ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 216, 2008

Maria Czogała[*]

ON THE PROPERTIES OF SOME PREDICTOR IN TIME SERIES ANALYSIS

ABSTRACT. Cieślak (1993) and Kohler and College (1988) considered a predictor being an arithmetic mean of a set of *k*-latest observations in time series, where *k* was constant. In this paper a modified predictor is presented and its properties are discussed. For each *t*-th observation the hypothesis that there is no change in the level of the time series is tested. When the hypothesis isn't rejected the predictor is an arithmetic mean of a set of *t*-latest observations otherwise the predictor is equal to the value of the last observation in the time series. The mean square error is used for assessing the error of prediction.

Key words: predictor, hypothesis testing, mean square error.

I. INTRODUCTION

Cieślak (1993) and Kohler and College (1988) considered a predictor being an arithmetic mean of a set of *k*-latest observations in time series, where *k* was constant. A change in level of time series is possible and in such situation the use of predictor being a moving average of a set of *k* observations would result in too high errors of prediction. A modification of the predictor which allows to reduce errors of prediction in situations when there is a change in a level of time series is proposed. For each *t*-th observation the hypothesis that there is no change in the level of this time series is verified. When this hypothesis is not rejected the predictor is an arithmetic mean of a set of *t*-latest observations, otherwise the predictor is equal to the value of the last observation – for which the change was notice ed. As always in hypothesis testing the type I error and type II error can occur and the error of prediction is to be the combination of the probability of making type I and type II error and the values of the errors of prediction in each case. The mean square error is used for assessing the error of prediction.

[*] MSc, Department of Statistics, University of Economics Katowice.

II. BASIC DEFINITIONS AND PROPERTIES OF THE PREDICTOR

Let us assume that a time series $\{Y_t, t = 1,2,...,n\}$ is given. Let us also assume that for the given time series the following assumptions are fulfilled:

$$\underset{t,h=1,2,...,n}{\forall} \quad E(Y_t) = \mu \quad \wedge \quad D^2(Y_t) = \sigma^2 \quad \wedge \quad Cov(Y_t, Y_h) = 0 \text{ for } t \neq h. \quad (1)$$

Let us consider predictor being an average of a set of k-latest observations (M. Cieślak, 1993):

$$Y_{Tp2}(k) = \frac{1}{k}\sum_{t=T-k}^{T-1} Y_t, \qquad T = k+1, k+2..., n+1. \quad (2)$$

The parameters of this predictor are as follows (J. Wywiał, 1995):

$$E\left(Y_{Tp2}(k)\right) = \mu \qquad D^2\left(Y_{Tp2}(k)\right) = \frac{\sigma^2}{k}. \quad (3)$$

$$Cov\left(Y_{Tp2}(k), Y_{Hp2}(k)\right) = \begin{cases} \dfrac{k - |H - T|}{k^2}\sigma^2 & dla \;\; |H - T| < k \\ 0 & dla \;\; |H - T| \geq k. \end{cases} \quad (4)$$

III. BASIC PROPERTIES OF THE PROPOSED TEST PREDICTOR

Let us assume that a time series $\{Y_t, t = 1,2,...,n\}$ is given. The assumptions set in (1) are fulfilled. A change in level of time series is possible. For each t-th observation in the time series the hypothesis that there is no change in the level of this time series is verified. The null hypothesis and the alternative hypothesis can be written as follows:

$$\begin{matrix} H_0: & \mu = \mu_1 \\ H_1: & \mu = \mu_2 \end{matrix}, \quad \mu_1 \neq \mu_2 \quad (5)$$

When we fail to reject the null hypothesis the proposed predictor is an arithmetic mean of a set of all t observations in the time series. When we accept the alternative hypothesis the predictor is equal to the value of only one – the last observation. This observation can be treated as the first observation of a new time series. Assuming that to the k-th observation there is no change in the level of time series and that the null hypothesis is not rejected to the k-th observation the test predictor can be defined as follows:

$$Y^{*}_{Tp2}(k)=\begin{cases}\frac{1}{k}\sum\limits_{t=T-k}^{T-1} Y_t, & \text{when } H_0 \text{ is accepted}\\ Y_{T-1}, & \text{when } H_0 \text{ is rejected}\end{cases} \tag{6}$$

As always in hypothesis testing two types of errors can occur. The error of prediction is to be the combination of the probability of making type I and II error and the values of the errors of prediction in each case. The mean square error is used for assessing the error of prediction.

