} F_{a_i} \left( m - 3 ; i = 1, 2, 3 \right) = (1 - \alpha)
\]

where \( F_{a_i} \) is chosen satisfying (2.10).

3. SIMULTANEOUS TESTS IN 4TH ORDER MARKOV PROCESS

Consider the model

\[
X_n = \mu_n + e_n,
\]

where

\[
\mu_n = \alpha + \beta_n,
\]

and the stationary process \(\{e_n\}\) of normal variates satisfies the conditions of Lemma 2. Now, substituting (3.2) in (1.10), we obtain the following expressions for the conditional expectation and variance of \(X_{(h+1)}\) given \(X_{(h+1)}\), \(X_{(h+1)}\), \(X_{(h+1)}\), \(X_{(h+1)}\) \((\rho = 1, 2, \ldots, h)\).

\[
E_i(X_{(h+1)}) = \beta_1 + \beta_2 K + \sum_{\rho=3}^{h+2} \beta_\rho X'_{\rho, K},
\]

\[
\text{Var}_i(X_{(h+1)}) = \sigma^2.
\]

where

\[
X'_{\rho, K} = \frac{X_{(h+1)} - \rho - X_{(h+1)} + \rho}{2},
\]

\[
\beta_1 = \alpha \left(1 - \sum_{\rho=1}^{h} b_\rho\right),
\]

\[
\beta_2 = \beta_2 \left[2 - (h+1) \sum_{\rho=1}^{h} b_\rho\right],
\]

\[
\beta_3 = b_1,
\]

\[
\ldots
\]

\[
\beta_{h+2} = b_h,
\]

and the \( b \)'s and \( \sigma^2 \) are given by (1.12) and (1.13) respectively.

Now, let

\[
H_0:\quad \beta_i = 0, \quad i = 1, 2, \ldots, h+2.
\]
Also, let

\[ F_i = \frac{(m-h-2)\hat{\beta}_i^2}{c_i S^2_i}, \quad i = 1, 2, \ldots, (h+2), \]

where \( \hat{\beta}_i \) is the least square estimate of \( \beta_i \) in the model (3.3), \( m \) is the size of the sample, \( S^2_i \) is the sum of squares due to deviation from the regression (3.3), and \( (c_i)_{(h+2) \times (h+2)} \) is the inverse of the matrix of the coefficients of normal equations used in estimating \( \beta_i \)'s. From classical regression theory, it is known that

\[ (c_{ij}) = \left( \sum_{k=1}^{m} Y_k Y_k \right)^{-1} \]

where \( Y_k = (1, K, X_{3,k}, \ldots, X_{n+2,k}) \) and \( Y_k^t \) is the transpose of \( Y_k \).

We shall now consider the problem of testing \( H_{01} \) and \( H_{02} \) simultaneously.

The hypothesis \( H_{01} : (i = 1, 2) \) is accepted or rejected according as \( F_i \leq F_a \) where \( F_a \) is chosen such that

\[ P \left[ F_i \leq F_a ; \quad i = 1, 2 \right] \bigwedge_{i=1}^{2} H_{0i} = (1 - \alpha). \tag{3.6} \]

Here we note that \( H_{01} \) and \( H_{02} \) are respectively equivalent to the hypotheses \( \alpha = 0 \) and \( \beta = 0 \). When \( \bigwedge_{i=1}^{2} H_{0i} \) is true, \( \hat{\beta}_1^2 c_1^2 \sigma^2_0 \) and \( \hat{\beta}_2^2 c_2^2 \sigma^2_0 \) are jointly distributed as a bivariate chi-square distribution with 1 degree of freedom. Also, \( S^2_k \) is another chi-square variate with \( (m-h-2) \) degrees of freedom distributed independently of \( \hat{\beta}_1^2 \) and \( \hat{\beta}_2^2 \). When \( H_{01} \cap H_{02} \) is true, we know from [4,5] that the joint distribution of \( F_1 \) and \( F_2 \) (holding \( X_{K+1,1+p} \) and \( X_{K+1,1+p} \) fixed for \( p = 1, 2, \ldots, h \)) is given by

\[ f_c(F_1, F_2) = \frac{(m-h-2)^{m-h-2} \Gamma((m-h-1)/2)}{\Gamma(m-h-1/2)^2} \frac{\sum_{i=0}^{\infty} \frac{(p_{12})^{2i} t^{2i}}{i!} \Gamma(i+1/2) \Gamma(i+1/2) \Gamma((m-h-2)(1-p_{12})+F_1+F_2)^{m-h-2}}{(m-h-2)(1-p_{12})+F_1+F_2} \]

where \( p_{12} = c_1^2 c_2^2 (c_1^2 c_2^2) \). Here \( p_{12} \) depends upon \( K \) and the fixed variates \( X_{K+1,1+p} \) and \( X_{K+1,1+p} \). The simultaneous confidence intervals associated with the above test are

\[ P \left[ \left| \hat{\beta}_i - \beta_i \right| \leq \sqrt{\frac{c_i \sigma_0^2 \sigma^2_0}{(m-h-2)}} F_a ; \quad i = 1, 2 \right] = (1-\alpha) \tag{3.8} \]

where \( F_a \) is chosen satisfying (3.6).

