

ARL 66-0214
OCTOBER 1966



Aerospace Research Laboratories

SIMULTANEOUS TESTS FOR TREND AND SERIAL CORRELATIONS
FOR GAUSSIAN MARKOV RESIDUALS

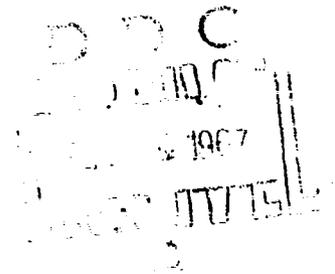
(Reprint from *Econometrica*, Vol. 34, No. 2, April 1966)

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



SIMULTANEOUS TESTS FOR TREND AND SERIAL CORRELATIONS
FOR GAUSSIAN MARKOV RESIDUALS

BY P. R. KRISHNAIAH AND V. K. MURTHY¹

In this paper exact tests are proposed for testing the trend in the presence of autocorrelation and 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 h th 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 h th 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 h th 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

$$(1.1) \quad X_n = \mu_n + \varepsilon_n,$$

where μ_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 μ_n is a function of the discrete time parameter n only. On the other hand (1.1) is called a multiple regression model if

$$(1.2) \quad \mu_n = \alpha + \beta_1 \xi_{1n} + \beta_2 \xi_{2n} + \dots + \beta_r \xi_{rn},$$

where the ξ '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

¹ The authors are grateful to the referee and Professor M. M. Rao for their valuable suggestions.

the case of a general μ_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

$$(1.3) \quad \begin{cases} E(\varepsilon_n) = 0, \\ \text{Var}(\varepsilon_n) = \sigma^2, \\ \text{COV}(\varepsilon_n, \varepsilon_{n+K}) = \sigma^2 \rho^K, \quad \rho < 1. \end{cases}$$

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

$$(1.4) \quad E_c(X_{2K}) = \mu_{2K} - b \frac{\mu_{2K-1} + \mu_{2K+1}}{2} + b \frac{X_{2K-1} + X_{2K+1}}{2},$$

and

$$(1.5) \quad \text{Var}_c(X_{2K}) = \sigma_0^2,$$

where

$$(1.6) \quad b = \frac{2\rho}{1+\rho^2},$$

$$(1.7) \quad \sigma_0^2 = \frac{\sigma^2(1-\rho^2)}{1+\rho^2}.$$

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

$$(1.8) \quad \rho_K + a_1 \rho_{K-1} + \dots + a_h \rho_{K-h} = 0, \quad K = 1, 2, \dots; \quad a_h \neq 0$$

and the a 's are such that the roots of the equation

$$(1.9) \quad z^h + a_1 z^{h-1} + \dots + a_{h-1} z + a_h = 0$$

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

$$(1.10) \quad E_c(X_{K(h+1)}) = \mu_{2K} - \sum_{p=1}^h b_p \frac{\mu_{K(h+1)-p} + \mu_{K(h+1)+p}}{2} + \sum_{p=1}^h b_p \frac{X_{K(h+1)-p} + X_{K(h+1)+p}}{2},$$

and

$$(1.11) \quad \text{Var}_c(X_{K(h+1)}) = \sigma_0^2,$$

where $\{b_p\}$, $p = 1, 2, \dots, h$, are given by

$$(1.12) \quad \begin{bmatrix} b_h \\ b_{h-1} \\ \cdot \\ \cdot \\ b_1 \\ b_1 \\ \cdot \\ \cdot \\ b_h \end{bmatrix} = 2 \begin{bmatrix} 1 & \dots & \rho_{h-1} & \rho_{h+1} & \dots & \rho_{2h} \\ \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ \rho_{h-1} & \dots & 1 & \rho_2 & \dots & \rho_{h+1} \\ \rho_{h+1} & \dots & \rho_2 & 1 & \dots & \rho_{h-1} \\ \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ \rho_{2h} & \dots & \rho_{h+1} & \rho_{h-1} & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_h \\ \rho_{h-1} \\ \cdot \\ \cdot \\ \rho_1 \\ \rho_1 \\ \cdot \\ \cdot \\ \rho_h \end{bmatrix},$$

and

$$(1.13) \quad \sigma_0^2 = \frac{1 + a_1 \rho_1 + \dots + a_h \rho_h}{1 + a_1^2 + \dots + a_h^2} \sigma^2.$$

