Tomáš Cipra
Investigation of periodicity for dependent observations

_Aplikace matematiky_, Vol. 29 (1984), No. 2, 134--142

Persistent URL: [http://dml.cz/dmlcz/104076](http://dml.cz/dmlcz/104076)

Terms of use:
© Institute of Mathematics AS CR, 1984

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these _Terms of use._

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project _DML-CZ: The Czech Digital Mathematics Library_ [http://project.dml.cz](http://project.dml.cz)
INVESTIGATION OF PERIODICITY FOR DEPENDENT OBSERVATIONS

Tomáš Cipra

(Received May 10, 1983)

It is proved that Hannan’s procedure [9] for statistical test of periodicity in the case of time series with dependent observations can be combined with Siegel’s improvement of the classical Fisher’s test of periodicity. Simulations performed in the paper show that this combination can increase the power of Hannan’s test when at least two periodicities are present in the time series with dependent observations.

1. INTRODUCTION

The first exact test for periodic components in time series was proposed by Fisher [5], [6]. It deals with observations \( x_1, \ldots, x_n \) arising from the model

\[
x_t = \zeta_t + \varepsilon_t, \quad t = 1, \ldots, n,
\]

where \( \zeta_t \) represents the deterministic unobservable component of the series and \( \varepsilon_t \) is the normal white noise (i.e. \( \varepsilon_t \sim \text{iid } \mathcal{N}(0, \sigma^2) \)) representing random errors due to measurement or other sources. When investigating periodicity of the series we are interested in periodic activity of \( \zeta_t \). The null hypothesis is that there is no periodic activity

\[
H_0 : \zeta_1 = \zeta_2 = \ldots = \zeta_n.
\]

Let \( I_n(\lambda, x) \) be the periodogram of \( x_t \) defined as

\[
I_n(\lambda, x) = a_n^2(\lambda, x) + b_n^2(\lambda, x),
\]

where

\[
a_n(\lambda, x) = \sqrt{\frac{2}{n}} \sum_{t=1}^{n} x_t \cos \lambda t,
\]

\[
b_n(\lambda, x) = \sqrt{\frac{2}{n}} \sum_{t=1}^{n} x_t \sin \lambda t.
\]
Fisher's test uses the values of this periodogram only for the frequencies

\[ \lambda_j = \frac{2\pi j}{n}, \quad j = 1, \ldots, s, \]

where \( s \) is the integer part of \( \frac{1}{2}(n - 1) \). These values must be normalized to the form

\[ Y_j = \frac{I_n(\lambda_j, x)}{\sum_{i=1}^{s} I_n(\lambda_i, x)} \]

(1.6)

to eliminate the effect of \( \sigma^2 \). The hypothesis \( H_0 \) in (1.2) is rejected at the given significance level \( \alpha \) if

\[ \max_{j=1,\ldots,s} Y_j > g_F(n, \alpha), \]

(1.7)

where \( g_F(n, \alpha) \) is the appropriate critical value calculated according to the exact distributional formula for \( \max_j Y_j \) given in [1] or [5] and tabulated in [1], [5] or [14]. The use of the \( r \)th largest value from \( Y_1, \ldots, Y_s, r > 1 \), in testing for periodicity is discussed in [7]. Walker [15] demonstrated that the dropping of the normality condition has little effect on the large sample distributions of the previous test statistics.

When there is a suspicion of activity at several frequencies in (1.1) (i.e., \( \zeta_t \) is composed of several periodic components), Siegel's extension of Fisher's test can be used (see [14]) since it has generally higher power for such compound periodicity. Its test statistic has the form

\[ T_\lambda = \sum_{j=1}^{s} \left( Y_j - \lambda g_F(n, \alpha) \right)_+ \],

(1.8)

where \( (t)_+ \) denotes \( \max(t, 0) \) and \( \lambda \) is a parameter chosen between 0 and 1. The exact distributional formula for \( T_\lambda \) under the null hypothesis (1.2) can be derived and used for the calculation of the critical values \( t_\lambda(n, \alpha) \) tabulated in [14], such that \( H_0 \) is rejected for

\[ T_\lambda > t_\lambda(n, \alpha). \]

(1.9)

Simulation studies demonstrate that the most advantageous value for \( \lambda \) is 0.6 (see [14]). It is shown in [13] that the asymptotic null distribution of \( T_\lambda \) is the so called noncentral chi-squared distribution with zero degrees of freedom.

