An Exact Test for Moving Averages

Ogawara, Masami
Meteorological Research Institute

https://doi.org/10.5109/12979

出版情報：統計数理研究. 7 (3/4), pp.77-83, 1957-03. Research Association of Statistical Sciences
バージョン：
権利関係：

AN EXACT TEST FOR MOVING AVERAGES

By

Masami OGAWARA

(*Received* *Jan.* 30, 1957)

§ **1. Introduction.** A stationary (wide sense) real process $x(t)$ with discrete parameter is called a process of *moving averages* (*m.a.*) of order $k(\geqq 1)$ if it has a form

$$x(t) = \mu + b_0 y(t) + b_1 y(t-1) + \cdots + b_k y(t-k), \quad b_0 b_k \neq 0, \tag{1.1}$$

where $y(t)$ is a non-autocorrelated stationary process with zero mean (WOLD [1]). The $x(t)$ is said to be a *regular m.a.* if

$$\begin{cases} \text{the roots of the equation} \\ b_0 z^k + b_1 z^{k-1} + \cdots + b_k = 0 \\ \text{are less than or equal to one in modulus.} \end{cases} \tag{1.2}$$

In the continuous parameter case a process of *m.a.* is defined by an expression of the form

$$x(t) = \int_{-\infty}^{\infty} f(s) d\xi(t-s), \tag{1.3}$$

where $f(s)$ is a Lebesgue measurable function, $\int_{-\infty}^{\infty} |f(s)|^2 ds < \infty$, and the $\xi(t)$ is a process of orthogonal increments, $E\{|d\xi(t)|^2\} = dt$ (DOOB [2]).

A large sample test for *m.a.* with discrete parameter was discussed by H. WOLD [3] in 1949. In this paper we shall first point out that an exact test for a discrete parameter normal *m.a.* (*d.p.n.m.a.*) is given by an application of the normal regression theory, and then we shall apply similar method to the continuous parameter *m.a.* (*c.p.m.a.*) under some restrictions.

A prediction scheme for the process of this type will be easily derived as well even if the size of sample is finite (OGAWARA [4]).

§ **2. Discrete Parameter Case.** For a *m.a.* (1.1), the (theoretical) autocorrelation coefficients are given by

$$\begin{aligned} \rho(\tau) &= \rho(-\tau) \\ &= \begin{cases} (b_0 b_\tau + b_1 b_{\tau+1} + \cdots + b_{k-\tau} b_k)/(b_0^2 + b_1^2 + \cdots + b_k^2) & \tau = 0, 1, \cdots, k \\ 0 & \tau \geqq k+1 \end{cases} \end{aligned} \tag{2.1}$$

Thus, for a *n.m.a.*, if $|t_1 - t_2| \geqq k+1$, $x(t_1)$ and $x(t_2)$ are independent. This fact leads us immediately to the following test schemes. Let $x_1, x_2, \cdots, x_N$

be a sample sequence of the *n.m.a.* of order k whose mean, variance and autocorrelation coefficients are μ, σ^2 and $\rho(\tau)$ respectively.

(i) *Test for the mean and the variance.* A subsequence $z_j \equiv x_{t_j}$ ($j = 1, 2, \cdots, n$), where $1 \leqq t_1 < t_2 < \cdots < t_n \leqq N$ and $t_{j+1} - t_j \geqq k+1$ ($j = 1, 2, \cdots, n-1$), forms a random sample from the normal population with the mean μ and the variance σ^2. Thus the ordinary test schemes can be applied.

(ii) *Test for the correlogram.* We consider here only a regular *m.a.*. In that case the coefficients b's are uniquely determined by the autocorrelation coefficients (theoretical correlogram), provided that we put $b_0 = 1$ without loss of generality (WOLD [1]). Consequently, the hypothesis specifing the $\rho(\tau)$ ($\tau = 1, 2, \cdots, k$), $H(\rho(1), \rho(2), \cdots, \rho(k))$, is equivalent to the hypothesis specifing the b_i ($i = 1, 2, \cdots, k$), $H(b_1, b_2, \cdots, b_k)$. This fact we denote as

$$H(\rho(1), \rho(2), \cdots, \rho(k)) \sim H(b_1, b_2, \cdots, b_k) \tag{2.1}$$

or simply $H(\rho) \sim H(b)$.

