Jiří Anděl

Statistical analysis of periodic autoregression

Aplikace matematiky, Vol. 28 (1983), No. 5, 364–385

Persistent URL: http://dml.cz/dmlcz/104048

STATISTICAL ANALYSIS OF PERIODIC AUTOREGRESSION

JIŘÍ ANDĚL

(Received January 20, 1983)

The autoregressive parameters of the classical autoregressive model are constants. In some seasonal time series it is possible to assume that the autoregressive coefficients are also periodic functions of time with the period corresponding to the seasonal component. Such a model is called a periodic autoregression. In the paper methods for estimating parameters and testing hypotheses in the periodic autoregression are proposed and investigated. Two models are considered, one with constant variances of the innovation process and the other with periodically changing variances. The statistical analysis is based on the Bayes approach. The parameters of the model are supposed to be random variables with a vague prior density. Theoretical results are demonstrated on numerical examples.

1. INTRODUCTION

The usual autoregressive process $\{X_t\}$ is given by the relation

$$X_t = b_1 X_{t-1} + \ldots + b_n X_{t-n} + Y_t, \tag{1.1}$$

where $\{Y_t\}$ is a white noise with vanishing mean and a variance $\sigma^2 > 0$. The vector $b = (b_1, \ldots, b_n)'$ and the parameter σ^2 are estimated from a realization $X_1, \ldots, X_N$.

The autoregressive model (1.1) can be also used in the analysis of seasonal time series (see Box and Jenkins [3]). In the typical cases the characteristic equation

$$z^n - b_1 z^{n-1} - \ldots - b_n = 0$$

has some roots with the absolute values equal to one and it is known in advance that some fixed autoregressive parameters are zeros.

In the seasonal time series the length of the longest period is known. For example, an economic time series consisting of monthly data is expected to have a periodic behaviour with the period 12. If we have a model of type (1.1) for such a series, it is possible to assume that the vector b of autoregressive parameters is not constant over a year but reflects the same periodicity. Taking into account that the longest period is p (say), we introduce p autoregressive vectors

$$b_1 = (b_{11}, \ldots, b_{1n})', \ldots, b_p = (b_{p1}, \ldots, b_{pn})' .$$

Consider a model with given variables $X_1, \ldots, X_n$ in which X_t for $t > n$ is generated by

$$X_{n+jp+k} = \sum_{i=1}^{n} b_{ki} X_{n+jp+k-i} + Y_{n+jp+k} , \tag{1.2}$$

where $k = 1, \ldots, p$ and $j = 0, 1, 2, \ldots$. Denote $b = (b_1', \ldots, b_p')'$. It is quite natural to introduce vectors

$$Z_s = (X_{n+ps+1}, X_{n+ps+2}, \ldots, X_{n+ps+p})' ,$$

$s = 1, 2, \ldots$. It can be derived from (1.2) that $\{Z_s\}$ is a p-dimensional autoregressive process. This model allows to decide whether the original process $\{X_t\}$ has an explosive behaviour or not. We shall use this approach in examples in Section 6.

For any process $\{X_t\}$ we can define new processes $\{\xi_{j,t}\}$ by $\xi_{j,t} = X_{pt+j}, j = 1 \ldots \ldots, p$. If the processes $\{\xi_{j,t}\}$ are stationary, $\{X_t\}$ is called a periodically correlated random sequence. This concept was introduced and investigated by Gladyshev [6] and [7]. Jones and Brelsford [10] considered the model for periodic autoregression (1.2). They expanded $b_{k1}, \ldots, b_{kn}$ into a Fourier series, the coefficients of which could be estimated by a regression method. The results were used for extrapolation of the process $\{X_t\}$. Pagano [12] investigated asymptotic properties of estimators of the covariance functions of the processes $\{\xi_{j,t}\}$. He proved that in the case of periodic autoregression the estimators for b_{ki} obtained from modified Yule-Walker equations are asymptotically efficient. Cleveland and Tiao [5] introduced a periodic ARMA model. Tiao and Grupe [13] investigated the errors of misclassification, when the periodic structure of an ARMA process was neglected.

It has been discovered that the method of periodic autoregression can substantially simplify the computation of estimates of autoregressive parameters in the classical multiple autoregressive models. This result is also important for estimating spectral characteristics in multiple stationary time series. Some details of this procedure were described by Newton [11].

In the present paper we apply a Bayes method for estimating parameters b and σ^2 in the periodic autoregression. The results are used for testing some hypotheses about the model. A method of estimation is given also in the case that some elements b_{ki} do not depend on k.

The Bayes approach has become popular in the time series analysis (see Zellner [14], for example). Being simple, this method is frequently used in the statistical research. For example, the intervention analysis was also built on the Bayes principle, see Box and Tiao [4]. In our paper we apply the methods used by Anděl [1], pp. 173–180, for the classical autoregressive processes.

2. PRELIMINARIES

In this part of the paper we collect some auxiliary assertions which will be used in the following sections.

Theorem 2.1. *Let* $K = \begin{Vmatrix} A, & B \\ B', & D \end{Vmatrix}$ *be a symmetric positive definite matrix such that A and D are square blocks. Denote*

$$P = D - B'A^{-1}B\,, \quad Q = A - BD^{-1}B'\,, \quad K^{-1} = \begin{Vmatrix} K^{11}, & K^{12} \\ K^{21}, & K^{22} \end{Vmatrix},$$

where K^{-1} and K are divided into blocks in the same manner. Then P and Q are symmetric and positive definite matrices, and

$$K^{11} = A^{-1} + K^{12}PK^{21}\,, \quad K^{12} = -A^{-1}BP^{-1}\,, \quad K^{21} = K^{12\prime}\,, \quad K^{22} = P^{-1}\,.$$

Other expressions for the blocks of K^{-1} are

$$K^{11} = Q^{-1}\,, \quad K^{12} = -Q^{-1}BD^{-1}\,, \quad K^{21} = K^{12\prime}\,, \quad K^{22} = D^{-1} + K^{21}QK^{12}\,.$$

Proof. Theorem is well known from the matrix theory, see Anděl [2], pp. 65–66. □

Theorem 2.2. *Let $Q_1, \ldots, Q_p$ be $n \times n$ symmetric positive definite matrices. Assume $p \geqq 2$ and introduce matrices $Q = Q_1 + \ldots + Q_p$,*

$$H = \begin{Vmatrix} Q_1, & 0, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & \ldots, & Q_{p-1} \end{Vmatrix} - \begin{Vmatrix} Q_1Q^{-1}Q_1, & \ldots, & Q_1Q^{-1}Q_{p-1} \\ \ldots & \ldots & \ldots \\ Q_{p-1}Q^{-1}Q_1, & \ldots, & Q_{p-1}Q^{-1}Q_{p-1} \end{Vmatrix},$$

$$K = \begin{Vmatrix} Q_1^{-1}, & 0, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & \ldots, & Q_{p-1}^{-1} \end{Vmatrix} - \begin{Vmatrix} Q^{-1}, & \ldots, & Q^{-1} \\ \ldots & \ldots & \ldots \\ Q^{-1}, & \ldots, & Q^{-1} \end{Vmatrix}.$$

Then H and K are symmetric positive definite matrices.

Proof. Let $z_1, \ldots, z_{p-1}$ be vectors with n components. Denote $z = (z'_1, \ldots, z'_{p-1})'$ and put

$$z_p = -Q^{-1} \sum_{k=1}^{p-1} Q_k z_k\,.$$

It can be verified that

$$z'Hz = \sum_{k=1}^{p-1} (z_k + z_p)'\, Q_k(z_k + z_p) + z'_pQ_pz_p$$

holds. Let $z \neq 0$. If $z_p \neq 0$, then we immediately obtain

$$z'Hz \geqq z'_pQ_pz_p > 0\,.$$

If $z_p = 0$, then there exists a vector $z_k (1 \leqq k \leqq p - 1)$ such that $z_k \neq 0$ because of our assumption $z \neq 0$. In this case $z_k + z_p \neq 0$ and we have

$$z'Hz \geqq (z_k + z_p)' Q_k(z_k + z_p) > 0 .$$

Therefore, H is a positive definite matrix.