Bolviken (see [2] or [4]) has tried, similarly to Siegel, to increase the power of Fisher's test for the compound periodicity. He has suggested to replace \( Y_j \) in (1.6) by

\[ I_n(\lambda_j, x)/\sum_{i=1}^{s-a} I_n(\lambda_{(i)}, x), \]

where \( a \) is a preselected constant and the periodogram ordinates \( I_n(\lambda_1), \ldots, I_n(\lambda_s) \) are ordered so that \( I_n(\lambda_1) \leq \cdots \leq I_n(\lambda_{s-a}) \). Experience shows that this modification of Fisher’s test consisting in putting away the largest periodogram values in the denominator of \( Y_j \) can further increase the power of the test when more than one periodic component are present.
Another improvement of the classical Fisher's test is given in [3]. It is based on the fact that the maximum among $I_n(\lambda_1), \ldots, I_n(\lambda_n)$ is frequently much smaller than the maximum of $I_n(\lambda)$ for all $-\pi \leq \lambda \leq \pi$.

The test for periodicity in multiple time series based on the Euclidean norm of the matrix of the periodogram is derived in [10].

All the tests mentioned have the common feature that the null hypothesis $H_0$ supposes the independence of the observations of the series. In practice this demand is often unrealistic and therefore tests of periodicity for dependent observations have been looked for. The most important of them are described in Section 2 of this paper. However, the main purpose of this paper is to propagate the combination of Hannan’s test from Section 2 with Siegel’s test described in this section as a test suitable for dependent observations with compound periodicity in the alternative hypothesis. The theoretical justification of this combination is given in Section 3 while the results of numerical simulations are reported in Section 4.

2. SOME TESTS OF PERIODICITY FOR DEPENDENT OBSERVATIONS

Let observations $x_1, \ldots, x_n$ arise from the model

$$x_t = \zeta_t + u_t, \quad t = 1, \ldots, n,$$

where $u_t = \sum_{j=0}^{\infty} x_j e_{t-j}$ is a normal linear process with a positive spectral density $f(\lambda)$. Since such a process $u_t$ under general assumptions fulfils the relation

$$I_n(\lambda, u) = 2\pi f(\lambda) I_n(\lambda, e) + O(n^{-1/2})$$

(see e.g. [8]), Whittle [16] suggested to replace $I_n(\lambda_j, x)$ in Fisher’s test by

$$K_n(\lambda_j, x) = I_n(\lambda_j, x)/[2\pi f(\lambda_j)], \quad j = 1, \ldots, s$$

(the corresponding distributional formula for the test statistic then holds, of course, only asymptotically). However, we must know apriori the spectral density $f(\lambda)$ for this procedure.

In the case of an unknown spectral density Whittle [17], [18] recommended to use an estimate $\hat{f}(\lambda)$ instead of $f(\lambda)$ in (2.3) but this approach has a great disadvantage: if the null hypothesis of nonexistence of periodicities in a time series is not true the estimate of $f(\lambda)$ in the neighbourhood of the significant frequencies can be inflated remarkably. This inflation of $f(\lambda)$ can reduce the values of (2.3), i.e. it can reduce the power of the test. Therefore Hannan [9] modified Whittle’s approach in such a way that the regression on the harmonic with the frequency $\lambda_j$ is taken out before computing the estimate of the spectral density at that frequency. The regression mentioned can be approximately carried out by using the corresponding value $I_n(\lambda_j)$ of the
periodogram: the previous estimate $\hat{f}(\lambda_j)$ is replaced by

$$f_n^*(\lambda_j) = \frac{\hat{f}_n(\lambda_j) - (2\pi/n) w_n(0) I_n(\lambda_j)}{1 - (8\pi^2/n) w_n(0)},$$

where $w_n(\lambda)$ is the spectral window used for the construction of the estimate $\hat{f}_n$.

A certain version of this Hannan's test is given in [11].

We must also mention the so-called Bartlett's grouped periodogram test described e.g. in [12]. Here the periodogram values $I_n(\lambda_1), \ldots, I_n(\lambda_s)$ are subdivided into several groups so that the spectral density corresponding to the frequencies in any of these groups can be considered approximately constant. Therefore it is possible to use in each of these groups the classical Fisher's test based on the usual periodogram values for the frequencies in the considered group.

