TIME SERIES ANALYSIS AND TRENDS BY USING SPSS PROGRAMME

RADMILA KOCURKOVÁ
Silesian University in Opava
School of Business Administration in Karviná
Department of Mathematical Methods in Economics
Czech Republic
Tel.: *+ 420 69 63 98 283*
Fax: *+ 420 69 63 12 069*
e-mail: *kocurkova@opf.slu.cz*

Key Words:
time series analysis, process ARIMA, unemployment, programme SPSS

Introduction

In my lecture I would like to tell you something about the time series, respectively about trends in the number of job applicants registered by labour offices in the Czech Republic. On the basis of results I will forecast the number of job applicants registered by labour offices till December 2000. My data was taken from statistic source: Ministry of Labour and Social Affairs.

At our university I teach the course Time Series Analysis for the students of system engineering and information specialisation. The main aim of this subject is: investigation of the regular course of time series; deviation investigation of expected regularity; prediction of the future time series; regular determination of the convenient control directions for influence of time series.

During the lessons I usually use MS Excel and statistic programme SPSS (Statistical Package for the Social Sciences). SPSS makes some operations with the time series possible in the basic module – Base; for deeper analysis it is necessary to join module *Trends*. For SPSS programme is the time series normal data file and it is presupposed that one row of the date nut contains the observation in one time and the rows ground in the way, that the oldest observation is the first, the youngest observation is the last row of the nut.

Time series analysis

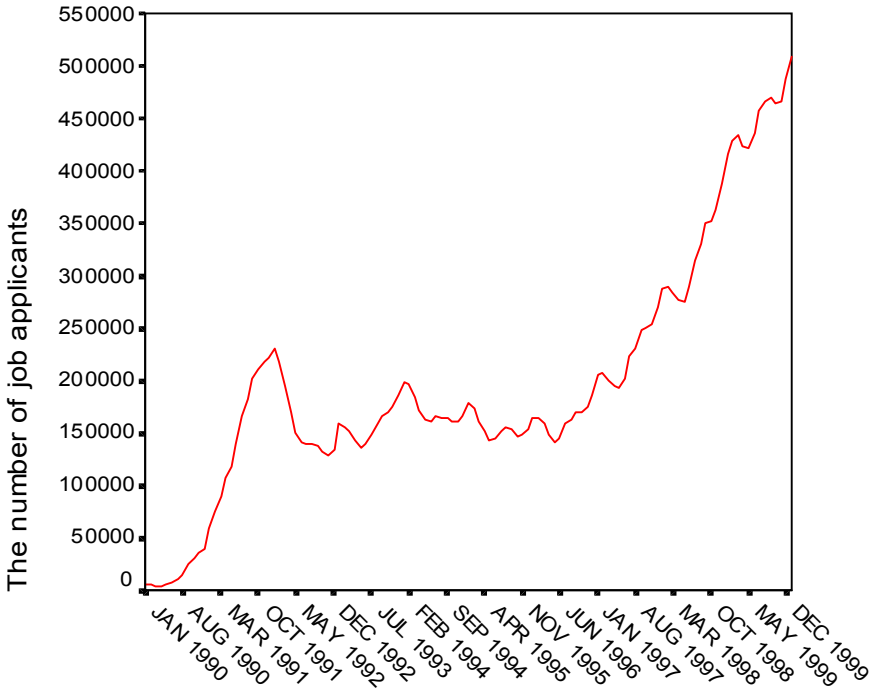
Firstly, we have to describe the trend of time series of the applicants registered by labour office in the Czech Republic.

From the Graph 1 we can say that January 1997 the number of job applicants grows much more than in the previous years. This time series is not stationary (which can be deleted by the seasonal difference of the time series).

Theory

Trend and prediction of time series can be computed by using ARIMA model. ARIMA (p,d,q) model is complex a linear model. There are three parts (they do not have to contain always all of these): AR (Autoregressive) – linear combination of the influence of previous values; I (Integrative) – random walk; MA (Moving average) – linear combination of previous errors. These models are very flexible, quite hard for computing and for the understanding of the results. They are demand quality and a large number of dealing dates (it is assumpted at least 50 dealing or observations).