Now, if we put

$$\left.\begin{array}{ll} z_j \equiv x_{j(2k+1)+k+1} & \\ x'_{jp} \equiv x_{j(2k+1)+p} & p = 1, 2, \cdots, k \end{array}\right\} \quad j = 1, 2, \cdots, n,$$

$(k+1)$-variate normal random variables $(z_j, x'_{j1}, \cdots, x'_{jk})$ ($j = 1, 2, \cdots, n$) are independent with the consequence that the z_j ($j = 1, 2, \cdots, n$) are conditionally independent relatively to the x'_{jp} ($p = 1, 2, \cdots, k$; $j = 1, 2, \cdots, n$), and the conditional expectation of z_j turns out

$$E\{z_j | x'_{jp},\ \ p = 1, 2, \cdots, k\} = c_0 + \sum_{p=1}^{k} c_p x'_{jp},$$

where $c_0 = \mu(1 - c_1 - c_2 - \cdots - c_k)$ and

$$\rho(\tau) + c_1\rho(\tau-1) + \cdots + c_k\rho(\tau-k) = 0 \qquad (\tau = 1, 2, \cdots, k) \tag{2.2}$$

or

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = - \begin{bmatrix} 1 & \rho(1) & \cdots & \rho(k-1) \\ \rho(1) & 1 & \cdots & \rho(k-2) \\ \vdots & \vdots & & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(k) \end{bmatrix} \tag{2.3}$$

or inversely

$$\begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \vdots \\ \rho(k-1) \\ \rho(k) \end{bmatrix} = - \begin{bmatrix} c_2 - 1 & c_3 & c_4 & \cdots & c_k & 0 \\ c_1 + c_3 & c_4 - 1 & c_5 & \cdots & 0 & 0 \\ c_2 + c_4 & c_1 + c_5 & c_6 - 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ c_{k-2} + c_k & c_{k-3} & c_{k-4} & \cdots & -1 & 0 \\ c_{k-1} & c_{k-2} & c_{k-3} & \cdots & c_1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{k-1} \\ c_k \end{bmatrix} \tag{2.4}$$

provided that the inverse matrix on the right hand side exists. Thus

(2.5) $$H(\rho) \sim H(c), \quad c=(c_1, c_2, \cdots, c_k)$$

and the hypothesis $H(c)$ can be tested by the test function

(2.6) $$F=\frac{\sum_{p,q=1}^{k} a_{pq}(\hat{c}_p-c_p)(\hat{c}_q-c_q)}{\sum_{j=1}^{n}(z_j-\hat{c}_0-\sum_{p=1}^{k}\hat{c}_p x'_{jp})^2}\cdot\frac{n-k-1}{k},$$

where

$$a_{pq}=a_{qp}=\sum_{j=1}^{n} x'_{jp}x'_{jq}, \quad (p, q=1, 2, \cdots, k),$$

$$a_{0p}=\sum_{j=1}^{k} x'_{jp}z_j, \quad (p=1, 2, \cdots, k),$$

$$\sum_{q=1}^{k} a_{pq}\hat{c}_q=a_{0p}, \quad (p=1, 2, \cdots, k),$$

$$\hat{c}_0=\bar{z}-\sum_{p=1}^{k}\hat{c}_p\bar{x}_p, \quad (\bar{x}_p=\sum_{j=1}^{n} x'_{jp}/n, \quad \bar{z}=\sum_{j=1}^{n} z_j/n),$$

that is we have the following

Theorem 1. *For a d.p.m.a.* (1.1), *the hypothesis* $H(\rho)$ *is tested by the test function* (2.6) *which follows the F distribution with the pair of degrees of freedom* $(k, n-k-1)$ *under the supposition that the* z_j*'s are random variables and the* x'_{jp}*'s are fixed variates.*

This method may be extended to various cases corresponding to various ways of selection of z_j and x'_{jp}, for instance, we may adopt $z_j\equiv x(u_j+v)$ and $x'_{jp}\equiv x(u_j+p)$ $(p=1, 2, \cdots, k)$, where $v\geqq k+1$ and $u_j-u_{j-1}=v+1$. (Among them, however, the scheme which makes the size of the confidence region for $(\rho(1), \rho(2), \cdots, \rho(k))$ as small as possible may be advisable and in order to get it the relations similar to (2.2)$\sim$(2.4) and (2.6) should be refered to.)