3.1. Properties of the test predictor in case when there is no change in the level of time series and the null hypothesis is not rejected

Let us assume that a time series $\{Y_t, t=1,2,...,n\}$ is given and that there is no change in the level of the time series. Let us also assume that the assumptions set in (1) are fulfilled and that the test predictor is defined as in (6). The parameters of this predictor are as follows:

$$E\left(Y^{*}_{Tp2}\right)=\mu. \tag{7}$$

$$D^2\left(Y^{*}_{Tp2}\right)=\frac{\sigma^2}{k}. \tag{8}$$

The error of prediction can be written as follows:

$$U_{k+1}=\overline{y}_k-Y_{k+1}. \tag{9}$$

The variance of the error of prediction is calculated as follows:

$$D^2\left(U_{k+1}\right)=\sigma^2\left(1+\frac{1}{k}\right). \tag{10}$$

The mean square error of prediction for $T = k+1$ period is calculated as follows:

$$MSE\left(Y^{*}_{Tp2}\right) = MSE\left(\overline{y}_{T-1}\right) = E\left(\overline{y}_{k} - Y_{k+1}\right)^{2}$$

and so we get:

$$MSE\left(Y^{*}_{Tp2}\right) = \sigma^{2}\left(1 + \frac{1}{k}\right). \qquad (11)$$

3.2. Properties of the test predictor in case when there is no change in the level of time series but the null hypothesis is rejected

Let us assume that a time series $\{Y_t, t = 1,2,...,n\}$ is given and that there is no change in the level of the time series. Let us also assume that the assumptions set in (1) are fulfilled. The predictor is defined as in (6). It is assumed that in the $k+1$ period the null hypothesis is rejected, although there is no change in this time series.

The mean square error of prediction for $T = k+1$ period is calculated as follows:

$$MSE\left(Y^{*}_{Tp2}\right) = MSE\left(\overline{y}_{T-1}\right) = E\left(\overline{y}_{k} - Y_{k+1}\right)^{2} = E\left(\frac{1}{k}\sum_{t=1}^{k} Y_t - Y_{k+1}\right)^{2}.$$

Consequently:

$$MSE\left(Y^{*}_{Tp2}\right) = \sigma^{2}\left(1 + \frac{1}{k}\right). \qquad (12)$$

As a result of calculation the mean square error of prediction for $T = k+2$ period is obtained:

$$MSE\left(Y^{*}_{Tp2}\right) = 2\sigma^{2}. \qquad (13)$$

The mean square error of prediction for $T = k+3$ period:

$$MSE\left(Y_{Tp2}^{\bullet}\right)=\frac{3}{2}\sigma^{2}. \tag{14}$$

The mean square error of prediction for $T=k+s$ period is evaluated as follows:

$$MSE\left(Y_{Tp2}^{\bullet}\right)=\frac{s}{s-1}\sigma^{2}. \tag{15}$$

3.3. Properties of the test predictor when there is a change in the level of time series but the null hypothesis is not rejected

Let us assume that a time series $\{Y_t, t=1,2,...,n\}$ is given and that there is a change in the level of the time series:

$$\begin{aligned} &E(Y_t)=\mu_1 \quad t=1,2,...,k,\\ &E(Y_t)=\mu_2 \quad t=k+1,k+2,...,n,\\ &\mu_1\neq\mu_2. \end{aligned} \tag{16}$$

Let us also assume that:

$$D^{2}(Y_t)=\sigma^{2}, \quad t=1,2,...,n, \tag{17}$$

$$Cov(Y_t,Y_h)=0 \quad \text{for} \quad t\neq h \quad \text{and} \quad t,h=1,2,...,n. \tag{18}$$

The predictor is defined as in (6). It is assumed that a change in the level of the time series is observed in the $k+1$ period and that the null hypothesis is not rejected.

The mean square error of prediction for $T=k+1$ period is calculated as follows:

$$MSE\left(Y_{Tp2}^{\bullet}\right)=MSE\left(\bar{y}_{T-1}\right)=E\left(\bar{y}_k-Y_{k+1}\right)^2=E\left(\frac{1}{k}\sum_{t=1}^{k}Y_t-Y_{k+1}\right)^2$$

and as a result we get:

$$MSE\left(Y_{Tp2}^{\bullet}\right)=\sigma^{2}\left(1+\frac{1}{k}\right)+\left(\mu_1-\mu_2\right)^{2}. \tag{19}$$

The mean square error of prediction for $T = k + 2$ period is calculated as follows:

$$MSE\left(Y_{Tp2}^{*}\right) = MSE\left(\overline{y}_{T-1}\right) = E\left(\overline{y}_{k+1} - Y_{k+2}\right)^2 = E\left(\frac{1}{k+1}\sum_{t=1}^{k+1} Y_t - Y_{k+2}\right)^2.$$

