ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 152, 2000

Bronisław Ceranka, Krystyna Katulska***

ANALYSIS OF VARIANCE WITH TIME SERIES DATA

Abstract. Successive observations over time are made on individual subjects classified into different groups. It is assumed that the mean response may vary between groups, that there is a random effect for each individual, and that successive observations on each individual follow a AR model. The likelihood-ratio criteria for testing the hypothesis of equality of the group means is considered.

1. INTRODUCTION

The usual analysis of variance F test is considered in the case when the variance-covariance matrix of errors is of the from $\sigma^2\mathbf{I}_n$. In many practical situations this assumption is not satisfied, for example, if observations are correlated, but we are interested of testing the hypothesis of equality of the group means. In this paper, we assume, that successive observations over time are made on individual subjects classified into different groups. This problem was considered also by Yong and Carter (1983). They proposed to use the usual analysis of variance F test for nested designs.

The purpose of this paper is considered the likelihood ratio for testing hypothesis of equality of the group means for correlated observations.

2. LIKELIHOOD RATIO TEST

Consider the experiment in which $m \times n$ experimental units are divided in m groups each of n units. Let Y_{ijt} be the observation at time t on the j-th individual from group i. Suppose that Y_{ijt} can be modelled by

* Agricultural University of Poznań, Department of Mathematical and statistical Methods.
** Adam Mickiewicz University, Faculty of Mathematics and Computer Science.

$$Y_{ijt} = \mu_i + \alpha_{ij} + \varepsilon_{ijt} \quad i = 1, \ldots, m, \; j = 1, \ldots, n, \; t = 1, 2, \ldots, T, \tag{1}$$

where μ_i are fixed group means, α_{ij} represent individual effects, and ε_{ijt} are the errors. Further, we assume that ε_{ijt} and α_{ij} are independent, α_{ij} are independent normal random variables with mean zero and variance σ_α^2 and ε_{ijt} are identical stationary Gaussian time series for fixed i and j.

We introduce the following notations. Let $\mathbf{Y}_{ij} = (Y_{ij1}, \ldots, Y_{ijT})'$ be T-dimensional vector of the observations on the j-th individual from group i and $\varepsilon_{ij} = (\varepsilon_{ij1}, \ldots, \varepsilon_{ijT})'$ denote the random vector, such that ε_{ij} is distributed according to $N_T(\mathbf{O}, \sigma_\varepsilon^2 \mathbf{A}_T)$, where $\sigma_\varepsilon^2 \mathbf{A}_T$ is the autocovariance matrix of a stationary Gaussian time series, where $N_p(\mu, \sigma^2 \mathbf{V})$ is a p-variate normal distribution with mean vector μ and the covariance matrix $\sigma^2 \mathbf{V}$.

Then, the model (1) can be written as

$$\mathbf{Y}_{ij} = \mu_i \mathbf{1}_T + \alpha_{ij} \mathbf{1}_T + \varepsilon_{ij}, \; i = 1, \ldots, m, \; j = 1, \ldots, n \tag{2}$$

where $\mathbf{1}_p$ denotes $p \times 1$ vector of ones. It implies that $\mathbf{Y}_{ij}$ is distributed according to $N_T(\mu_i \mathbf{1}_T, \sigma_\varepsilon^2(\gamma_\alpha^2 \mathbf{J}_T + \mathbf{A}_T))$, where $\mathbf{J}_T = \mathbf{1}_T \mathbf{1}_T'$ and $\gamma_\alpha^2 = \sigma_\alpha^2 / \sigma_\varepsilon^2$.