2. EXACT TESTS FOR LINEAR TREND AND SERIAL CORRELATION IN FIRST ORDER MARKOV PROCESS

Consider

$$(2.1) \quad X_n = \mu_n + \varepsilon_n,$$

where

$$(2.2) \quad \mu_n = \alpha + \beta n,$$

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 (1.4) we obtain for the conditional expectation and variance of X_{2K} given X_{2K-1} and X_{2K+1} the following:

$$(2.3) \quad E_c(X_{2K}) = \beta_1 + \beta_2 K + \beta_3 X'_K, \quad K = 1, 2, \dots, m;$$

where

$$(2.4) \quad \begin{cases} \beta_1 = \alpha(1-b), \\ \beta_2 = 2\beta(1-b), \\ \beta_3 = b = \frac{2\rho}{1+\rho^2}, \end{cases}$$

$$X'_K = \frac{X_{2K-1} + X_{2K+1}}{2},$$

and

$$(2.5) \quad \text{Var}_c(X_{2k}) = \frac{\sigma^2(1-\rho^2)}{1+\rho^2}.$$

Equations (2.3)-(2.5) constitute the usual regression model for independent residuals with β_1 and β_2 corresponding to fitting a linear trend and β_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.,

$$H_0: \begin{matrix} \alpha = 0, \\ \beta = 0. \end{matrix}$$

Let

$$H_{01}: \alpha = 0$$

and

$$H_{02}: \beta = 0.$$

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 $H_{03}: \rho = 0$.

Using the symbol \Leftrightarrow to denote "implies and is implied by" it is clear from (2.4) that

$$(2.6) \quad \begin{matrix} \alpha = 0 \Leftrightarrow \beta_1 = 0, \\ \beta = 0 \Leftrightarrow \beta_2 = 0, \\ \rho = 0 \Leftrightarrow \beta_3 = 0. \end{matrix}$$

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 β_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 β_1, β_2 , and β_3 and the covariance matrix $\Sigma = \sigma_0^2(c_{ij})$ where σ_0^2 is given in (1.7) and (c_{ij}) denotes the inverse of the matrix of the coefficients of normal equations used in estimating β_1, β_2 , and β_3 . Now, let $H_{0i}: \beta_i = 0$ for $i = 1, 2, 3$ and let S_E^2 be the sum of squares due to deviations from regression (2.6). In addition, let

$$F_i = \frac{(m-3)\hat{\beta}_i^2}{c_{ii}S_E^2}, \quad i = 1, 2, 3.$$

When H_0 is true, $\hat{\beta}_1^2/c_{11}\sigma_0^2$ and $\hat{\beta}_2^2/c_{22}\sigma_0^2$ are jointly distributed as a bivariate chi-square distribution with 1 degree of freedom. Also, S_E^2/σ_0^2 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

$$(2.7) \quad f_c(F_1, F_2) = \frac{(m-3)^{(m-3)/2} (1-\rho_{12}^2)^{(m-2)/2}}{\sqrt{\pi} \Gamma(m-3)/2} \cdot \sum_{i=0}^{\infty} \frac{(\rho_{12})^{2i}}{i! \Gamma(i+1) [(m-3)(1-\rho_{12}^2) + F_1 + F_2]^{\frac{(m-1)}{2} + 2i}}$$

where $\rho_{12} = c_{12}/(c_{11}c_{22})^{1/2}$. Here we note that ρ_{12} depends upon K and the fixed variates $X_{2,K-1}$ 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 $F_i \leq F_\alpha$ where F_α is chosen such that

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

The total hypothesis H_0 is accepted if and only if H_{01} and H_{02} are accepted. The critical values F_α 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

$$(2.9) \quad P\{|\beta_i - \beta_j| \leq \sqrt{F_\alpha c_{ii} S_{jj}^2 / (m-3)}; i = 1, 2\} = (1-\alpha)$$

where F_α 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

$$P[F_i^* \leq F_\alpha; i = 1, 2 | A] = P[F_i \leq F_\alpha; i = 1, 2 | H_0]$$

where

$$F_i^* = \frac{(\beta_i - \beta_j)^2 (m-3)}{c_{ii} S_{jj}^2}, \quad A_i: \beta_i \neq 0 \text{ and } A = \bigcap_{i=1}^2 A_i.$$

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 $F_i \leq F_\alpha$ where F_α is chosen such that

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

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 $\bigcap H_{0i}$ 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]. Krish-

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

$$(2.11) \quad P[|\hat{\beta}_i - \beta_i| \leq \sqrt{c_{ii} S_i^2 F_{\alpha} / (m-3)}; i = 1, 2, 3] = (1-\alpha)$$

where F_{α} is chosen satisfying (2.10).