§ 3. Continuous Parameter Case.

Here we consider a real *c.p.m.a.* which has the form

(3.1) $$x(t)=\int_{-\infty}^{\infty} f(s)d\xi(t-s),$$

where $f(s)$ is a real function such that it is continuous except an enumerable set of points, for a fixed t_0,

$$f(s)\equiv 0 \quad \text{for} \quad s>t_0>0 \quad \text{and for} \quad s<0$$

and it has continuous and non-zero point in $(0, \varepsilon)$ and in $(t_0-\varepsilon, t_0)$, ε being an arbitrary positive number such that $\varepsilon<t_0$ and

$$\int_0^{t_0} |f(s)|^2 ds<\infty,$$

and where $\xi(t)$ is a Wiener process on $(-\infty, \infty)$, $E\{|d\xi(t)|^2\} = dt$. Then we get

$$(3.2)\qquad E\{x(t)x(t+\tau)\} = \int_{-\infty}^{\infty} f(u)f(u+\tau)du \qquad |\tau| < t_0,$$
$$= 0 \qquad |\tau| \geqq t_0,$$

that is, for $|t_1 - t_2| \geqq t_0$, $x(t_1)$ and $x(t_2)$ are independent, and the variance of the $x(t)$ is given by

$$(3.3)\qquad \sigma^2 = \int_0^{t_0} |f(u)|^2 du .$$

Let $(l+1)\Delta t = t_0$, l being an arbitrary positive integer, and put

$$(3.4)\qquad b_j y(t + j\Delta t) = \int_0^{\Delta t} f(j\Delta t + s) d\xi(t - j\Delta t - s), \quad j = 0, 1, \cdots, l,$$

where the b_j's are constants such that the $y(t+j\Delta t)$ $(j = 0, 1, \cdots, l)$ have the same variance σ_y^2. We may put $b_0 = 1$ without loss of generality, then

$$(3.5)\qquad \sigma_y^2 = \int_0^{\Delta t} |f(s)|^2 ds$$

and

$$(3.6)\qquad b_j^2 = \int_0^{\Delta t} |f(j\Delta t + s)|^2 ds / \sigma_y^2 .$$

Because of the independency of $y(t+j\Delta t)$ and $y(t+j'\Delta t)$ $(j \neq j')$,

$$(3.7)\qquad x(t' + i\Delta t) = \sum_{j=0}^{l} \int_0^{\Delta t} f(j\Delta t + s) d\xi(t' + i\Delta t - j\Delta t - s)$$
$$= \sum_{j=0}^{l} b_j y(t' + i\Delta t - j\Delta t) \qquad i = 0, \pm 1, \pm 2, \cdots$$

is a *d.p.m.a.* of oder l at most, where t' is an arbitrary constant. The coefficient b_j is a square root of (3.6) and its sign is determined sucessively by the comparison between two types of expressions (2.1) and (3.2) devided by (3.3) for the autocorrelation coefficients in the following way.

$$(3.8)\qquad \begin{aligned} b_l &= \int_{-\infty}^{\infty} f(u)f(u + l\Delta t) du / \sigma_y^2, \\ b_{l-1} + b_1 b_l &= \int_{-\infty}^{\infty} f(u)f(u + (l-1)\Delta t) du / \sigma_y^2, \\ b_{l-2} + b_1 b_{l-1} + b_2 b_l &= \int_{-\infty}^{\infty} f(u)f(u + (l-2)\Delta t) du / \sigma_y^2, \\ &\cdots \\ b_1 + b_1 b_2 + \cdots + b_{l-1} b_l &= \int_{-\infty}^{\infty} f(u)f(u + \Delta t) du / \sigma_y^2, \end{aligned}$$

where

$$\sigma_y^2 = \sigma^2/(1 + b_1^2 + \cdots + b_l^2) = \int_0^{\Delta t} |f(s)|^2 ds .$$