Graph 1
Trends in the number of job applicants registered by the labour office: January 1990 – January 2000



Source: Ministry of Labour and Social Affairs

At first, we have to identify the type of the model and the values of the parameters. We can do it by using autocorelate (ACF) and partial autocorelation (PACF) function of the stationary time series. Autocorelate means the correlation's between time series and the same time series lag. Partial autocorrelations are also correlation coefficients between the basic time series and the same time series lag and we will eliminate the influence of the members between. This coefficient presents only about direct structure, for example between y_t and y_{t-2} with the elimination of transmission over the observation y_{t-1}.

In output of SPSS we compute signitications for the individuals parameters in ARIMA model. From these values we can determine if the parameters can be used in this model or not. In the case of computing more models we choose the model where AIC (Akai information criteria), respectively SBC (Schwartz-Bayes criteria) are minimal and Log likelihood is maximal. At the end we verify if the residual component is the white noise. In ARIMA models, we assume dependence between the quantities $y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t-2},...$

If this process contains the seasonal fluctuation, as it is in this model, we can expect also the dependence seasons: $y_{t-2s}, y_{t-s}, y_{ts}, y_{t+s}, y_{t-2s}, \ldots$, where s is the length of the period (in this case 12).

This process is called SARIMA (p,d,q) $(P,D,Q)_s$, where

p is order of process AR,
q is the order of process MA,
d is the order of difference,
P is order of seasonal process AR,
Q is the order of MA,
D is order of seasonal difference,
s is the length of seasonal period.

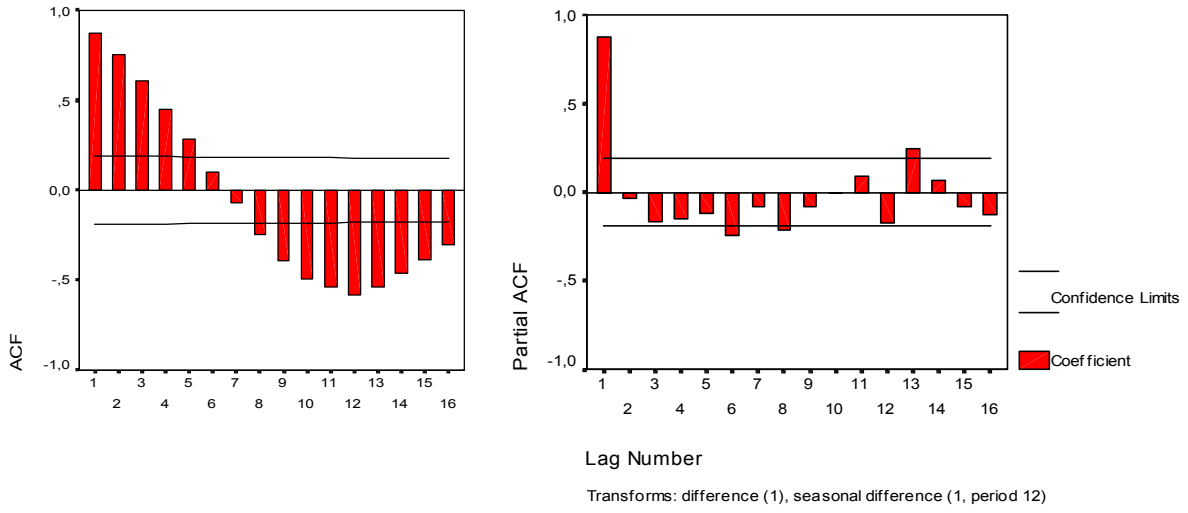
The equation of this model is: $\Phi_P\left(B^s\right)\phi_p(B)(1-B)^d\left(1-B^s\right)^D y_t = \theta_q(B)\Theta_Q\left(B^s\right)a_t$, where

$\phi_p(B)$ is autoregressive operator,
$\theta_q(B)$ is the operator of moving averages,
$\Phi_P\left(B^s\right)$ is seasonal autoregressive operator,
$\Theta_Q\left(B^s\right)$ is seasonal operator of moving averages,
$\{a_t\}$ is white noise.