3. SIMULTANEOUS TESTS IN h TH ORDER MARKOV PROCESS

Consider the model

$$(3.1) \quad X_n = \mu_n + \varepsilon_n,$$

where

$$(3.2) \quad \mu_n = \alpha + \beta_n,$$

and the stationary process $\{\varepsilon_n\}$ 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 $X_{K(h+1)}$ given $X_{K(h+1)-p}, \dots, X_{K(h+1)+p}$ ($p = 1, 2, \dots, h$).

$$(3.3) \quad E_c(X_{K(h+1)}) = \beta_1 + \beta_2 K + \sum_{p=3}^{h+2} \beta_p X'_{p,K},$$

$$(3.4) \quad \text{Var}_c(X_{K(h+1)}) = \sigma_0^2,$$

where

$$(3.5) \quad \begin{cases} X'_{p,K} = \frac{X_{K(h+1)-p} + X_{K(h+1)+p}}{2}, \\ \beta_1 = \alpha \left(1 - \sum_{p=1}^h b_p\right), \\ \beta_2 = \beta \left[2 - (h+1) \sum_{p=1}^h b_p\right], \\ \beta_{h+1} = b_1, \\ \dots \\ \dots \\ \dots \\ \beta_{h+2} = b_h, \end{cases}$$

and the b 's and σ_0^2 are given by (1.12) and (1.13) respectively.

Now, let

$$H_{0i}: \beta_i = 0, \quad i = 1, 2, \dots, h+2.$$

Also, let

$$F_i = \frac{(m-h-2)\beta_i^2}{c_{ii}S_h^2}, \quad i = 1, 2, \dots, (h+2),$$

where β_i is the least square estimate of β_i in the model (3.3), m is the size of the sample, S_h^2 is the sum of squares due to deviation from the regression (3.3), and $(c_{ij}) : (h+2) \times (h+2)$ is the inverse of the matrix of the coefficients of normal equations used in estimating β 's. From classical regression theory, it is known that

$$(c_{ij}) = \left(\sum_{K=1}^m \underline{Y}_K \underline{Y}_K' \right)^{-1}$$

where $\underline{Y}_K' = (1, K, X'_{3,K}, \dots, X'_{k+2,K})$ and \underline{Y}_K is the transpose of \underline{Y}_K' .

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

The hypothesis H_{0i} ($i = 1, 2$) is accepted or rejected according as $F_i \leq F_\alpha$ where F_α is chosen such that

$$(3.6) \quad P[F_i \leq F_\alpha; i = 1, 2 | \bigcap_{i=1}^2 H_{0i}] = (1-\alpha).$$

Here we note that H_{01} and H_{02} are respectively equivalent to the hypotheses $\alpha = 0$ and $\beta = 0$. When $\bigcap_{i=1}^2 H_{0i}$ is true, $\beta_1^2/c_{11}\sigma_0^2$ and $\beta_2^2/c_{22}\sigma_0^2$ are jointly distributed as a bivariate chi-square distribution with 1 degree of freedom. Also, S_h^2 is another chi-square variate with $(m-h-2)$ degrees of freedom distributed independently of β_1^2 and β_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 $X_{K(h+1)+p}$ and $X_{K(h+1)-p}$ fixed for $p = 1, 2, \dots, h$) is given by

$$(3.7) \quad f_c(F_1, F_2) = \frac{(m-h-2)^{(m-h-2)/2} (1-\rho_{12}^2)^{(m-h-1)/2}}{\sqrt{\pi} \Gamma(m-h-1)/2} \\ \cdot \sum_{i=0}^{\infty} \frac{(\rho_{12})^{2i} \Gamma\left[2i + \frac{(m-h)}{2}\right] (F_1 F_2)^{i-1}}{i! \Gamma(i+\frac{1}{2}) [(m-h-2)(1-\rho_{12}^2) + F_1 + F_2] \frac{m-h}{2} + 2i}$$

where $\rho_{12} = c_{12}/(c_{11}c_{22})^{1/2}$. Here ρ_{12} depends upon K and the fixed variates $X_{K(h+1)+p}$ and $X_{K(h+1)-p}$, $p = 1, 2, \dots, h$. The simultaneous confidence intervals associated with the above test are

$$(3.8) \quad P\left[|\hat{\beta}_i - \beta_i| \leq \sqrt{\frac{c_{ii}\sigma_0^2 S_h^2}{(m-h-2)} F_\alpha}; i = 1, 2\right] = (1-\alpha)$$

where F_α is chosen satisfying (3.6).