Denote $D = \mathrm{Diag}(Q_1, \dots, Q_{p-1})$. This matrix is regular and symmetric. Since $K = D^{-1}HD^{-1}$, the matrix K is also positive definite. □

Theorem 2.3. *Let V be an $n \times n$ symmetric positive definite matrix and let a random vector $X = (X_1, \dots, X_n)'$ have the density*

$$q(x) = c(1 + x'Vx)^{-m/2} , \tag{2.1}$$

where c is a constant and $m \geqq n + 1$. Introduce a random vector

$$Z = (Z_1, \dots, Z_s)' = (X_{i_1}, \dots, X_{i_s})' ,$$

where $1 \leqq i_1 < i_2 < \dots < i_s \leqq n$, $1 \leqq s < n$. Let W be the matrix arising from the rows $i_1, \dots, i_s$ and from the columns $i_1, \dots, i_s$ of the matrix V^{-1}. Then the marginal density of the vector Z is

$$q_1(z) = c_1(1 + z'W^{-1}z)^{-(m-n+s)/2} ,$$

where c_1 is a constant.

Proof. First we prove the assertion in the case that $(i_1, \dots, i_s) = (n - s + 1, \dots, n)$ i.e. for $Z = (X_{n-s+1}, \dots, X_n)'$. Denote

$$V^{-1} = R = \left\| \begin{matrix} R_{11}, & R_{12} \\ R_{21}, & R_{22} \end{matrix} \right\| ,$$

where R_{22} is an $s \times s$ block. Put

$$T = R_{11} - R_{12}R_{22}^{-1}R_{21} , \quad S = \left\| \begin{matrix} T^{-1/2}, & -T^{1/2}R_{12}R_{22} \\ 0, & R_{22}^{-1/2} \end{matrix} \right\| .$$

Theorem 2.1 gives $S'S = V$, and thus $S'^{-1}VS^{-1} = I$ (the unit matrix). Consider the transformation $U = SX$. The Jacobian is a constant and we get that the density of U is

$$q_2(u) = c_2(1 + u'S'^{-1}VS^{-1}u)^{-m/2} = c_2(1 + u'u)^{-m/2} .$$

Let $U = (U_1', U_2')'$, where U_2 has s components. From $U = SX$ we get $U_2 = R_{22}^{-1/2}Z$. The marginal distribution of U_2 is

$$q_3(u_2) = \int_{R_{n-s}} c_2(1 + u_1'u_1 + u_2'u_2)^{-m/2} \,\mathrm{d}u_1 =$$

$$= c_2(1 + u_2'u_2)^{-m/2} \int_{R_{n-s}} [1 + u_1'(1 + u_2'u_2)^{-1} u_1]^{-m/2} \,\mathrm{d}u_1 ,$$

where R_{n-s} is the Euclidean $(n-s)$-dimensional space. For calculating the integral we put

$$t = (1 + u_2'u_2)^{-1/2} u_1 .$$

The Jacobian is $(1 + u_2'u_2)^{(n-s)/2}$, and thus

$$q_3(u_2) = c_2(1 + u_2'u_2)^{-(m-n+s)/2} \int_{R_{n-s}} (1 + t't)^{-m/2} \,\mathrm{d}t =$$
$$= c_3(1 + u_2'u_2)^{-(m-n+s)/2} .$$

Since $U_2 = R_{22}^{-1/2} Z$, the density of Z is

$$q_1(z) = c_1(1 + z'R_{22}^{-1}z)^{-(m-n+s)/2} , \tag{2.2}$$

and for the special choice of $i_1, \ldots, i_s$ the assertion is proved.

Now consider the general case. Introduce a matrix J which has in each row an element equal to 1 while all the other elements are zero. The units are subsequently placed in the columns

$$1, 2, \ldots, i_1 - 1, i_1 + 1, \ldots, i_2 - 1, i_2 + 1, \ldots, i_s - 1, i_s + 1, i_1, i_2, \ldots, i_s . \tag{2.3}$$

Put $Y = JX$. The elements of the vector Y are X_i in the order (2.3). The vector $Z = (X_{i_1}, \ldots, X_{i_s})'$ is placed at the end of the vector Y. Since $JJ' = I$, we have $J^{-1} = J'$ and the density of Y is

$$q_4(y) = c_4(1 + y'JVJ'y)^{-m/2} .$$

In this case we have proved that the density $q_1(z)$ is given by the formula (2.2), where R_{22} is the $s \times s$ right-down corner of the matrix $(JVJ')^{-1}$. Since $(JVJ')^{-1} = JV^{-1}J'$, R_{22} is the matrix arising from the rows $i_1, \ldots, i_s$ and from the columns $i_1, \ldots, i_s$ of the matrix V^{-1}. □

Theorem 2.4. *Let a vector $X = (X_1, \ldots, X_n)'$ have the density (2.1). Then the random variable*

$$F = \frac{m-n}{n} X'VX$$

has the Fisher-Snedecor $F_{n,m-n}$ distribution.

Proof. The density of $Y = V^{1/2}X$ is

$$g_1(y) = c_1(1 + y'y)^{-m/2} .$$

Consider the transformation

$$\begin{aligned} y_1 &= r^{1/2} \cos \Theta_1 , \\ y_2 &= r^{1/2} \sin \Theta_1 \cos \Theta_2 , \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

$$y_{n-1} = r^{1/2} \sin \Theta_1 \sin \Theta_2 \dots \sin \Theta_{n-2} \cos \Theta_{n-1} ,$$
$$y_n = r^{1/2} \sin \Theta_1 \sin \Theta_2 \dots \sin \Theta_{n-2} \sin \Theta_{n-1} ,$$

where

$$r \geqq 0 , \quad 0 \leqq \Theta_1, \Theta_2, \dots, \Theta_{n-2} < \pi , \quad 0 \leqq \Theta_{n-1} < 2\pi .$$

Denote $\Theta = (\Theta_1, \dots, \Theta_{n-1})'$. Since the Jacobian of the last transformation is $r^{(n/2)-1} h(\Theta)$, where h is a non-negative measurable function, the simultaneous density of r and Θ is

$$g_2(r, \Theta) = c_2 \, r^{(n/2)-1}(1 + r)^{-m/2} \, h(\Theta) .$$

The marginal density of r is

$$g_3(r) = \int g_2(r, \Theta) \, \mathrm{d}\Theta = c_3 \, r^{(n/2)-1}(1 + r)^{-m/2}$$

for $r > 0$. From $F = (m - n) \, r/n$ we obtain the density of F

$$g_4(f) = c_4 \, f^{(n/2)-1} \left(1 + \frac{n}{m - n} f\right)^{-m/2} , \quad f > 0 ,$$

which is the density of the $F_{n,m-n}$ distribution. □

Let $Y = (Y_1, \dots, Y_n)'$ be a random vector with the density

$$p(y; \nu, P) = \frac{\Gamma[(\nu + n)/2]}{(\pi\nu)^{n/2} \, \Gamma(\nu/2) \, |P|^{1/2}} \left(1 + \nu^{-1} \, y'P^{-1}y\right)^{-(\nu+n)/2} ,$$

where P is an $n \times n$ symmetric positive definite matrix. The function p is called the density of the n-dimensional Student t distribution with ν degrees of freedom (see Johnson and Kotz [9], pp. 132–150). Let a vector $X = (X_1, \dots, X_n)'$ have the density q defined in (2.1). Then $Y = (m - n)^{1/2} X$ has the density $p(y; m - n, V^{-1})$, i.e. the n-dimensional t distribution with $m - n$ degrees of freedom. From this point of view, Theorems 2.3 and 2.4 express some properties of the multivariate t distribution.

3. ANALYSIS OF THE MODEL

Let us consider the model (1.2). We shall assume that the variables $\{Y_t\}$ are independent and $Y_t \sim N(0, \sigma^2)$. Moreover, let X_s and Y_t be independent for $s < t$. We shall analyze a realization $X_1, \dots, X_N$. Usually, N is large in comparison with np.