All the above mentioned procedures for the case of dependent observations are certain generalizations of the classical Fisher's test so that they may have rather low power in the case of compound periodicity similarly as the classical Fisher's test. So far no procedure considering this fact has been proposed for testing periodicity in dependent observations, although such a test would be desirable for practical purposes (see e.g. [4]). This has motivated this paper in which we try to combine the above mentioned Hannan's test with Siegel's improvement of Fisher's test described in Section 1. The numerical simulations in Section 4 show that this method can actually improve the power of Hannan's test in the case of compound periodicity.

3. THEORETICAL RESULTS

Let us consider the model (2.1), where the process $u_t$ fulfills the following assumptions:

$$u_t = \sum_{j=0}^{\infty} \alpha_j e_{t-j},$$

$$\varepsilon_t \sim \text{iid } \mathcal{N}(0, \sigma^2),$$

$$\sum_{j=0}^{\infty} |\alpha_j| j^{1/2} < \infty,$$

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{\infty} \alpha_j e^{ij\lambda} \right|^2 > 0, \quad -\pi \leq \lambda \leq \pi.$$
The following Lemma shows that in the case of dependent observations the expression $Y_j$ in Siegel's statistic (1.8) can be replaced for large samples by the expression $K_n(\lambda_j, x)\sum_{i=1}^{s} K_n(\lambda_i, x)$.

**Lemma.** Let the assumptions (3.1)–(3.4) be fulfilled. Then under the null hypothesis (1.2),

\[
\lim_{n \to \infty} P \left\{ \sum_{j=1}^{s} \left( \frac{K_n(\lambda_j, x)}{\sum_{i=1}^{s} K_n(\lambda_i, x)} - \lambda g_f(n, x) \right)_+ > t_j(n, x) \right\} = \alpha.
\]

**Proof.** Since $|y_+ - (z)_+| \leq |y - z|$ we can write

\[
\sum_{j=1}^{s} \left( \frac{K_n(\lambda_j, u)}{\sum_{i=1}^{s} K_n(\lambda_i, u)} - \lambda g_f(n, x) \right)_+ - \sum_{j=1}^{s} \left( \frac{I_n(\lambda_j, \varepsilon)}{\sum_{i=1}^{s} I_n(\lambda_i, \varepsilon)} - \lambda g_f(n, x) \right)_+ \leq
\]

\[
\leq \sum_{j=1}^{s} \left| \frac{K_n(\lambda_j, u)}{\sum_{i=1}^{s} K_n(\lambda_i, u)} - \frac{I_n(\lambda_j, \varepsilon)}{\sum_{i=1}^{s} I_n(\lambda_i, \varepsilon)} \right| \leq \frac{1}{\sum_{i=1}^{s} I_n(\lambda_i, \varepsilon)} \left[ \sum_{i=1}^{s} I_n(\lambda_i, \varepsilon) \right]^2
\]

\[
\leq 1 + \sum_{i=1}^{s} \left| \frac{K_n(\lambda_i, u) - I_n(\lambda_i, \varepsilon)}{\sum_{i=1}^{s} I_n(\lambda_i, \varepsilon)} \right| - \frac{1}{\sum_{i=1}^{s} I_n(\lambda_i, \varepsilon)} \left[ \sum_{i=1}^{s} I_n(\lambda_i, \varepsilon) \right]^2
\]

\[
= 2 \left[ \sum_{i=1}^{s} \left| \frac{K_n(\lambda_i, u) - I_n(\lambda_i, \varepsilon)}{\sum_{i=1}^{s} I_n(\lambda_i, \varepsilon)} \right| \right] - \frac{1}{\sum_{i=1}^{s} I_n(\lambda_i, \varepsilon)} \left[ \sum_{i=1}^{s} I_n(\lambda_i, \varepsilon) \right]^2
\]

We have (see e.g. [1])

\[
\sum_{i=1}^{s} I_n(\lambda_i, \varepsilon)/s = \sum_{i=1}^{n} e_i^2/s
\]
so that

\[(3.7) \quad \sum_{i=1}^{s} I_n(\lambda_i, \epsilon) / s \to 2\sigma^2\]

in probability for \(n \to \infty\) according to the weak law of large numbers since \(s\) is the integer part of \((n - 1)/2\).