It is then equal to:

$$MSE\left(Y_{Tp2}^{*}\right) = \sigma^2\left(1 + \frac{1}{k+1}\right) + \left(\frac{\mu_1 - \mu_2}{1 + \frac{1}{k}}\right)^2. \tag{20}$$

Consequently the mean square error of prediction for $T = k + s$ period is evaluated as follows:

$$MSE\left(Y_{Tp2}^{*}\right) = MSE\left(\overline{y}_{T-1}\right) = \sigma^2\left(1 + \frac{1}{k+s-1}\right) + \left(\frac{\mu_1 - \mu_2}{1 + \frac{s-1}{k}}\right)^2. \tag{21}$$

3.4. Properties of the test predictor in case when there is a change in the level of time series and the null hypothesis is rejected

Let us assume that a time series $\{Y_t, t = 1,2,...,n\}$ is given and that there is a change in the level of the time series. The assumptions set in (16), (17) and (18) are fulfilled. The predictor is defined as in (6). It is assumed that the change in the level of the time series is observed in the *k+1* period and that the alternative hypothesis is accepted.

The mean square error of prediction for $T = 1,2,...,k$ is as follows:

$$MSE\left(Y_{Tp2}^{*}\right) = MSE\left(\overline{y}_{T-1}\right) = \sigma^2\left(1 + \frac{1}{T-1}\right). \tag{22}$$

The mean square error of prediction for $T = k + 1$ period is calculated as follows:

$$MSE(Y^{\bullet}_{Tp2}) = MSE(\bar{y}_{T-1}) = E(\bar{y}_k - Y_{k+1})^2 = E\left(\frac{1}{k}\sum_{t=1}^{k} Y_t - Y_{k+1}\right)^2$$

and we obtain:

$$MSE\left(Y^{\bullet}_{Tp2}\right) = \sigma^2\left(1+\frac{1}{k}\right) + (\mu_1 - \mu_2)^2. \tag{23}$$

The mean square error of prediction for $T = k + 2$ period is equal to:

$$MSE\left(Y^{\bullet}_{Tp2}\right) = 2\sigma^2. \tag{24}$$

The mean square error of prediction for $T = k + 3$ period is equal to:

$$MSE\left(Y^{\bullet}_{Tp2}\right) = \frac{3}{2}\sigma^2. \tag{25}$$

The mean square error of prediction for $T = k + s$ period is evaluated as follows:

$$MSE\left(Y^{\bullet}_{Tp2}\right) = MSE\left(\bar{y}_{T-1}\right) = \frac{s}{s-1}\sigma^2. \tag{26}$$

3.5. Properties of the test predictor in general case

As always in hypothesis testing two types of errors can occur. The error of prediction is to be the combination of the probability of making type I and II error and the values of the errors of prediction in each case. It can be written as follows:

$$MSE\left(\bar{y}_t\right) = MSE\left(\bar{y}_t\right)|_{\text{failure of rejection the } H_0} \cdot P\left(\text{failure of rejection the } H_0\right) + \\ + MSE\left(\bar{y}_t\right)|_{\text{rejection the } H_0} \cdot P\left(\text{rejection the } H_0\right). \tag{27}$$

Let:

$$\alpha = P(U \geq u_\alpha \mid H_0), \tag{28}$$

$$\beta = P(U < u_\alpha \mid H_1). \tag{29}$$

Let us assume that there is no change in level of the time series. The null hypothesis and the alternative hypothesis are defined as in (5). Let us also assume that to the k-th observation there was no rejection of the null hypothesis. As the mean square error can be defined as in (27) the mean square error for the $T = k+1$ observation is calculated as follows:

$$MSE(Y^{*}_{Tp2}) = MSE(\bar{y}_k) = \alpha \cdot \sigma^2\left(1+\frac{1}{k}\right) + (1-\alpha)\sigma^2\left(1+\frac{1}{k}\right) = \sigma^2\left(1+\frac{1}{k}\right). \quad (30)$$

The mean square error for the $T = k+2$ observation is calculated as follows:

$$\begin{aligned} &MSE(Y^{*}_{Tp2}) = MSE(\bar{y}_{k+1}) = (1-\alpha)^2 MSE(\bar{y}_{k+1}) + \alpha(1-\alpha)MSE(\bar{y}_2) + \\ &+\alpha^2 MSE(\bar{y}_1) + \alpha(1-\alpha)MSE(\bar{y})_1 = \\ &= (1-\alpha)^2 MSE(\bar{y}_{k+1}) + \alpha(1-\alpha)MSE(\bar{y}_2) + MSE(\bar{y}_1)(\alpha^2+\alpha-\alpha^2) = \\ &= (1-\alpha)^2 \sigma^2\left(1+\frac{1}{k+1}\right) + \alpha\sigma^2\left[(1-\alpha)\left(1+\frac{1}{2}\right)+2\right]. \end{aligned} \quad (31)$$