Furthermore, let $\mathbf{Y} = (\mathbf{Y}_{11}', \ldots, \mathbf{Y}_{mn}')'$ $\mu = (\mu_1', \ldots, \mu_m')'$ $\varepsilon = (\varepsilon_{11}', \ldots, \varepsilon_{mn}')'$ and $\alpha = (\alpha_{11}, \ldots, \alpha_{mn})'$. Then (2) is of the form

$$\mathbf{Y} = \mu \otimes \mathbf{1}_{nT} + \alpha \otimes \mathbf{1}_T + \varepsilon \tag{3}$$

where $\otimes$ denotes Kronecker product. Under these notations $\mathbf{Y}$ is distributed according to N_{mnT} $(\mu \otimes \mathbf{1}_{nT}, \; \sigma_\varepsilon^2 \mathbf{I}_{mn} \otimes (\gamma_\alpha^2 \mathbf{J}_T + \mathbf{A}_T))$, where $\mathbf{I}_p$ is the $p \times p$ identity matrix.

In this paper we assume that $\mathbf{A}_T$ is the known matrix, while σ_ε^2 and γ_α^2 are unknown parameters.

The purpose of the data analysis is to test the hypothesis H_0: $\mu_1 = \mu_2 = \ldots = \mu_m$ against the general alternative H_1: H_0 is not true.

Lemma 1. If $\mathbf{Y}$ is given by (3) then the maximum likelihood estimates of μ_i, γ_α^2, σ_ε^2 can be written as

(a)
$$\hat{\mu}_i = \frac{1}{nb} \sum_{i=1}^{n} \mathbf{Y}'_{ij} \mathbf{A}_T^{-1} \mathbf{1}_T, \; i = 1, \ldots, m,$$

$$\hat{\gamma}_\alpha^2 = -\frac{1}{b} + \frac{(T-1)\left\{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} c_{ij} - \frac{1}{n}\left(\sum_{j=1}^{n} \mathbf{Y}_{ij}' \mathbf{A}_T^{-1} \mathbf{1}_T\right)^2\right)\right\}}{b\left(b \sum_{i=1}^{m}\sum_{j=1}^{n} d_{ij} - \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\right)},$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{mn(T-1)} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} d_{ij} - \frac{1}{b} \sum_{j=1}^{n} c_{ij} \right),$$

where

$$b = 1_T' A_T^{-1} 1_T, \quad c_{ij} = \mathbf{Y}_{ij}' \mathbf{A}_T^{-1} \mathbf{J}_T \mathbf{A}_T^{-1} \mathbf{Y}_{ij}, \quad d_{ij} = \mathbf{Y}_{ij}' \mathbf{A}_T^{-1} \mathbf{Y}_{ij}.$$

(b) If H_0 is true, then

$$\tilde{\mu} = \frac{1}{mnb} \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{Y}_{ij}' \mathbf{A}_T^{-1} \mathbf{1}_T,$$

$$\hat{\gamma}_\alpha = -\frac{1}{b} + \frac{(T-1)\left\{ \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} - \frac{1}{mn} (\sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{Y}_{ij}' \mathbf{A}_T^{-1} \mathbf{1}_T)^2 \right\}}{b\left(b \sum_{i=1}^{m} \sum_{j=1}^{n} d_{ij} - \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} \right)},$$

$$\tilde{\sigma}_\varepsilon^2 = \hat{\sigma}_\varepsilon^2.$$

Proof. (a) The likelihood function of the vector $\mathbf{Y}$ given in (3) can be written as

$$p(\mathbf{Y}_{ij}) = (2\pi)^{-\frac{mnT}{2}} \{(1+\gamma_\alpha^2 b)\sigma_\varepsilon^{2T}\}^{-\frac{mn}{2}} (\det \mathbf{A}_T)^{-\frac{mn}{2}} \times$$

$$\times \exp\left(-\frac{1}{2} \sigma_\varepsilon^{-2} \sum_{i=1}^{m} \sum_{j=1}^{n} (\mathbf{Y}_{ij} - \mu_i \mathbf{1}_T)' (\mathbf{A}_T^{-1} - \frac{\gamma_\alpha^2}{1+\gamma_\alpha^2 b} \mathbf{A}_T^{-1} \mathbf{J}_T \mathbf{A}_T^{-1})(\mathbf{Y}_{ij} - \mu_i \mathbf{1}_T) \right).$$