We shall assume that b and σ are random variables. The conditional density of $X_{n+1}, \dots, X_N$, given $X_1 = x_1, \dots, X_n = x_n$ and given b and σ, is

$$(3.1) \qquad f(x_{n+1}, \dots, x_N \mid x_1, \dots, x_n, b, \sigma) = (2\pi)^{-(N-n)/2} \, \sigma^{-(N-n)} \times$$
$$\times \exp\left\{-\frac{1}{2\sigma^2} \sum_{k=1}^{p} \sum_{j=1}^{\alpha_k} \left[x_{n+k+(j-1)p} - \sum_{i=1}^{n} b_{ki} x_{n+k+(j-1)p-i}\right]^2\right\} ,$$

where

$$\alpha_k = \left[\frac{N - n - k}{p}\right] + 1 . \tag{3.2}$$

If we denote $b_{k0} = -1$ for $k = 1, \ldots, p$, then (3.1) can be written in the form

$$(2\pi)^{-(N-n)/2}\, \sigma^{-(N-n)} \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{p}\sum_{i=0}^{n}\sum_{j=0}^{n} q_{ij}^{(k)} b_{ki} b_{kj}\right\}, \tag{3.3}$$

where

$$q_{ij}^{(k)} = \sum_{h=1}^{\alpha_k} x_{n+k+(h-1)p-i}\, x_{n+k+(h-1)p-j} .$$

Introduce the matrices, vectors and variables

$$\Omega_k = \|q_{ij}^{(k)}\|_{i,j=0}^{n}, \quad Q_k = \|q_{ij}^{(k)}\|_{i,j=1}^{n}, \quad q_k = (q_{01}^{(k)}, \ldots, q_{0n}^{(k)})',$$
$$b_k^* = (b_{k1}^*, \ldots, b_{kn}^*)' = Q_k^{-1} q_k, \quad b^* = (b_1^{*\prime}, \ldots, b_p^{*\prime})',$$
$$v_k = q_{00}^{(k)} - b_k^{*\prime} q_k, \quad v = v_1 + \ldots + v_p .$$

We shall assume throughout the paper that all the matrices Ω_k are positive definite. Then the vectors b_k^* are well defined and it follows from Theorem 2.1 that $v_k > 0$ for all k. The symbols c and c_i will denote constants, which may depend on a realization $x = (x_1, \ldots, x_N)'$ of the random vector $(X_1, \ldots, X_N)'$.

Theorem 3.1. *Let b and σ have the prior density*

$$\pi(b, \sigma) = \sigma^{-1}$$

for $\sigma > 0$ and $b \in R_n$ (and equal to zero otherwise) and let (b, σ) be independent of $(X_1, \ldots, X_n)$. Let $N \geqq np + 1$. Then the posterior density $g(b, \sigma \mid x)$ is

$$g(b, \sigma \mid x) = c\sigma^{-N+n-1} \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{p}\left[(b_k - b_k^*)' Q_k (b_k - b_k^*) + v_k\right]\right\} \tag{3.4}$$

for $\sigma > 0$ and zero otherwise. The modus of the posterior distribution is b^* and σ^*, where

$$\sigma^{*2} = v/(N - n + 1) .$$

The marginal posterior density of the vector b is

$$g_1(b \mid x) = c_1\left[1 + v^{-1}\sum_{k=1}^{p}(b_k - b_k^*)' Q_k (b_k - b_k^*)\right]^{-(N-n)/2}, \tag{3.5}$$

and the marginal posterior density of the parameter σ is

$$g_2(\sigma \mid x) = c_2 \sigma^{-N+n-1+np} \exp\{-v/(2\sigma^2)\}, \quad \sigma > 0 . \tag{3.6}$$

Proof. Formula (3.4) immediately follows from (3.3) by using the Bayes theorem. Since all Q_k are positive definite, it is clear that b^* is the modus, because $g(b^*, \sigma \mid x) \geqq g(b, \sigma \mid x)$ for every b and σ. The maximization of $g(b^*, \sigma \mid x)$ over $\sigma \in (0, \infty)$ gives the value of σ^*. In order to prove (3.5), we put

$$a = \sum_{k=1}^{p} [(b_k - b_k^*)' Q_k(b_k - b_k^*) + v_k] .$$

Then

$$g_1(b \mid x) = \int_0^\infty g(b, \sigma \mid x) \, d\sigma = c \int_0^\infty \sigma^{-N+n-1} \exp\{-a/(2\sigma^2)\} \, d\sigma .$$

After the substitution $u = a^{1/2} \sigma^{-1}$ we obtain formula (3.5). Further, we have

$$g_2(\sigma \mid x) = \int_{R_{np}} g(b, \sigma \mid x) \, db =$$

$$= c\sigma^{-N+n-1} \exp\left\{-\frac{v}{2\sigma^2}\right\} \prod_{k=1}^{p} \int_{R_n} \exp\left\{-\frac{1}{2\sigma^2}(b_k - b_k^*)' Q_k(b_k - b_k^*)\right\} db_k .$$

After the substitution $\beta_k = \sigma^{-1}(b_k - b_k^*)$ we get

$$g_2(\sigma \mid x) = c\sigma^{-N+n-1} \exp\left\{-\frac{v}{2\sigma^2}\right\} \prod_{k=1}^{p} \left(\sigma^n \int_{R_n} \exp\left\{-\frac{1}{2}\beta' Q_k \beta_k\right\} d\beta_k\right) =$$

$$= c_2 \sigma^{-N+n-1+np} \exp\{-v/(2\sigma^2)\} . \qquad \square$$

The modus b^* and σ^* can be used as a point estimator of b and σ. As for the decision about the individual parameters, the following assertions can be useful.

Theorem 3.2. *Let $q^{(k)ii}$ be the (i, i)-th element of the matrix Q_k^{-1}. Then the posterior distribution of*

$$T_{ki} = \left[\frac{N - n - np}{vq^{(k)ii}}\right]^{1/2} (b_{ki} - b_{ki}^*)$$

is the Student t distribution with $N - n - np$ degrees of freedom for $k = 1, \ldots, p$ and $i = 1, \ldots, n$.

Proof. Using Theorem 2.3 we obtain the posterior density of b_{ki} in the form

$$c[1 + (b_{ki} - b_{ki}^*)^2/(vq^{(k)ii})]^{-(N-n-np+1)/2} .$$

From here we immediately get the density of T_{ki}. $\square$

Theorem 3.3. *The posterior distribution of v/σ^2 is the χ^2 distribution with $N - n - np$ degrees of freedom.*

Proof. The assertion follows from (3.6). $\square$

Theorem 3.4. *The random variable*

$$F = \frac{N - n - np}{npv} \sum_{k=1}^{p} (b_k - b_k^*)' \, Q_k(b_k - b_k^*)$$

has the posterior $F_{np,N-n-np}$ *distribution.*

Proof. The assertion follows from (3.5) and Theorem 2.4. □

This last theorem can be used as a test of goodness-of-fit in the case that a hypothesis specifies the value of b (e.g. in a simulation). Alternatively, F can be used for testing independence in a time series against the alternative of periodic autoregression, when we put $b = 0$. In both the cases the large values of F are significant.

For testing the fit of a model we can use either the statistics T_{ki} or F. It is important to know a relation among them. Denote

$$D = \mathrm{Diag}\,\{Q_1, \ldots, Q_p\}\,, \quad \Theta = \left\| \begin{matrix} b_1 - b_1^* \\ \ldots\ldots \\ b_p - b_p^* \end{matrix} \right\| .$$

Let Θ_{ki} be the $[(k-1)\,n + i]$-th element of Θ. For any fixed vector $h \in R_{np}$ we have the inequality

$$(h'\Theta)^2 = [(D^{-1/2}h)'\,(D^{1/2}\Theta)]^2 \leqq h'D^{-1}h.\ \Theta'D\Theta\,.$$

For $h = (0, \ldots, 0, 1, 0, \ldots, 0)'$, where 1 is the $[(k-1)\,n + i]$-th element, we get

$$\Theta_{ki}^2 \leqq q^{(k)ii}\Theta'D\Theta\,.$$

This yields

$$T_{ki}^2 \leqq npF\,, \quad k = 1, \ldots, p\,; \quad i = 1, \ldots, n\,.$$

In fact, this result corresponds to Scheffé's theorem concerning multiple comparison.