The mean square error for the $T = k+ s$ observation is equal to:

$$MSE(Y^{*}_{Tp2}) = MSE(\bar{y}_{k+s-1}) = \sigma^2\left[(1-\alpha)^s\left(1+\frac{1}{k+s-1}\right) + \alpha\sum_{n=1}^{s}(1-\alpha)^{n-1}\left(1+\frac{1}{n}\right)\right]. \quad (32)$$

Let us now assume that in the $k+1$-th observation there is a change in the level of the time series. Let us also assume that to the k-th observation there was no change in the level of the time series and that there was no rejection of the null hypothesis.

The mean square error for the $T = k+1$ observation is calculated as follows:

$$\begin{aligned} &MSE(Y^{*}_{Tp2}) = MSE(\bar{y}_k) = (1-\beta)\sigma^2\left[\left(1+\frac{1}{k}\right)+(\mu_1-\mu_2)^2\right] + \\ &+\beta\sigma^2\left[\left(1+\frac{1}{k}\right)+(\mu_1-\mu_2)^2\right] = \\ &= \sigma^2\left[\left(1+\frac{1}{k}\right)+(\mu_1-\mu_2)^2\right]. \end{aligned} \quad (33)$$

The mean square error for the $k+2$ observation is calculated as follows:

$$MSE\left(Y_{Tp2}^{*}\right)=MSE\left(\overline{y}_{k+1}\right)=\beta^{2}MSE\left(\overline{y}_{k+1}\right)+\beta\left(1-\beta\right)MSE\left(\overline{y}_{2}\right)+$$
$$+\left(1-\beta\right)^{2}MSE\left(\overline{y}_{1}\right)+\beta\left(1-\beta\right)MSE\left(\overline{y}_{1}\right)= \tag{34}$$
$$=\beta^{2}\left[\sigma^{2}\left(1+\frac{1}{k+1}\right)+\left(\frac{\mu_{1}-\mu_{2}}{1+\frac{1}{k}}\right)^{2}\right]+\beta\left(1-\beta\right)\sigma^{2}\left(1+\frac{1}{2}\right)+2\sigma^{2}\left(1-\beta\right).$$

The mean square error for the $k+t$ observation is equal to:

$$MSE\left(Y_{Tp2}^{*}\right)=MSE\left(\overline{y}_{k+t-1}\right)=\beta^{t}MSE\left(\overline{y}_{k+t-1}\right)+\sum_{n=1}^{t}\beta^{n-1}\left(1-\beta\right)MSE\left(\overline{y}_{n}\right). \tag{35}$$

IV. CONCLUDING REMARKS

As nowadays being able to forecast the future from the present information has an essential role the methods of prediction are in constant progress. The better predictor we have the more precise information about future we can get from the information we have today and the less is the risk that we will take wrong decisions. The proposed predictor allows reducing the errors of prediction as it is sensitive for the information from the present period. The analysis shows that the predictor gives better results than the classic predictor.

REFERENCES

Cieślak M., Dittman P., Kania-Gospodarowicz A., Krupowicz J., Kwiatkowska D., Radzikowska B. (1993): *Prognozowanie gospodarcze*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

Kohler H., College A. (1988): *Essentials of Statistics*, Scott, Foresman & Company, Glenview, Illinois.

Wywiał J. (1995): *Weryfikacja hipotez o błędach predykcji adaptacyjnej*, Wydawnictwo Ossolineum, Wrocław.

Maria Czogała

WŁASNOŚCI PEWNEGO PREDYKTORA W ANALIZIE SZEREGÓW CZASOWYCH

W swoich pracach Cieślak (1993) oraz Kohler i College (1988) rozważali predyktor będący średnią arytmetyczną k-ostatnich obserwacji szeregu czasowego, gdzie k jest stałe. W pracy przedstawiona jest modyfikacja wspomnianego predyktora oraz omówione są jego własności. Zaproponowany predyktor jest średnią arytmetyczną k-ostatnich obserwacji szeregu czasowego, przy czym k nie jest wielkością stałą. Dla każdej t-kolejnej obserwacji szeregu czasowego weryfikowana jest hipoteza, twierdząca, że w poziomie szeregu czasowego nie nastąpiła zmiana. Gdy hipoteza zerowa nie jest odrzucona predyktor jest wyznaczany jako średnia arytmetyczna z wszystkich t-ostatnich obserwacji, w przypadku przeciwnym predyktor jest równy ostatniej obserwacji tego szeregu czasowego. Do oceny błędów predykcji wykorzystany jest błąd średniokwadratowy.