Let us call the process (3.7) the *d.p.m.a.* of formal order *l belongs to* the continuous parameter process (3.1), and if the process (3.7) is regular *m.a.* of order l the continuous parameter process (3.1) may be said to be *regular at the order l.* Since the expression of discrete parameter regular *m.a.* ($b_0 = 1$) is uniquely determined by its autocorrelation coefficients, we have the following

Theorem 2. *To a c.p.m.a.* (3.1) *belongs a d.p.m.a. of arbitrary (formal) order and if it is regular at the order l the d.p,m.a. of order l belongs to it is unique.*

Now, we assume that the functional form of $f(s)$ in (3.1) is known and it has k unknown parameters, $\theta_1, \theta_2, \cdots, \theta_k$, that is $f(s) \equiv f(s\,;\,\theta)$, $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$. In that case we have

Theorem 3. *If the c.p.m.a.* (3.1) *is regular at the order k and if* $H(\theta) \sim H(b)$, $b = (b_1, b_2, \cdots, b_k)$, *the hypothesis* $H(\theta)$ *can be tested by the F-test under the same supposition enunciated in Theorem* 1.

For the purpose of the test for the hypothesis $H(\theta)$, however, it is not necessary to use a *d.p.m.a.*. Let

$$D: \quad 0 < t_1 < t_2 < \cdots < t_k < t_0$$

be a division of the interval $(0, t_0)$ and put

$$z_j \equiv x(2jt_0) = \int_{-\infty}^{\infty} f(s\,;\,\theta) d\xi(2jt_0 - s) ,$$

$$x'_{jp} \equiv x(2jt_0 - t_p) = \int_{-\infty}^{\infty} f(s\,;\,\theta) d\xi(2jt_0 - t_p - s) , \quad p = 1, 2, \cdots, k ,$$

then the regression coefficients of z_j, $c_1, c_2, \cdots, c_k$, on x'_{jp} ($p = 1, 2, \cdots, k$) are the functions of

$$\int_{-\infty}^{\infty} f(s\,;\,\theta) f(t_p + s\,;\,\theta) ds \quad \text{and} \quad \int_{-\infty}^{\infty} f(s\,;\,\theta) f(t_p - t_q + s\,;\,\theta) ds ,$$
$$(p, q = 1, 2, \cdots, k) ,$$

consequently they are the functions of $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$. Thus the test for the hypothesis $H(\theta)$ is reduced to that of $H(c)$, if $H(\theta) \sim H(c)$. It will be an important problem to be solved what division D gives the most efficient test. In this connection, we shall consider a simple case in the following section.

§ 4. A Special Case. Let us consider a simple *c.p.m.a.*

$$x(t)=\int_0^{t_0} f(s;\theta)\,d\xi(t-s), \tag{4.1}$$

$$\begin{aligned} f(s;\theta)&=e^{-\theta s} \quad \text{for} \quad 0\leqq s\leqq t_0, \\ &=0 \quad \text{for} \quad s<0 \ \text{and} \ s>t_0, \quad (\theta>0), \end{aligned} \tag{4.2}$$

$$\sigma^2=(1-e^{-2\theta t_0})/2\theta, \tag{4.3}$$

$$\begin{aligned} \rho(\tau)&=(e^{-\theta|\tau|}-e^{\theta|\tau|-2\theta t_0})/(1-e^{-2\theta t_0}) \qquad |\tau|<t_0, \\ &=0 \qquad |\tau|\geqq t_0. \end{aligned} \tag{4.4}$$

In this one parameter case, if we take $\Delta t=t_0/2$ and put

$$y(t)=\int_0^{\Delta t} e^{-\theta s}\,d\xi(t-s),$$

$$b\,y(t-\Delta t)=\int_0^{\Delta t} e^{-\theta\Delta t-\theta s}\,d\xi(t-\Delta t-s),$$

then

$$\sigma_y^2=(1-e^{-\theta t_0})/2\theta,$$
$$b^2=e^{-\theta t_0}$$

and

$$x(t)=y(t)+b\,y(t-\Delta t), \qquad (t=i\Delta t,\ i=0,\ \pm1,\ \pm2,\ \cdots) \tag{4.5}$$

is a *d.p.m.a.*, where b should be positive and equal to $\exp(-\theta t_0/2)$, because $\rho(\Delta t)=\rho(t_0/2)=b/(1+b^2)>0$ according to (4.4).