From the definition of the maximum likelihood estimates of μ_i, γ_α^2, σ_ε^2 and the properties of the function $p(\mathbf{Y}_{ij})$ it follows that $\hat{\mu}_i$, $\hat{\gamma}_\alpha^2$, $\hat{\sigma}_\varepsilon^2$ are the solution of the system

$$\frac{\partial p(\mathbf{Y}_{ij})}{\partial \mu_i} = 0, \quad i = 1, \ldots, m,$$

$$\frac{\partial p(\mathbf{Y}_{ij})}{\partial \gamma_\alpha^2} = 0,$$

$$\frac{\partial p(\mathbf{Y}_{ij})}{\partial \sigma_\varepsilon^2} = 0.$$

(b) If H_0 is true then the likelihood function is

$$p^*(\mathbf{Y}_{ij}) = (2\pi)^{-\frac{mnT}{2}}\left\{(1+\gamma_\alpha^2 b)\sigma_\varepsilon^{2T}\right\}^{-\frac{mn}{2}}(\det \mathbf{A}_T)^{-\frac{mn}{2}} \times$$

$$\times \exp\left(-\frac{1}{2}\sigma_\varepsilon^{-2}\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{Y}_{ij}-\mu\mathbf{1}_T)'\left(\mathbf{A}_T^{-1}-\frac{\gamma_\alpha^2}{1+\gamma_\alpha^2 b}\mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}\right)(\mathbf{Y}_{ij}-\mu\mathbf{1}_T)\right).$$

The further part of the proof is analogous to the proof of (a).

The purpose of the data analysis is to test the hypothesis H_0: $\mu_1 = \mu_2 = \ldots = \mu_m$ against the general alternative H_1: H_0 is not true. The likelihood – ratio principle (S c h e f f e, 1959, p. 33) may be used to derive the statistical test.

The likelihood-ratio test consists in rejecting H_0 if $\lambda < \lambda_0$, where the constant λ_0 is chosen to give the desired significance level α i.e. $R = \{\mathbf{Y}: \lambda < \lambda_0\}$, where $P(\lambda < \lambda_0) = \alpha$ and

$$\lambda = \frac{p^*(\tilde{\mu}, \tilde{\gamma}_\alpha^2, \tilde{\sigma}_\varepsilon^2)}{p(\hat{\mu}_i, \hat{\gamma}_\alpha^2, \hat{\sigma}_\varepsilon^2)} = \left(\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{Y}_{ij}-\hat{\mu}_i\mathbf{1}_T)'\mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}(\mathbf{Y}_{ij}-\hat{\mu}_i\mathbf{1}_T)}{\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{Y}_{ij}-\tilde{\mu}\mathbf{1}_T)'\mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}(\mathbf{Y}_{ij}-\tilde{\mu}\mathbf{1}_T)}\right)^{\frac{mn}{2}}.$$

Since

$$P\left(\left(\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{Y}_{ij}-\hat{\mu}_i\mathbf{1}_T)'\mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}(\mathbf{Y}_{ij}-\hat{\mu}_i\mathbf{1}_T)}{\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{Y}_{ij}-\tilde{\mu}\mathbf{1}_T)'\mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}(\mathbf{Y}_{ij}-\tilde{\mu}\mathbf{1}_T)}\right)^{\frac{mn}{2}} < \lambda_0\right) =$$

$$= P\left(\frac{nb^2\sum_{i=1}^{m}\hat{\mu}_i^2 - mnb^2\tilde{\mu}^2}{\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{Y}_{ij}'\mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}\mathbf{Y}_{ij} - nb^2\sum_{i=1}^{m}\hat{\mu}_i^2} > \lambda_0^{-\frac{2}{mn}} - 1\right)$$

then $\alpha = P(\lambda < \lambda_0) = P(\lambda^* > \lambda_0^*)$ and

$$\lambda^* = \frac{nb^2\sum_{i=1}^{m}\hat{\mu}_i^2 - mnb^2\tilde{\mu}^2}{\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{Y}_{ij}'\mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}\mathbf{Y}_{ij} - nb^2\sum_{i=1}^{m}\hat{\mu}_i^2} \tag{4}$$

Hence, the optimal test is to reject H_0 at the α level of significance if $\lambda^* > \lambda_0^*$.