It is very important to have a test statistic for the decision whether the model (1.2) is really necessary or whether it is possible to reduce it to the classical homogeneous model (1.1). For this purpose we prove the following assertion.

Theorem 3.5. *Let* $\Delta_k = b_k - b_p - (b_k^* - b_p^*)$ *for* $k = 1, \ldots, p-1$, *and introduce a vector* $\Delta = (\Delta_1', \ldots, \Delta_{p-1}')'$. *Let* Q *and* H *be the matrices defined in Theorem 2.2. Then the posterior distribution of the statistic*

$$F_\Delta = \frac{N - n(p-1)}{n(p-1)\,v}\,\Delta'H\Delta$$

is $F_{n(p-1),N-n(p-1)}$.

Proof. Put $\Delta_p = b_p - b_p^*$. The Jacobian of the transformation $(b_1', \ldots, b_p')' \to$ $\to (\Delta_1', \ldots, \Delta_p')'$ is equal to 1. From the formula (3.5) for $g_1(b \mid x)$ we obtain the density

$$g_3(\Delta_1, \ldots, \Delta_p \mid x) = c_3\{1 + v^{-1}[\Delta_p'Q_p\Delta_p + \sum_{k=1}^{p-1} (\Delta_k + \Delta_p)'\,Q_k(\Delta_k + \Delta_p)]\}^{-(N-n)/2}.$$

Denote

$$h = \sum_{k=1}^{p-1} Q_k \Delta_k \,, \quad G = \sum_{k=1}^{p-1} \Delta_k' Q_k \Delta_k - h' Q^{-1} h \,.$$

Then we can write

$$g_3(\Delta_1, \ldots, \Delta_p \mid x) = c_3[1 + v^{-1} G + v^{-1}(\Delta_p + Q^{-1}h)' \, Q(\Delta_p + Q^{-1}h)]^{-(N-n)/2} \,.$$

The posterior density of the vector Δ is

$$g_4(\Delta \mid x) = \int_{R_n} g_3(\Delta_1, \ldots, \Delta_p \mid x) \, \mathrm{d}\Delta_p =$$

$$= c_3(1 + v^{-1}G)^{-(N-n)/2} \,.$$

$$\cdot \int_{R_n} [1 + (1 + v^{-1}G)^{-1} \, v^{-1}(\Delta_p + Q^{-1}h)' \, Q(\Delta_p + Q^{-1}h)]^{-(N-n)/2} \, \mathrm{d}\Delta_p \,.$$

The Jacobian of the substitution

$$w = (1 + v^{-1}G)^{-1/2} \, v^{-1/2} Q^{1/2}(\Delta_p - Q^{-1}h)$$

is

$$(1 + v^{-1}G)^{n/2} \, v^{n/2} |Q|^{-1/2}$$

and thus

$$g_4(\Delta \mid x) = c_3 v^{-n/2} |Q|^{-1/2} \, (1 + v^{-1}G)^{-N/2} \int_{R_n} (1 + w'w)^{-(N-n)/2} \, \mathrm{d}w \,.$$

After some computations we get $G = \Delta' H \Delta$, and thus

$$g_4(\Delta \mid x) = c_4(1 + v^{-1} \Delta' H \Delta)^{-N/2} \,,$$

where

$$c_4 = c_3 v^{-n/2} |Q|^{-1/2} \int_{R_n} (1 + w'w)^{-(N-n)/2} \, \mathrm{d}w \,.$$

Since the matrix H is positive definite, the assertion follows from Theorem 2.4. $\square$

The above result can be used for testing the hypothesis $H_0 : b_1 = b_2 = \ldots = b_p$. Under H_0, we have

$$\Delta_k = b_p - b_k \,, \quad \Delta = (\Delta_1', \ldots, \Delta_{p-1}')' \,.$$

If

$$F_\Delta = \frac{N - n(p-1)}{n(p-1)\, v} \Delta' H \Delta \tag{3.7}$$

exceeds the critical value $F_{n(p-1), N-n(p-1)}(\alpha)$, we reject H_0 on the level α.

In some cases it can be important to test a hypothesis that only some components of the vectors $b_1, \ldots, b_p$ do not depend on the first subscript. For example, consider a set $i_1, \ldots, i_s$ such that

$$1 \leqq i_1 < i_2 < \ldots < i_s \leqq n \,, \quad 1 \leqq s < n \,,$$

and introduce subvectors

$$B_1 = (b_{1i_1}, \ldots, b_{1i_s})', \quad B_*^1 = (b_{1i_1}^*, \ldots, b_{1i_s}^*)',$$
$$\ldots\ldots\ldots\ldots \quad \ldots\ldots\ldots\ldots$$
$$B_p = (b_{pi_1}, \ldots, b_{pi_s})', \quad B_p^* = (b_{pi_1}^*, \ldots, b_{pi_s}^*)'.$$

We are interested in testing the hypothesis $H_0^* : B_1 = B_2 = \ldots = B_p$.

Theorem 3.6. *Let M_k be the inverse of the matrix arising from the rows $i_1, \ldots, i_s$ and from the columns $i_1, \ldots, i_s$ of the matrix Q_k^{-1}, $k = 1, \ldots, p$. Then the marginal posterior density of the vector $B = (B_1', \ldots, B_p')'$ is*

$$q_1(B \mid x) = c[1 + v^{-1} \sum_{k=1}^{p} (B_k - B_k^*)' M_k(B_k - B_k^*)]^{-[N-n-p(n-s)]/2}.$$

Proof. The posterior density of b given in (3.5) can be written in the form

$$g_1(b \mid x) = c_1[1 + (b - b^*)' K(b - b^*)]^{-(N-n)/2},$$

where $K = v^{-1} \operatorname{Diag}\{Q_1, \ldots, Q_p\}$. Our assertion immediately follows from Theorem 2.3. □

Obviously, the case $s = 1$ will occur most frequently. We put briefly $i_i = i$. The result can be formulated as follows.

Corollary 3.7. *Let $q^{(k)ii}$ be the (i, i)-th element of the matrix Q_k^{-1}. Then the marginal posterior density of the elements $b_{1i}, \ldots, b_{pi}$ is*

$$q_2(b_{1i}, \ldots, b_{pi} \mid x) = c[1 + v^{-1} \sum_{k=1}^{p} (b_{ki} - b_{ki}^*)^2 / q^{(k)ii}]^{-[N-n-p(n-1)]/2}.$$

Proof. Corollary is a special case of Theorem 3.6. □

Theorem 3.8. Put $\delta_k = B_k - B_p - (B_k^* - B_p^*)$ for $k = 1, \ldots, p - 1$, and denote $\delta = (\delta_1', \ldots, \delta_{p-1}')'$. Introduce the matrices $M = M_1 + \ldots M_p$ and

$$L = \begin{Vmatrix} M_1, & 0, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & \ldots, & M_{p-1} \end{Vmatrix} - \begin{Vmatrix} M_1 M^{-1} M_1, & \ldots, & M_1 M^{-1} M_{p-1} \\ \ldots & \ldots & \ldots \\ M_{p-1} M^{-1} M_1, & \ldots, & M_{p-1} M^{-1} M_{p-1} \end{Vmatrix}.$$

Then the posterior distribution of the variable

$$F_\delta = \frac{N - (n + 1)p + 2s}{s(p - 1)v} \delta' L \delta$$

is the $F_{s(p-1), N-(n+1)p+2s}$ distribution.

Proof is analogous to that of Theorem 3.5. □

For $s = 1$ and $i_1 = i$ we have

$$M_k = 1/q^{(k)ii}, \quad M = \sum_{k=1}^{p} (1/q^{(k)ii}).$$

Corollary 3.9. *If $s = 1$ and $i_1 = i$, then the posterior distribution of the variable*

$$F_\delta = \frac{N - (n+1)p + 2}{(p-1)v} \{ \sum_{k=1}^{p-1} (\delta_k^2/q^{(k)ii}) - M^{-1} [\sum_{k=1}^{p-1} (\delta_k/q^{(k)ii})]^2 \}$$

is $F_{p-1, N-(n+1)p+2}$.