Practice

For the fist we have to estimate the parameters of the model by using autocorrelation and partial autocorrelation functions. Because time series is not stationary we have to differentiate it, by the first order. We also assume the season dynamics that is why we differentiate time series also by seasonal fist order.

Graph 2, 3
Autocorrelation and partial autocorrelation function time series of the number of job applicants



Source: own calculations

On the basis of the form functions we choose model SARIMA (1,1,0) $(1,1,0)_{12}$ or model SARIMA (1,1,0) $(0,1,0)_{12}$. Akai information criteria and Schwartz-Bayes criteria are lower at the model SARIMA (1,1,0) $(1,1,0)_{12}$, we choose this model.

The shorter SPSS output for model SARIMA (1,1,0) $(1,1,0)_{12}$

```
Variable:   The number of unemployment applicant
Regressors: NONE
Non-seasonal differencing: 1          Seasonal differencing: 1
Length of Seasonal Cycle:  12         No missing data.

FINAL PARAMETERS:

Number of residuals   108             Standard error    4950,653
Log likelihood        -1072,8442
AIC                   2149,6884       SBC               2155,0526

           Variables in the Model:

               B           SEB       T-RATIO      APPROX. PROB.
AR1      ,83996600    ,05157667    16,285775       ,00000000
SAR1    -,43481456    ,09038694    -4,810591       ,00000501

11 new cases have been added.
```

General equation of the chosen model is: $\Phi_1\left(B^{12}\right)\phi_1(B)(1-B)^1\left(1-B^{12}\right)^1 y_t = a_t$.

After the modification of the equation and substitution of estimated values, we get the following equation and it describes the dynamics of our time series:

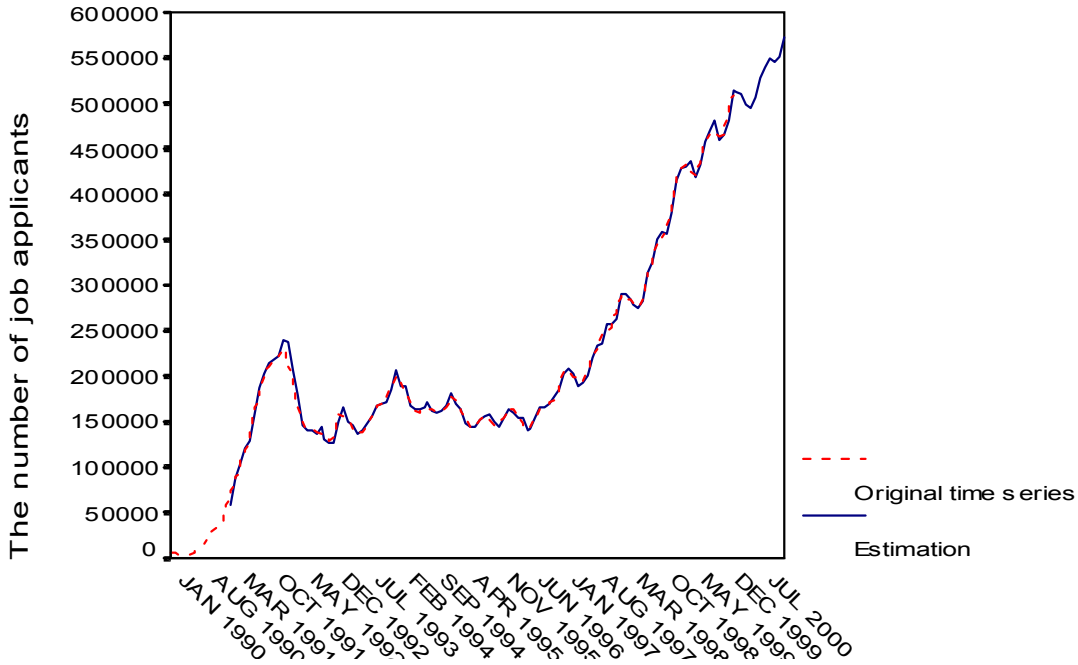
$$y_t = 1{,}84y_{t-1} - 0{,}84y_{t-2} + 0{,}57y_{t-12} - 1{,}1y_{t-13} + 0{,}48y_{t-14} + 0{,}43y_{t-24} - 0{,}79y_{t-25} - 0{,}84y_{t-26} + a_t$$

We can say, that the number of job applicants this month is much more influenced by the number of job applicants in the last month.