Further,

\[(3.8) \quad \max_{j=1,\ldots,s} |K_n(\lambda_j, u) - I_n(\lambda_j, \epsilon)| \to 0\]

in probability for \(n \to \infty\) according to [8].

Since \(K_n(\lambda_j, x) = K_n(\lambda_j, u)\) under the null hypothesis \((1.2)\), we have due to \((3.6)-(3.8)\)

\[\sum_{j=1}^{s} \left( \frac{K_n(\lambda_j, x)}{\sum_{i=1}^{s} K_n(\lambda_i, x)} - \lambda g_f(n, \alpha) \right) + \sum_{j=1}^{s} \left( \frac{I_n(\lambda_j, \epsilon)}{\sum_{i=1}^{s} I_n(\lambda_i, \epsilon)} - \lambda g_f(n, \alpha) \right) \to 0\]

in probability for \(n \to \infty\) so that \((3.5)\) holds because this formula holds for \(K_n(\lambda_i, x)\) replaced by \(I_n(\lambda_i, \epsilon), i = 1, \ldots, s\) (the critical values \(t_j(n, \alpha)\) converge to a constant for \(n \to \infty\) since \(T_\alpha\) has under \((1.2)\) the asymptotic distribution described in Section 1). This completes the proof of Lemma.

On the basis of the previous discussion we suggest in the case of dependent observations with a suspicion of compound periodicity to use the test of \(H_0\) with the critical region

\[\sum_{j=1}^{s} \left( \frac{K_n^*(\lambda_j, x)}{\sum_{i=1}^{s} K_n^*(\lambda_i, x)} - \lambda g_f(n, \alpha) \right) > t_j(n, \alpha), \quad \text{(3.9)}\]

where

\[K_n^*(\lambda_j, x) = I_n(\lambda_j, x)/\{2\pi f_n^*(\lambda_j)\}, \quad j = 1, \ldots, s, \quad \text{(3.10)}\]

and \(f_n^*(\lambda_j)\) is defined in \((2.4)\).

### 4. SIMULATION STUDY

To verify the result of our lemma we performed the simulations given in Table 2. Some simple normal MA or AR models without periodicities were chosen and 100 replications were carried out for each of these models at the computer ADT 4130 at the Department of Statistics of Charles University. The constants \(\lambda = 0.6\) and \(\alpha = 0.05\) were chosen. We used critical values taken from [14], which are given in Table 1. Since we have known the spectral density for each of these models we could calculate directly the values \(K_n(\lambda_j, x)\) according to \((2.3)\) without the modification.
The empirical significance levels corresponding to (3.5) are recorded in Table 2. They are really in very good accordance with the used theoretical significance level $\alpha = 0.05$ so that our lemma is empirically justified.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\lambda$</th>
<th>$\alpha$</th>
<th>$g_F(n, \alpha)$</th>
<th>$t_4(n, \alpha)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>41</td>
<td>0.4</td>
<td>0.01</td>
<td>0.330</td>
<td>0.222</td>
</tr>
<tr>
<td>41</td>
<td>0.6</td>
<td>0.05</td>
<td>0.270</td>
<td>0.116</td>
</tr>
<tr>
<td>51</td>
<td>0.4</td>
<td>0.01</td>
<td>0.278</td>
<td>0.194</td>
</tr>
<tr>
<td>51</td>
<td>0.6</td>
<td>0.05</td>
<td>0.228</td>
<td>0.0997</td>
</tr>
<tr>
<td>81</td>
<td>0.6</td>
<td>0.05</td>
<td>0.157</td>
<td>0.0721</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>$n$</th>
<th>Empirical significance level corresponding to (3.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_t = e_t + 0.6e_{t-1}$</td>
<td>51</td>
<td>0.06</td>
</tr>
<tr>
<td>$x_t = e_t + 0.8e_{t-1}$</td>
<td>81</td>
<td>0.05</td>
</tr>
<tr>
<td>$x_t = 0.5x_{t-1} + e_t$</td>
<td>51</td>
<td>0.07</td>
</tr>
<tr>
<td>$x_t = 1.1x_{t-1} - 0.5x_{t-2} + e_t$</td>
<td>51</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Table 2. The frequencies of rejections of $H_0$ for models with known spectral densities and without periodicities ($e_t \sim \text{iid } N(0, 1), \lambda = 0.6, \alpha = 0.05$)