Theorem 1. If $\mathbf{Y}$ is distributed according to $N_{mnT}(\mu \otimes \mathbf{1}_{nT}, \sigma_\varepsilon^2 \mathbf{I}_{mn} \otimes (\gamma_\alpha^2 \mathbf{J}_T + \mathbf{A}_T))$ then λ^* given in (4) has a noncentral $\mathbf{F}$ distribution with m–1, $m(n$–1) degrees of freedom and noncentrality parameter

$$\frac{nb}{\sigma_\varepsilon^2(1+b\gamma_\alpha^2)} \sum_{i=1}^{m} (\mu_i - \bar{\mu})^2 \tag{5}$$

where $\bar{\mu} = \frac{1}{m} \sum_{i=1}^{m} \mu_i$.

Proof. The statistic λ^* can be written as the ratio of two quadratic forms

$$\lambda^* = \frac{c\mathbf{Y}'(m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} - \mathbf{J}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1})\mathbf{Y}}{c\mathbf{Y}'(mn \otimes \mathbf{I}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} - m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1})\mathbf{Y}},$$

where $c = (mn\sigma_\varepsilon^2(1+b\gamma_\alpha^2)b)^{-1}$. From Theorem 1 (R a o, M i t r a, 1971, p. 171) we conclude that if $\mathbf{Y}$ is distributed according to $N_{mnT}(\mu \otimes \mathbf{1}_{nT}, \sigma_\varepsilon^2 \mathbf{I}_{mn} \otimes (\gamma_\alpha^2 \mathbf{J}_T + \mathbf{A}_T))$ then the statistic $\mathbf{Y}'\mathbf{B}\mathbf{Y}$ has the $\chi^2(k, \delta)$ distribution if and only if $\mathbf{B}(\sigma_\varepsilon^2 \mathbf{I}_{mn} \otimes (\gamma_\alpha^2 \mathbf{J}_T - \mathbf{A}_T)$ is idempotent, in which case $k = \mathrm{tr}(\mathbf{B}(\sigma_\varepsilon^2 \mathbf{I}_{mn} \otimes (\gamma_\alpha^2 \mathbf{J}_T + \mathbf{A}_T))$ and $\delta = (\mu \otimes \mathbf{I}_{nT})'\ \mathbf{B}(\mu \otimes \mathbf{1}_{nT})$.
Consider the quadratic form

$$c\mathbf{Y}'(m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} - \mathbf{J}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1})\mathbf{Y} = \mathbf{Y}'\mathbf{B}\mathbf{Y} \tag{6}$$

Since $\mathbf{B}(\sigma_\varepsilon^2 \mathbf{I}_{mn} \otimes (\gamma_\alpha^2 \mathbf{J}_T + \mathbf{A}_T)) = (mnb)^{-1}\ \ (m\mathbf{I}_m \otimes \mathbf{J}_n - \mathbf{J}_{mn}) \otimes \mathbf{A}_T^{-1}\mathbf{J}_T$ it is easy to verify that this matrix is idempotent. Hence the quadratic form (6) has the $\chi^2(k\delta)$ distribution, in which case $k = (mnb)^{-1}\mathrm{tr}\{m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T - \mathbf{J}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\} = m - 1$,

$$\delta = (\mu \otimes \mathbf{1}_{nT})'(c(m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} - \mathbf{J}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}))(\mu \otimes \mathbf{1}_{nT}) =$$
$$= \frac{nb}{\sigma_\varepsilon^2(1+b\gamma_\alpha^2)} \left(\sum_{i=1}^{m} \mu_i^2 - \frac{1}{m} \Big(\sum_{i=1}^{m} \mu_i \Big)^2 \right) = \frac{nb}{\sigma_\varepsilon^2(1+b\gamma_\alpha^2)} = \sum_{i=1}^{m} (\mu_i - \bar{\mu})^2.$$

Further, we investigate the distribution of the quadratic form

$$c\mathbf{Y}'(mn\mathbf{I}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} - m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1})\mathbf{Y} = \mathbf{Y}'\mathbf{B}\mathbf{Y} \tag{7}$$