The procedure for testing \( H_{01}, \ldots, H_{0,h+2} \) simultaneously is as follows: Accept
or reject $H_{i0}, i = 1, 2, \ldots, h + 2$, according as $F_i \leq F$, where $F_i$ is chosen such that

$$P[F_i \leq F; \ i = 1, 2, \ldots, h + 2; \ \bigcap_{i=1}^{h+2} H_{i0}] = (1 - \alpha).$$

(3.9)

When $\bigcap_{i} H_{i0}$ is true, the joint distribution of $F_1, \ldots, F_{h+2}$ is a central $(h + 2)$ variate $F$ distribution with $(1, m - h - 2)$ degrees of freedom. So, the critical values $F_i$ can be obtained from [7]. The simultaneous confidence intervals associated with the above test are

$$P[R_i - \beta_i \leq \sqrt{F_{i0}S_i^2(m - h - 2); \ \ i = 1, 2, \ldots, h + 2}] = (1 - \alpha)$$

where $F_i$ is chosen satisfying (3.9). Here we note that the hypothesis $\mu_1 = \ldots = \mu_h = 0$ is equivalent to the hypothesis $\bigcap_{i=1}^{h+2} H_{i0}$.

If one is interested in testing just the hypotheses, $\bigcap_{i=3}^{h+2} H_{i0}, H_{03}, \ldots, H_{0,h+2}$ simultaneously, the critical value $F_i$ should be chosen such that

$$P[F_i \leq F; \ i = 3, \ldots, h + 2; \ \bigcap_{i=3}^{h+2} H_{i0}] = (1 - \alpha).$$

4. GENERAL REMARKS

Under model (2,1), one can use the more usual $F$ statistic with $(2, m - 3)$ degrees of freedom to test $H_{01}$ and $H_{02}$ (defined in Section 2) simultaneously. The $F$ statistic in this case (using the notation of Section 2) is

$$F = (\hat{\beta}_1, \hat{\beta}_2) \left[ \begin{array}{cc} C_{11} & C_{12} \\ C_{21} & C_{22} \end{array} \right]^{-1} \left[ \begin{array}{c} \hat{\beta}_1 \\ \hat{\beta}_2 \end{array} \right] \left[ \begin{array}{c} m - 3 \\ 2 \end{array} \right] S_i^{-1}$$

But using the methods in [4,5], it is seen that the lengths of the confidence intervals associated with the simultaneous test procedure proposed in Section 2 are shorter than the lengths of the corresponding confidence intervals associated with the overall $F$ test procedure. Similar remarks can be made for testing $H_{01}, H_{02},$ and $H_{03}$ simultaneously under the model (2.1) and for testing various hypotheses simultaneously under the model (3.1). The authors are not aware of any other alternative test procedures for testing various hypotheses simultaneously under the models (2.1) and (3.1). The optimum properties of the powers of the simultaneous test procedures considered in this paper are under investigation.

Sometimes the experimenter may be interested in testing individual hypotheses separately. For the first order Markov process, the test procedure in this case is to accept or reject $H_i$ for any given $i$, according as $F_i \leq F_i^{*}$, where $F_i$ and $H_i$ are defined in Section 2 and $F_i^{*}$ is chosen such that

$$P[F_i \leq F_i^{*}|H_i] = (1 - \alpha).$$
The associated confidence bound is given by

\[ P [ | \beta_j - \beta_j | \leq \sqrt{F^*_\alpha \cdot c_{n_n} S^2 / (m - 3)} ] = (1 - \alpha). \]

Here we note that \( F^*_\alpha \) is the upper \( \alpha \) per cent value of the central \( F \) distribution with \((1, m-3)\) degrees of freedom. Similar test procedures can be proposed for the \( h \)th order Markov process.

Aerospace Research Laboratories, Wright-Patterson AFB, Ohio and the Douglas Aircraft Company, California

REFERENCES