Résumé

We can describe graphically the original time series of the number of job applicants and the predicted time series by SARIMA (1,1,0) $(1,1,0)_{12}$.

Graph 4
The original and estimation values of the number of job applicants



Source: own calculations

Table 1 shows forecasting the time series in the number of job applicants registered by labour office in the Czech Republic.

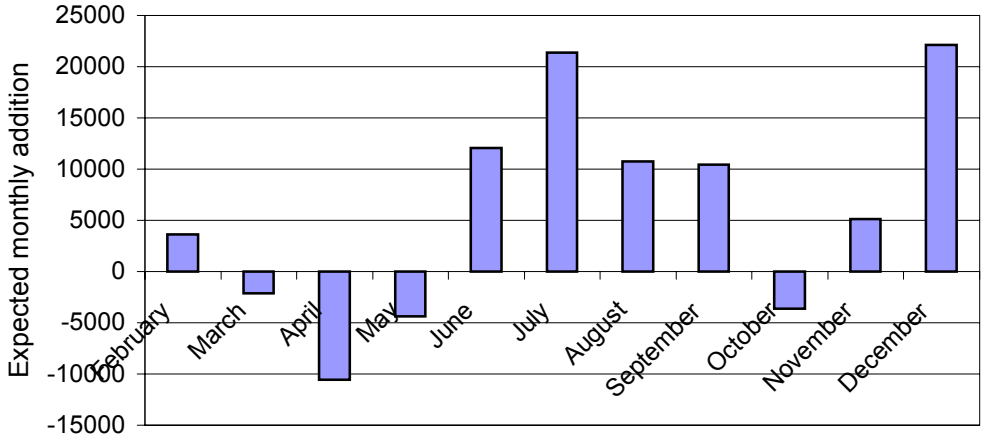
Table 1
Expected numbers of job applicants in year 2000

Month	The number of job applicants
February	512 094
March	509 978
April	499 388
May	495 024
June	507 073
July	528 465
August	539 225
September	549 669
October	546 019
November	551 165
December	573 290

Source: own calculations

In year 2000 we can expect the highest the numbers of job of applicants in December 573, 290. In comparison with the same period of year 1999 the number of job of applicants has extended by about 85, 667 applicants. The lowest number of job applicants we can expect in May 2000 and that is 495, 024. This is the consequence of the seasonal unemployment. In this period, there are some jobs for construction workers, agriculture workers, and young people will go abroad for a job.

Graph 5
Expected monthly addition of the numbers of job applicants in year 2000



Source: own calculations

In year 2000 we can expect the highest addition number of job applicants in December around 22, 125, July around 21, 392 (in this period graduated students from every school are registered by labour offices). The lowest number of job applicants we can expect in April 2000 – around 10, 590 and in October 3, 650 applicants.

Summary

The unemployment is a big problem for almost every country. That is way I decided to use SPSS software for the analysis of this problem in the conditions of the Czech Republic. After the Velvet revolution, the Czech Republic has the problem with unemployment, especially in its three previous years. Very dangerous is the situation in some place in the Czech Republic, where industry was the main sector of employment. School of Business Administration is in one of such parts country (Karviná) and there is about 69.4 applicants to one work place.

Literature

NORUŠIC, M. J. *SPSS Professional Statistics 7.5.* Chicago: SPSS Inc., 1997. ISBN 0-13-656935-8.

SPSS INC. *SPSS Trends 6.1.* Chicago: SPSS Inc., 1994. ISBN 0-13-201055-0.

DORNBUSCH, R., FISCHER, S. *Makroekonomie.* Praha: SPN, 1994. ISBN 80-04-25556-6.