Since $\mathbf{B}(\sigma_\varepsilon^2 \mathbf{I}_{mn} \otimes (\gamma_\alpha^2 \mathbf{J}_T + \mathbf{A}_T)) = (nb)^{-1}\ \ (n\mathbf{I}_{mn} - \mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T$, it can be easily seen that this matrix is idempotent. Therefore, the quadratic form (7) has the $\chi^2(k, \delta)$ distribution, in which case

$$k = (nb)^{-1}\mathrm{tr}(n\mathbf{I}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T - \mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T = m(n-1),$$

$$\delta = (\mu \otimes \mathbf{1}_{nT})'(c(mn\mathbf{I}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} - m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1})(\mu \otimes \mathbf{1}_{nT}) = 0.$$

It implies, that the quadratic form (7) has the central χ^2 distribution on $m(n-1)$ degrees of freedom.
Moreover, $c(m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} - \mathbf{J}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1})(\sigma_e^2\mathbf{I}_{mn} \otimes (\gamma_\alpha^2\mathbf{J}_T + \mathbf{A}_T)) \times c(mn\mathbf{I}_{mn} \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} - m\mathbf{I}_m \otimes \mathbf{J}_n \otimes \mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1}) = 0$.
It implies (Theorem 9.4.1, Rao, Mitra, 1971, p. 178) that the quadratic forms are independently distributed. The proof is complete.

Corollary 1. If H_0 is true then the statistic λ^* has the central F distribution with $m-1$, $m(n-1)$ degrees of freedom.

From Theorem 1 it follows that the optimal test is to reject H_0 at the α level of significance if

$$\lambda^* > F_{\alpha; m-1,\ m(n-1)}.$$

3. STATISTIC λ^* IN THE CASE AR(p)

In applications of the statistic λ^* for testing H_0: $\mu_1 = \ldots = \mu_m$ against alternative H_1: H_0 is not

$$\mathbf{A}_T = \begin{bmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{T-1} \\ \rho & 1 & \rho & \ldots & \rho^{T-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \ldots & 1 \end{bmatrix},$$

(Goldberger, 1972, p. 203).
Hence (Goldberger, 1972, p. 305)

$$\mathbf{A}_T^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho & 0 & \ldots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -\rho & 1 \end{bmatrix}$$

and

$$\mathbf{A}_T^{-1}\mathbf{J}_T\mathbf{A}_T^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & 1-\rho & \ldots & 1-\rho & 1 \\ 1-\rho & (1-\rho)^2 & \ldots & (1-\rho)^2 & 1-\rho \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1-\rho & (1-\rho)^2 & \ldots & (1-\rho)^2 & 1-\rho \\ 1 & 1-\rho & \ldots & 1-\rho & 1 \end{bmatrix}.$$

If ε_{ij} is the p order autoregressive process AR(p), i.e. $a_0\varepsilon_{ijt} + a_1\varepsilon_{ijt-1} + \ldots + a_p\varepsilon_{ijt-p} = z_t$, where z_t is a white noise, then to calculate $\mathbf{A}_T^{-1}$ the recurring algorithm given by Siddiqui (1958) can be used.

The testing of hypothesis of equality of the group means considered in this paper can be used for economical records, which often contain repeated measurements of one or more variables over time. In these situations, we have correlated observations.

REFERENCES

Goldberger A. S. (1972), *Teoria ekonometrii*, PWE, Warszawa.

Rao C. R, Mitra S. K. (1971), *Generalized Inverse of Matrices and its Applications*, John Wiley and Sons, New York.

Scheffe H. (1959), *The Analysis of Variance*, Wiley, New York.

Siddiqui M. M. (1958), *On the Inversion of the Sample Covariance Matrix in a Stationary Autoregressive Process*, Ann. Math. Stat., **29**, 585–588.

Yang M. C. K., Carter R. L. (1983), *One Way Analysis of Variance with Time Series Data*, "Biometrics" **39**, 747–751.