Table 3 concerns the general case with unknown spectral density when the test (3.9) is recommended. Parzen’s estimates of the spectral densities were used with points of truncation ranging from $n/6$ to $n/5$ in accordance with recommendations in the literature (see e.g. [1]). The power of the suggested test (3.9) can be compared with the power of Hannan’s test by means of Table 3. One can see that the suggested tests has systematically higher power than Hannan’s test in the case of compound periodicity. Under the null hypothesis $H_0$, the empirical significance levels of both tests exceed unpleasantly the theoretical level $\alpha$ (see first two rows of Table 3), which is caused by the choice of relatively small $n$ in the simulations (the previous results are only asymptotical and, moreover, the estimates of spectral densities are imperfect for small $n$); nevertheless, also in this case the suggested test gives slightly better results than Hannan’s test.
Table 3. The comparison of the power of the test (3.9) and Hannan’s test

<table>
<thead>
<tr>
<th>Model</th>
<th>$n$</th>
<th>$\lambda$</th>
<th>$z$</th>
<th>Point of truncation in $\hat{F}(\lambda)$</th>
<th>Power of test (3.9)</th>
<th>Power of Hannan’s test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_t = \varepsilon_t + 0.6\varepsilon_{t-1}$</td>
<td>51</td>
<td>0.6</td>
<td>0.05</td>
<td>10</td>
<td>0.18</td>
<td>0.19</td>
</tr>
<tr>
<td>$x_t = 0.5x_{t-1} + \varepsilon_t$</td>
<td>81</td>
<td>0.6</td>
<td>0.05</td>
<td>15</td>
<td>0.16</td>
<td>0.18</td>
</tr>
<tr>
<td>$x_t = \varepsilon_t + 0.6\varepsilon_{t-1} + 0.75\cos(2\pi 5t/n) +$</td>
<td>41</td>
<td>0.6</td>
<td>0.05</td>
<td>8</td>
<td>0.70</td>
<td>0.65</td>
</tr>
<tr>
<td>$x_t = \varepsilon_t + 0.6\varepsilon_{t-1} + 0.75\cos(2\pi 5t/n) +$</td>
<td>51</td>
<td>0.6</td>
<td>0.05</td>
<td>10</td>
<td>0.63</td>
<td>0.57</td>
</tr>
<tr>
<td>$x_t = 0.5x_{t-1} + \varepsilon_t + 0.75\cos(2\pi 5t/n) +$</td>
<td>41</td>
<td>0.4</td>
<td>0.01</td>
<td>8</td>
<td>0.64</td>
<td>0.56</td>
</tr>
<tr>
<td>$x_t = 0.5x_{t-1} + \varepsilon_t + 0.5\cos(2\pi 5t/n) +$</td>
<td>51</td>
<td>0.6</td>
<td>0.05</td>
<td>10</td>
<td>0.42</td>
<td>0.32</td>
</tr>
<tr>
<td>$x_t = 1.1x_{t-1} - 0.5x_{t-2} + \varepsilon_t + 0.5\cos(2\pi 5t/n) +$</td>
<td>51</td>
<td>0.4</td>
<td>0.01</td>
<td>10</td>
<td>0.67</td>
<td>0.59</td>
</tr>
<tr>
<td>$x_t = 1.1x_{t-1} - 0.5x_{t-2} + \varepsilon_t + 0.75\cos(2\pi 5t/n) +$</td>
<td>51</td>
<td>0.6</td>
<td>0.05</td>
<td>10</td>
<td>0.59</td>
<td>0.50</td>
</tr>
<tr>
<td>$x_t = 1.1x_{t-1} - 0.5x_{t-2} + \varepsilon_t + \cos(2\pi 5t/n) +$</td>
<td>51</td>
<td>0.6</td>
<td>0.05</td>
<td>10</td>
<td>0.92</td>
<td>0.82</td>
</tr>
</tbody>
</table>
References


Author's address: RNDr. Tomáš Cipra, CSc., Matematicko-fyzikální fakulta UK, Sokolovská 83, 186 00 Praha 8.

Souhrn

VYŠETŘOVÁNÍ PERIODICITY PRO ZÁVÍSLÁ POZOROVÁNÍ

TOMÁŠ CIPRA


Author's address: RNDr. Tomáš Cipra, CSc., Matematicko-fyzikální fakulta UK, Sokolovská 83, 186 00 Praha 8.

142