Proof. Corollary is a special case of Theorem 3.8. □

It can happen that we finally choose a model in which some parameters are periodically changing and some are constant. This situation was described before Theorem 3.6. If we decide for a model corresponding to a hypothesis H_0, it is convenient to have also formulas for estimating the parameters. For two most important cases, when the constant parameters are either at the beginning or at the end of the vectors b_k, we give explicit solutions.

Theorem 3.10. *Let $b_k = (\beta_k', \omega')'$, where the vector ω has s components. Introduce blocks Q_k^{ij} and q_k^i for $i, j = 1, 2$ by*

$$Q_k = \left\| \begin{matrix} Q_k^{11}, & Q_k^{12} \\ Q_k^{21}, & Q_k^{22} \end{matrix} \right\|, \quad q_k = \left\| \begin{matrix} q_k^1 \\ q_k^2 \end{matrix} \right\|, \quad k = 1, \ldots, p,$$

where Q_k^{22} is an $s \times s$ block and q_k^2 is a vector with s components. Denote

$$M = \left\| \begin{matrix} Q_1^{11}, & 0, & \ldots, & 0, & Q_1^{12} \\ 0, & Q_2^{11}, & \ldots, & 0, & Q_2^{12} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0, & 0, & \ldots, & Q_p^{11}, & Q_p^{12} \\ Q_1^{21}, & Q_2^{21}, & \ldots, & Q_p^{21}, & \sum_{k=1}^{p} Q_k^{22} \end{matrix} \right\|, \quad u = \left\| \begin{matrix} q_1^1 \\ q_2^1 \\ \cdots \\ q_p^1 \\ \sum_{k=1}^{p} q_k^2 \end{matrix} \right\|,$$

$$P = \sum_{k=1}^{p} [Q_k^{22} - Q_k^{21}(Q_k^{11})^{-1} Q_k^{12}], \quad \eta = (\beta_1', \ldots, \beta_p', \omega')'.$$

If the prior density of η and σ is σ^{-1} for $\sigma > 0$ and zero otherwise and if (η, σ) is independent of $(X_1, \ldots, X_n)$, then for $N \geqq p(n - s) + s + 1$ the posterior density of (η, σ) is

$$r_1(\eta, \sigma \mid x) = c\sigma^{-N+n-1} \exp\left\{ -\frac{1}{2\sigma^2} [(\eta - \eta^*)' M(\eta - \eta^*) + w] \right\}, \quad \sigma > 0,$$

where

$$\eta^* = M^{-1}u, \quad w = \sum_{k=1}^{p} q_{00}^{(k)} - \eta^{*\prime} u.$$

The matrix

$$M^{-1} = \begin{Vmatrix} M^{11}, & M^{12} \\ M^{21}, & M^{22} \end{Vmatrix}$$

can be calculated by the formulas

$$M^{22} = P^{-1}, \quad M^{21} = -P^{-1}\| Q_1^{21}(Q_1^{11})^{-1}, \ldots, Q_p^{21}(Q_p^{11})^{-1} \| ,$$

$$M^{12} = M^{21\prime}, \quad M^{11} = \begin{Vmatrix} (Q_1^{11})^{-1}, & 0, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & \ldots, & (Q_p^{11})^{-1} \end{Vmatrix} + M^{12}PM^{21}.$$

The modus of the posterior distribution is η^* and σ^*. where

$$\sigma^{*2} = w/(N - n + 1).$$

Proof. From (3.3) we get the posterior density

$$r_1(\eta, \sigma \mid x) = c\sigma^{-N+n-1} \exp\left\{-\frac{1}{2\sigma^2} \sum_{k=1}^{p} (-1, b_k')\, \Omega_k(-1, b_k')'\right\}.$$

Further we have

$$\sum_{k=1}^{p} (-1, b_k')\, \Omega_k(-1, b_k')' = \sum_{k=1}^{p} (q_{00}^{(k)} - b_k' q_k - q_k' b_k + b_k' Q_k b_k) =$$

$$= \sum_{k=1}^{p} [q_{00}^{(k)} - (\beta_k' q_k^1 + \omega' q_k^2) - (q_k^{1\prime}\beta_k + q_k^{2\prime}\omega) +$$

$$+ (\beta_k' Q_k^{11}\beta_k + \beta_k' Q_k^{12}\omega + \omega' Q_k^{21}\beta_k + \omega' Q_k^{22}\omega)] =$$

$$= \eta' M\eta - \eta' u - u'\eta + \sum_{k=1}^{p} q_{00}^{(k)} = (\eta - \eta^*)'\, M(\eta - \eta^*) + w.$$

The formula for M^{-1} follows from Theorem 2.1. The proof that (η^*, σ^*) is the modus is the same as the proof of Theorem 3.1, since the matrix M is positive definite. □

The marginal posterior density of η is

$$r_2(\eta \mid x) = c_2[1 + w^{-1}(\eta - \eta^*)'\, M(\eta - \eta^*)]^{-(N-n)/2}$$

and the posterior distribution of

$$F' = \frac{N - n - p(n - s) - s}{p(n - s) + s}\, w^{-1}(\eta - \eta^*)'\, M(\eta - \eta^*)$$

is $F_{p(n-s)+s,N-n-p(n-s)-s}$.

A similar assertion holds for the case that the constant parameters are placed at the beginning of the vectors $b_1, \ldots, b_p$.

Theorem 3.11. *Let $b_k = (\omega', \beta_k)'$, where the vector ω has s components. Introduce the blocks Q_k^{ij} in the same way as in Theorem 3.10. Denote*

$$N = \left\| \begin{matrix} \sum_{k=1}^{p} Q_k^{11}, & Q_1^{12}, & Q_2^{12}, & \ldots, & Q_p^{12} \\ Q_1^{21}, & Q_1^{22}, & 0, & \ldots, & 0 \\ Q_2^{21}, & 0, & Q_2^{22}, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ Q_p^{21}, & 0, & 0, & \ldots, & Q_p^{22} \end{matrix} \right\|, \quad y = \left\| \begin{matrix} \sum_{k=1}^{p} q_k^1 \\ q_1^2 \\ q_2^2 \\ \ldots \\ q_p^2 \end{matrix} \right\|,$$

$$S = \sum_{k=1}^{p} \left[Q_k^{11} - Q_k^{12} (Q_k^{22})^{-1} Q_k^{21} \right], \quad v = (\omega', \beta_1', \ldots, \beta_p')'.$$

If the prior density of v and σ is σ^{-1} for $\sigma > 0$ and zero otherwise and if (v, σ) is independent of $(X_1, \ldots, X_n)$, then for $N \geqq p(n - s) + s + 1$ the posterior density of (v, σ) is

$$r_2(v, \sigma \mid x) = c\sigma^{N-n-1} \exp \left\{ -\frac{1}{2\sigma^2} \left[(v - v^*)' N (v - v^*) + z \right] \right\}, \quad \sigma > 0,$$

where

$$v^* = N^{-1} y, \quad z = \sum_{k=1}^{p} q_{00}^{(k)} - v^{*\prime} y.$$

The matrix

$$N^{-1} = \left\| \begin{matrix} N^{11}, & N^{12} \\ N^{21}, & N^{22} \end{matrix} \right\|$$

can be calculated by the formulas

$$N^{11} = S^{-1}, \quad N^{12} = -S^{-1} \left\| Q_1^{12} (Q_1^{22})^{-1}, \ldots, Q_p^{12} (Q_p^{22})^{-1} \right\|,$$

$$N^{21} = N^{12\prime}, \quad N^{22} = \left\| \begin{matrix} (Q_1^{22})^{-1}, & 0, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & \ldots, & (Q_p^{11})^{-1} \end{matrix} \right\| + N^{21} S N^{12}.$$

The modus of the posterior distribution is v^* and σ^*, where

$$\sigma^{*2} = z / (N - n + 1).$$

Proof is analogous to that of the previous theorem. □

The marginal posterior density of v is

$$r_3(v \mid x) = c_3 \left[1 + z^{-1} (v - v^*)' N (v - v^*) \right]^{-(N-n)/2}$$

and the posterior distribution of

$$F'' = \frac{N - n - p(n - s) - s}{p(n - s) + s} z^{-1} (v - v^*)' N (v - v^*)$$

is $F_{p(n-s)+s, N-n-p(n-s)-s}$.