The regression coefficient of $x(t)$ on $x(t-\Delta t)$ is given by

$$c=\rho(t_0/2)=e^{-\theta t_0/2}/(1+e^{-\theta t_0}). \tag{4.6}$$

While, if generally $0<\Delta t<t_0$, the regression coefficient of $x(t)$ on $x(t-\Delta t)$ is given by

$$c(\Delta t)=\rho(\Delta t)=e^{-\theta\Delta t}(1-e^{-2\theta(t_0-\Delta t)})/(1-e^{-2\theta t_0}). \tag{4.7}$$

When the sample function $x(t)$ is given on the interval $(0,T)$, the test function we use is

$$F=\frac{(\hat{c}-c)^2\sum_{j=1}^{n}(x_j'-\bar{x}')^2(n-2)}{\sum_{j=1}^{n}\{z_j-\bar{z}-\hat{c}(x_j'-\bar{x}')\}^2}, \tag{4.8}$$

where

$$(n-1)(t_0+\Delta t)+\Delta t\leqq T<n(t_0+\Delta t), \tag{4.9}$$

$$\left.\begin{aligned} z_j&\equiv x((j-1)(t_0+\Delta t)+\Delta t) \\ x_j'&\equiv x((j-1)(t_0+\Delta t)) \end{aligned}\right\} \quad j=1,\ 2,\ \cdots,\ n,$$

$$\bar{z}=\sum_j z_j/n, \qquad x_j'=\sum_j x_j/n$$

and

$$\hat{c}=\sum_j(z_j-\bar{z})(x'_j-\bar{x}')/\sum_j(x'_j-\bar{x}')^2 .$$

From (4.8), the confidence interval for c with confidence coefficient $1-\alpha$ is given by

$$\hat{c}\pm[F^1_{n-2}(\alpha)D^2/(n-2)]^{1/2}, \tag{4.10}$$

where $F^1_{n-2}(\alpha)$ is the $100\,\alpha\,\%$ point of the F distribution with the pair of degrees of freedom $(1, n-2)$, $D^2=\sum_j\{z_j-\bar{z}-\hat{c}(x'_j-x')\}^2/\sum_j(x'_j-\bar{x}')^2$ and for a n large it will be shown by the limit theorem of probability theory under the supposition that x_j's are fixed

$$D^2\sim(1-c^2). \tag{4.11}$$

Hence if we take Δt very small, $1-c^2$ becomes small according to (4.7), and we can take n large in (4.9), consequently we can make the confidence interval length for c small. However, in that case, the confidence interval length for θ itself does not necessarily become small, because, from (4.7), we can easily see that the $c(\Delta t)$ is a monotone decreasing function of θ, $c(\Delta t)\to 1-\Delta t/t_0(\theta\to 0)$ and $c(\Delta t)\to 1\,(\Delta t\to 0+)$ uniformly for $0<\theta<\theta_1$, θ_1 being an arbitrary constant. On the other hand, $c(\Delta t)\to 0(\Delta t\to t_0-0)$ uniformly for $0<\theta<\infty$. An appropriate value of Δt for a given value of t_0 may be determined by refering to the family of functions of θ (4.7) for various values of Δt and the formulas (4.9), (4.10) and (4.11).

METEOROLOGICAL RESEARCH INSTITUTE.

References

[1] H. WOLD; A study in the analysis of stationary time series, Stockholm, Almqvist & Wiksell (1938).

[2] J. L. DOOB; Stochastic processes, New York, J. Wiley (1953).

[3] H. WOLD; *A large sample test for moving averages*, J. Roy. Stat. Soc. (B) **11** (1949), 297–305.

[4] M. OGAWARA; *A general stochastic prediction formula*, Papers in Meteorology and Geophysics, **5** (1955), 193-202.