The procedure for testing $H_{01}, \dots, H_{0, h+2}$ simultaneously is as follows: Accept

or reject H_{0i} , $i = 1, 2, \dots, h+2$, according as $F_i \leq F_\alpha$, where F_α is chosen such that

$$(3.9) \quad P[F_i \leq F_\alpha; i = 1, 2, \dots, h+2 | \bigcap_{i=1}^{h+2} H_{0i}] = (1-\alpha).$$

When $\bigcap_{i=1}^{h+2} H_{0i}$ is true, the joint distribution of F_1, \dots, F_{h+2} is a central $(h+2)$ variate F distribution with $(1, m-h-2)$ degrees of freedom. So, the critical values F_α can be obtained from [7]. The simultaneous confidence intervals associated with the above test are

$$P[|\hat{\beta}_i - \beta_i| \leq \sqrt{F_\alpha c_{ii} S_i^2 / (m-h-2)}; i = 1, 2, \dots, h+2] = (1-\alpha)$$

where F_α is chosen satisfying (3.9). Here we note that the hypothesis $\rho_1 = \dots = \rho_h = 0$ is equivalent to the hypothesis $\bigcap_{i=1}^{h+2} H_{0i}$.

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

$$P[F_i \leq F_\alpha; i = 3, \dots, h+2 | \bigcap_{i=3}^{h+2} H_{0i}] = (1-\alpha).$$

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

$$F = (\hat{\beta}_1, \hat{\beta}_2) \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}^{-1} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} (m-3) / 2 S_i^2.$$

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 power functions 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 $F_i \leq F_\alpha^*$ where F_i and H_i are defined in Section 2 and F_α^* is chosen such that

$$P[F_i \leq F_\alpha^* | H_i] = (1-\alpha).$$

The associated confidence bound is given by

$$P[|\beta_i - \beta_{il}| \leq \sqrt{F_{\alpha}^* c_{ii} S_E^2 / (m-3)}] = (1-\alpha).$$

Here we note that F_{α}^* is the upper α 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

- [1] ARMITAGE, J. V., AND P. R. KRISHNAIAH: *Tables for the Studentized Largest Chi-Square Distribution and Their Applications*, ARL 64-188, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio (1964).
- [2] GHOSH, M. N.: "Simultaneous Tests of Linear Hypotheses," *Biometrika*, Vol. 42 (1955), pp. 441-449.
- [3] JOHN, S.: "On the Evaluation of the Probability Integral of the Multivariate t -distribution," *Biometrika*, Vol. 48 (1961), pp. 409-417.
- [4] KRISHNAIAH, P. R.: *Simultaneous Tests and the Efficiency of Generalized Balanced Incomplete Block Designs*, ARL 63-174, Aerospace Research Laboratories, Wright-Patterson Air Force Base (1963).
- [5] ———: "On the Simultaneous ANOVA and MANOVA Tests," *Annals of the Institute of Statistical Mathematics*, Vol. 17 (1965).
- [6] ———: *Multiple Comparison Tests in Multivariate Case*, ARL 64-124, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio (1964).
- [7] KRISHNAIAH, P. R., AND J. V. ARMITAGE: "Probability Integrals of the Multivariate F Distribution with Tables for Special Cases." Presented at the Central Regional Meeting of the Institute of Mathematical Statistics held at Chicago, December, 1964.
- [8] MURTHY, V. K.: *On Fitting a Linear Trend and Testing Independence When the Residuals Form a Markov Process*, Technical Report No. 6, Applied Mathematics and Statistics Laboratory, Stanford University, Stanford, California (1960).
- [9] OGAWARA, M.: "A Note on the Test of Serial Correlation Coefficients," *Annals of Mathematical Statistics*, Vol. 22 (1953), pp. 115-118.