From the previous two theorems we could derive marginal distributions and test statistics similar to those given in Theorems 3.2–3.7.

4. MODEL WITH NON-EQUAL VARIANCES

In the model (1.2) it is sometimes assumed that Var $Y_{n+jp+k} = \sigma_k^2 > 0$ depends on k, $k = 1, \ldots, p$. The statistical analysis of the model is similar to the procedure given in Section 3.

Theorem 4.1. *Denote $\sigma = (\sigma_1, \ldots, \sigma_p)'$. If the prior density of (b, σ) is $\sigma_1^{-1} \ldots \sigma_p^{-1}$ for $\sigma_1 > 0, \ldots, \sigma_p > 0$ (and zero otherwise) independently of $(X_1, \ldots, X_n)$, then the posterior density of (b, σ) for positive $\sigma_1, \ldots, \sigma_p$ is*

$$g(b, \sigma \mid x) = c\sigma_1^{-\alpha_1-1} \ldots, \sigma_1^{-\alpha_p-1} \exp\left\{-\sum_{k=1}^{p} \frac{1}{2\sigma_k^2} [(b_k - b_k^*)' Q_k(b_k - b_k^*) + v_k]\right\},$$

where b_k, b_k^*, Q_k and v_k are given in Theorem 3.1 and α_k is defined in (3.2). The modus of the posterior distribution is b^* and $\sigma_1^*, \ldots, \sigma_p^*$ where

$$\sigma_k^{*2} = v_k/(\alpha_k + 1), \quad k = 1, \ldots, p.$$

The marginal posterior density of b is

$$g_1(b \mid x) = c_1 \prod_{k=1}^{p} [1 + v_k^{-1}(b_k - b_k^*)' Q_k(b_k - b_k^*)]^{-\alpha_k/2} \tag{4.1}$$

and the marginal posterior density of σ is

$$g_2(\sigma \mid x) = c_2 \prod_{k=1}^{p} \sigma_k^{-\alpha_k-1+n} \exp\left\{-\frac{v_k}{2\sigma_k^2}\right\} \tag{4.2}$$

for positive $\sigma_1, \ldots, \sigma_p$.

Proof. Theorem can be proved in the same way as the assertions in Section 3. □

Theorem 4.2. *Denote*

$$F_k = \frac{\alpha_k - n}{n} v_k^{-1}(b_k - b_k^*)' Q_k(b_k - b_k^*), \quad k = 1, \ldots, p.$$

Let H_k be the distribution function of the F_{n,α_k-n} distribution. Put $\pi_k = 1 - H_k(F_k)$. Then the posterior distribution of the random variable

$$\varrho = -2\sum_{k=1}^{p} \ln \pi_k$$

is χ_{2p}^2.

Proof. From (4.1) we can see that, given x, the vectors $b_1, \ldots, b_p$ are independent and the density of b_k has the form

$$c_{(k)}[1 + v_k^{-1}(b_k - b_k^*)' Q_k(b_k - b_k^*)]^{-\alpha_k/2}.$$

Then $F_k \sim F_{n,\alpha_k - n}$ (see Theorem 2.4) and given x, the variables $F_1, \ldots, F_p$ are independent. Further we use the well known construction of a test of significance based on several goodness-of-fit tests (see Janko [8], p. 27). □

The result can be used for testing the hypothesis that the theoretical values of autoregressive vectors are $b_1, \ldots, b_p$. If $\varrho \geqq \chi^2_{2p}(\alpha)$, we reject this hypothesis on the level α.

Theorem 4.3. *Denote* $\hat{\sigma}_k^2 = v_k/(\alpha_k - n)$, $k = 1, \ldots, p$. *Then, given* x, *the variables* $\hat{\sigma}_1^2, \ldots, \hat{\sigma}_p^2$ *are independent and*

$$(\alpha_k - n)\, \hat{\sigma}_k^2/\sigma_k^2 \sim \chi^2_{\alpha_k - n} \,.$$

Proof. The conditional independence immediately follows from (4.2). The marginal posterior density of σ_k is

$$c_k \sigma_k^{-\alpha_k - 1 + n} \exp\{-v_k/(2\sigma_k^2)\} \,.$$

Hence we get that $v_k/\sigma_k^2 \sim \chi^2_{\alpha_k - n}$. □

In many cases $\alpha_k = \alpha_0$ does not depend on k. In such a case we can test the hypothesis $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_p^2$ using the Cochran test (see Janko [8], p. 62). We calculate the statistic

$$g' = \left(\max_{1 \leqq k \leqq p} \hat{\sigma}_k^2\right) \Big/ \sum_{k=1}^{p} \hat{\sigma}_k^2 \,.$$

The critical values of g' for the levels $\alpha = 0{\cdot}05$ and $\alpha = 0{\cdot}01$ are tabulated in Janko [8], Tab. 16a and 16b.

This test shows whether the model with non-equal variances is necessary or whether it is possible to reduce it to the model investigated in Section 3.

Let us remark that for $p = 2$ in the case $\sigma_1^2 = \sigma_2^2$ the ratio $\hat{\sigma}_1^2/\hat{\sigma}_2^2$ has the posterior $F_{\alpha_1 - n, \alpha_2 - n}$ distribution and we can simply use the classical F test for comparing two variances instead of the Cochran test.

In order to simplify the procedure described in Theorem 4.2 and to derive a test corresponding to F_Δ in (3.7) we use some approximations.

Again, consider a random vector X with the density $q(x) = c(1 + x'Vx)^{-m/2}$ which was introduced in Theorem 2.3. Then the random vector $Y = m^{1/2}X$ has the density $c'(1 + y'Vy/m)^{-m/2}$, and we can see that the distribution of Y converges to $N(0, V^{-1})$ as $m \to \infty$. In view of this fact we shall approximate the distribution of X by $N(0, m^{-1}V^{-1})$, i.e. $q(x)$ will be approximated by the density $c'' \exp\{-\frac{1}{2}x'mVx\}$.

Denote

$$U_k = v_k^{-1}\alpha_k Q_k \,, \quad U = U_1 + \ldots + U_p \,,$$

$$L = \begin{Vmatrix} U_1, & 0, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & \ldots, & U_{p-1} \end{Vmatrix} - \begin{Vmatrix} U_1 U^{-1} U_1, & \ldots, & U_1 U^{-1} U_{p-1} \\ \ldots & \ldots & \ldots \\ U_{p-1} U^{-1} U_1, & \ldots, & U_{p-1} U^{-1} U_{p-1} \end{Vmatrix} .$$

The marginal posterior density $g_1(b \mid x)$ in (4.1) can be approximated by the density

$$(4.3) \qquad \lambda(b \mid x) = c \exp\left\{-\frac{1}{2}\sum_{k=1}^{p} (b_k - b_k^*)' U_k(b_k - b_k^*)\right\}.$$

Hence we immediately get, that the random variable

$$(4.4) \qquad \gamma = \sum_{k=1}^{p} (b_k - b_k^*)' U_k(b_k - b_k^*)$$

has approximately the posterior χ^2_{np} distribution. If γ exceeds the critical value $\chi^2_{np}(\alpha)$, we reject the hypothesis that the parameters of the model are $b_1, \ldots, b_p$ on a level which is approximately α. The statistic γ can be used instead of ϱ given in Theorem 4.2.

Now, we derive a test statistic for testing the hypothesis $H_0 : b_1 = b_2 = \ldots = b_p$. The derivation is similar to the proof of Theorem 3.5 and thus it is only sketched. Put

$$\Delta_k = b_k - b_p - (b_k^* - b_p^*) \quad \text{for} \quad k = 1, \ldots, p - 1\,, \quad \Delta_p = b_p - b_p^*\,,$$

$$\Delta = (\Delta'_1, \ldots, \Delta'_{p-1})'\,, \quad h = \sum_{k=1}^{p-1} U_k \Delta_k\,, \quad G = \sum_{k=1}^{p-1} \Delta'_k U_k \Delta_k - h' U^{-1} h\,.$$

Let b have the density (4.3). Then the density of $(\Delta'_1, \ldots, \Delta'_p)'$ is

$$\lambda_1(\Delta'_1, \ldots, \Delta'_p \mid x) = c_1 \exp\{-\tfrac{1}{2}[G + (\Delta_p + U^{-1}h)' U(\Delta_p + U^{-1} h)]\}\,.$$

Since $G = \Delta' L \Delta$, the marginal density of Δ is

$$\lambda_2(\Delta \mid x) = c_2 \exp\{-\tfrac{1}{2}\Delta' L \Delta\}\,.$$

Then

$$(4.5) \qquad r = \Delta L' \Delta$$

has the posterior $\chi^2_{n(p-1)}$ distribution. If $r \geqq \chi^2_{n(p-1)}(\alpha)$, we reject the hypothesis H_0 on a level which is approximately equal to α.

Let $u^{(k)ij}$ be the (i, j)-th element of matrix U_k^{-1}. It follows from (4.3) that the posterior distribution of $(b_{ki} - b_{ki}^*)/[u^{(k)ii}]^{1/2}$ is approximately $N(0, 1)$. This result can be used for testing hypotheses or for constructing confidence intervals for b_{ki}.

5. SUBSET PERIODIC AUTOREGRESSION

Consider again the model (1.2). In some cases it can be known in advance that some parameters b_{ki} are zeros. Assume that only the parameters b_{ki} for $i \in I = \{i_1, \ldots, i_m\}$ can be non-vanishing. Then we have the model

$$X_{n+jp+k} = \sum_{i \in I} b_{ki} X_{n+jp+k-i} + Y_{n+jp+k}\,.$$

All the procedures from Sections 3 and 4 remain valid, only the type of matrices Q_k is $m \times m$, their elements are defined only for $i, j \in I$ and starting from Theorem 3.1 we must write everywhere m instead of n and after that $N - n + m$ instead of N.

6. EXAMPLES

Consider a special case of the periodic autoregression with $n = 2$, $p = 2$, i.e.

$$(6.1) \qquad \begin{aligned} X_{2t+1} &= b_{11}X_{2t} + b_{12}X_{2t-1} + Y_{2t+1}\,, \\ X_{2t+2} &= b_{21}X_{2t+1} + b_{22}X_{2t} + Y_{2t+2}\,. \end{aligned}$$

If we denote

$$\xi_t = \begin{Vmatrix} X_{2t-1} \\ X_{2t} \end{Vmatrix}, \quad \eta_t = \begin{Vmatrix} Y_{2t-1} \\ Y_{2t} \end{Vmatrix}, \quad A_0 = \begin{Vmatrix} 1, & 0 \\ -b_{21}, & 1 \end{Vmatrix}, \quad A_1 = \begin{Vmatrix} -b_{12}, & -b_{11} \\ 0, & -b_{22} \end{Vmatrix},$$

the model can be written in the form

$$A_0\xi_{t+1} + A_1\xi_t = \eta_{t+1}\,.$$

Thus we have the classical two-dimensional autoregressive process. (For general n and p see Pagano [12].) Obviously,

$$|A_0 z + A_1| = z^2 - (b_{11}b_{21} + b_{12} + b_{22})\,z + b_{12}b_{22}\,.$$

Let z_1 and z_2 be the roots of the equation

$$(6.2) \qquad z^2 - (b_{11}b_{21} + b_{12} + b_{22})\,z + b_{12}b_{22} = 0\,.$$

It is well known that $\{\xi_t\}$ is stationary provided $|z_1| < 1$, $|z_2| < 1$. If this condition is not satisfied, $\{\xi_t\}$ has an explosive behaviour. Let us remark, however, that the results of previous sections do not depend on whether the condition $|z_1| < 1$, $|z_2| < 1$ holds or not.

A realization of the process (6.1) with $Y_t \sim N(0, 1)$, $b_{11} = 0{\cdot}2$, $b_{21} = 0{\cdot}6$, $b_{12} = b_{22} = 0{\cdot}7$ for $t = 1, 2, \ldots, 80$ is given in Fig. 1. The roots of equation (6.2) are $z_1 = 1{\cdot}056$ and $z_2 = 0{\cdot}464$. The explosive character of the series is clearly visible from Fig. 1. The results of the analysis are

$$b^*_{11} = 0{\cdot}096\,, \quad b^*_{12} = 0{\cdot}896\,, \quad b^*_{21} = 0{\cdot}561\,, \quad b^*_{22} = 0{\cdot}706\,, \quad \sigma^{*2} = 0{\cdot}963\,.$$

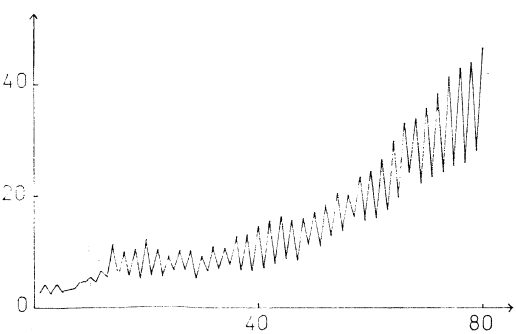

Fig. 1. A realization of the process $X_{2t+1} = 0.2X_{2t} + 0.7X_{2t-1} + Y_{2t+1}$, $X_{2t+2} = 0.6X_{2t+1} + 0.7X_{2t} + Y_{2t+2}$.

None of the statistics T_{ij} and F from Theorems 3.2 and 3.4 are significant at the level 0·05. To test the hypothesis $b_1 = b_2$ we used formula (3.7) with the result $F_\Delta = 53{\cdot}6$, which exceeds the critical value $F_{2,78}(0{\cdot}05) = 3{\cdot}11$. The difference between b_{12}^* and b_{22}^* is not significant at the level 0·05 (the correspondig F statistic was calculated by using Theorem 3.8). If we assume that $b_{12} = b_{22}$, we can use Theorem 3.10 for calculating new estimates. In our case we have got $\eta^* = (0{\cdot}185, 0{\cdot}499, 0{\cdot}746)'$ as an estimate of $\eta = (0{\cdot}2, 0{\cdot}6, 0{\cdot}7)'$.

The resut of simulation of the process (6.1) with $b_{11} = 0{\cdot}5$, $b_{12} = 0{\cdot}6$, $b_{21} = 0{\cdot}2$ and $b_{22} = 0{\cdot}6$ for $t = 1, \ldots, 80$ is given in Fig. 2. The roots of (6.2) are $z_1 = 1{\cdot}019$,

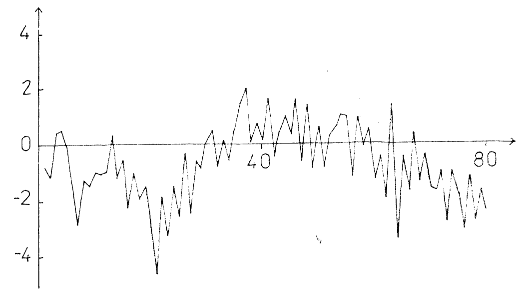

Fig. 2. A realization of the process $X_{2t+1} = 0.5X_{2t} + 0{\cdot}6X_{2t-1} + Y_{2t+1}$, $X_{2t+2} = 0{\cdot}2X_{2t+1} + 0.6X_{2t} + Y_{2t+2}$.

$z_2 = 0{\cdot}481$. Again, the process should have an explosive character, but this time the length of the realization is too short to recognize it visually. The estimates are

$$b_{11}^* = 0{\cdot}380\,, \quad b_{12}^* = 0{\cdot}540\,, \quad b_{21}^* = 0{\cdot}010\,, \quad b_{22}^* = 0{\cdot}789\,, \quad \sigma^{*2} = 1{\cdot}098\,.$$

The statistics T_{ij} and F from Theorems 3.2 and 3.4 are not significant at the level 0·05. However, $F_\Delta = 2{\cdot}34$ is not significant either.

The processes similar to that in Fig. 1 occur in economics. Our simulations show that for a given length of the record even a small change of the parameters of the model can lead from a clearly significant periodic autoregression to the decision for a classical autoregressive model.

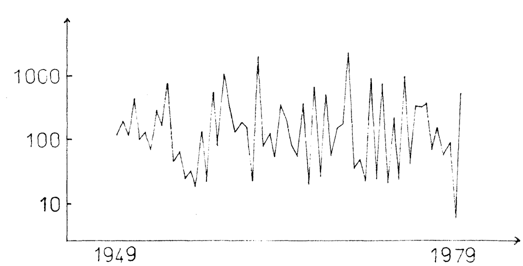

Fig. 3. Numbers of deaths due to influenza in Czech countries.

Two real time series are given in Tab. 1. Series A contains numbers of deaths due to influenza in Czech countries from 1949 to 1979, which were obtained from more detailed data given by the Czech Statistical Institute (Český úřad statistický). The first column of A summarizes data from April to September, the second column from October to March of the next year. Series B is formed by numbers of cases

Tab. 1.

A — numbers of deaths due to influenza in Czech countries
B — numbers of cases of hepatitis in Czech countries

	A		B	
Year	Apr.-Sept.	Oct.-March	Jan.-June	July-Dec.
1949	116	208		
1950	125	450	3 901	15 435
1951	96	136	14 983	18 757
1952	65	285	16 334	13 852
1953	159	709	9 678	19 352
1954	44	69	16 218	21 326
1955	25	32	14 815	19 491
1956	18	131	15 347	13 699
1957	23	522	8 713	8 326
1958	79	1 119	6 574	9 973
1959	288	139	7 313	12 749
1960	189	152	10 281	15 422
1961	22	2 259	11 688	15 641
1962	78	116	10 096	11 134
1963	56	372	9 078	16 664
1964	221	80	12 058	15 058
1965	54	364	9 531	8 020
1966	22	641	5 457	6 417
1967	26	471	4 432	4 904
1968	56	150	4 023	4 436
1969	183	2 146	3 803	5 012
1970	35	48	4 402	5 840
1971	22	906	4 461	5 074
1972	25	706	4 059	5 543
1973	23	217	5 879	5 325
1974	24	1 040	4 807	5 436
1975	42	316	4 410	4 662
1976	334	374	4 432	4 582
1977	71	159	4 370	4 348
1978	60	94	3 510	3 456
1979	6	570	4 432	30 422
1980			6 288	4 849
1981			4 027	3 326

of hepatitis in Czech countries from 1950 to 1981. The data were obtained from a more detailed information given by the Institute of Health Information and Statistics (Ústav zdravotnických informací a statistiky).

Series A is plotted on the logarithmic scale in Fig. 3, series B on the usual scale in Fig. 4.

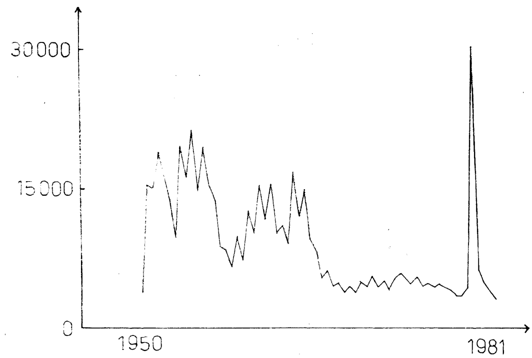

Fig. 4. Numbers of cases of hepatitis in Czech countries.

We analyzed not only series A and B, but also their logarithms. The results of the analysis are summarized in Tab. 2. In all four cases the hypothesis $\sigma_1^2 = \sigma_2^2$ is rejected at the level 0·05 and thus the model with non-equal variances must be applied.

Tab. 2. Results of statistical analysis

	A	$\ln A$	B	$\ln B$
N	62	62	64	64
b_{11}^*	0·0449	0·3597	0·3757	0·6385
b_{12}^*	0·4249	0·4726	0·4561	0·3424
b_{21}^*	2·5149	0·5244	1·5233	1·5827
b_{22}^*	0·1024	0·6036	−0·2022	−0·5433
$\hat{\sigma}_1^2$	9 103	0·8239	6 236 150	0·0578
$\hat{\sigma}_2^2$	454 795	2·5600	28 336 173	0·1379
r	4·86	9·35	15·3	29·7

The critical value for r at the level 0·05 is 5·99. Except for the series A, in all remaining three cases the hypothesis $b_1 = b_2$ is rejected. It seems that the non-significant value of r in the series A is caused (at least partially) by the large value of the ratio $\hat{\sigma}_2^2/\hat{\sigma}_1^2$, because the estimates b_1^* and b_2^* differ considerably.

Acknowledgement. The author is very grateful to RNDr. *Eva Švandová* from the Institute of Hygiene and Microbiology, Prague, for providing the data in Tab.1.

References

[1] *J. Anděl:* The Statistical Analysis of Time Series. SNTL, Prague 1976 (in Czech).
[2] *J. Anděl:* Mathematical Statistics. SNTL, Prague, 1978 (in Czech).
[3] *G. E. P. Box, G. M. Jenkins:* Time Series Analysis, Forecasting and Control. Holden Day, San Francisco, 1970.
[4] *G. E. P. Box, G. C. Tiao:* Intervention analysis with applications to economic and environmental problems. J. Amer. Statist. Assoc. 70 (1975), 70—79.
[5] *W. P. Cleveland, G. C. Tiao:* Modeling seasonal time series. Rev. Economic Appliquée 32 (1979), 107—129.
[6] *E. G. Gladyshev:* Periodically correlated random sequences. Soviet Math. 2 (1961), 385—388.
[7] *E. G. Gladyshev:* Periodically and almost periodically correlated random processes with continuous time parameter. Theory Prob. Appl. 8 (1983), 173—177.
[8] *J. Janko:* Statistical Tables. NČSAV, Prague, 1958 (in Czech).
[9] *N. L. Johnson, S. Kotz:* Distributions in Statistics: Continuous Multivariate Distributions. Wiley, New York, 1972.
[10] *R. H. Jones, W. M. Brelsford:* Time series with periodic structure. Biometrika 54 (1967), 403—408.
[11] *H. J. Newton:* Using periodic autoregression for multiple spectral estimation. Technometrics 24 (1982), 109—116.
[12] *M. Pagano:* On periodic and multiple autoregression. Ann. Statist. 6 (1978), 1310—1317.
[13] *C. G. Tiao, M. R. Grupe:* Hidden periodic autoregressive-moving average models in time series data. Biometrika 67 (1980), 365—373.
[14] *A. Zellner:* An Introduction to Bayesian Inference in Econometrics. Wiley, New York, 1971.

Souhrn

STATISTICKÁ ANALÝZA PERIODICKÉ AUTOREGRESE

Jiří Anděl

V klasickém autoregresním modelu jsou autoregresní parametry konstantní. Při vyšetřování sezónních časových řad lze však v některých případech předpokládat, že autoregresní koeficienty jsou periodickými funkcemi času s odpovídající periodou. Pak jde o tzv. periodickou autoregresi. V práci jsou navrženy metody pro odhad parametrů a testování hypotéz v modelu periodické autoregrese. Je vyšetřován jak model s konstantními rozptyly inovačního procesu, tak i model s periodicky se měnícími rozptyly. Statistická analýza je založena na bayesovském přístupu. O parametrech modelu se předpokládá, že to jsou náhodné veličiny s nevlastní apriorní hustotou. Teoretické výsledky jsou demonstrovány na simulovaných i reálných datech.

Author's address: Doc. RNDr. *Jiří Anděl*, DrSc., Matematicko-fyzikální fakulta Univerzity Karlovy, Sokolovská 83, 186 00 Praha 8.