Journal of Business and Economics, ISSN 2155-7950, USA
November 2014, Volume 5, No. 11, pp. 2052-2056
DOI: 10.15341/jbe(2155-7950)/11.05.2014/009

http://www.academicstar.us

Using the Transfer Function Model in Analyzing and Forecasting Students' Studying Achievement

Fang Po Huang[1], *Berlin Wu*[2]
(1. National ChiNan University, Taiwan; 2. National Chengchi University, Taiwan)

Abstract: In recent years, the innovation and improvement of forecasting techniques have caught more attention, and also provide indispensable information in decision-making process. Since student's achievement time series may be influenced by many factors, it is difficult to find an appropriate linear regression model to get a better forecasting result for students' studying achievement. In this paper, we use ARIMA model and transfer function model in analyzing and forecasting student's achievement. In the empirical study, we demonstrate a novel approach to forecast the student's achievement through transfer model. The mean absolute forecasting accuracy method is defined and used in evaluating the performance forecasting. The comparison with moving average model is also illustrated.

Key words: transfer function model; forecasting; students' studying achievement
JEL codes: I21

1. Introduction

Many researchers have been dedicated to developing and improving time series forecasting models over the past several years. Being as an active research method, time series prediction has drawn (找出) significant attention for applications in variety of studies. There are many forecasting techniques including: exponential smoothing, autoregressive integrated moving average (ARIMA) model, GARCH model, neural networks and genetic algorithm and so on. Those methods, however, have drawbacks and advantages. ARIMA model is one of the most principal and widely used time series models. ARIMA models can be used to forecast water quality, air quality, epidemiology, consumers' expenditure, sales forecasting, energy price, ozone levels, ammonia concentration and so on. ARIMA models are applicablewhen the time series are stationary without missing data. However, they have limited accuracy due to its failure to forecast extreme cases or nonlinear relationship. On the other hand, linear regression model have been suggested as an alternative method for forecasting. These models also can change their structure based on internal or external information that flow through the network or system. These advantages make them attractive in predicting nonparametric nonlinear time series models.

There are not many literature discussing the students' achievement. It's revealed that unlike regression-based

Fang Po Huang, Department of Education Policy and Advertisement, National ChiNan University; research areas/interests: education policy. E-mail: hfb@mail.ttvs.ntct.edu.tw.
Berlin Wu, Professor, Department of Mathematical Sciences, National Chengchi University; research areas/interests: nonlinear time series analysis and forecasting; artificial intelligence; statistics and its applications; financial statistics, econometrics. E-mail: berlin@math.nccu.edu.tw.

model, transfer function model did not generate any prior assumption regarding any functional form or produced the time-span for prediction.

The results exhibited satisfactory accurate prediction on student's performance. Furthermore, student's chinese performance and english performance are correlated, it will not be the single determinant of chinese performance. The research completed by Starr et al. showed that the relationship between defense spending and inflation was mutually related in France and Germany. Chan's study proved that spending tended to be more import-demanding in the developing countries and was possible to generate domestic inflation.

2. Methodology

To analyze time series data and make accurate forecasting are motivated many researchers in several fields, ranging from the natural sciences, economics, and management related disciplines. It is well noticed that ARIMA models are designed for predicting linear data, while ANNs models are suitable for data with nonparametric and nonlinear patterns. Obviously each model possesses its own strength and has different applications. Based on single variable ARIMA models and single variable as well as multivariate ANNs models to conduct forecasting. This study applies the mean absolute error (MAE) approach to evaluate prediction accuracy.

2.1 Why We Use Transform Functions

For more than three decades, ARIMA linear models have dominated many fields of time series prediction. In an ARIMA (p, d, q) model, the future value of a variable is supposed to be a linear function of several past observations and random errors. The general form is shown as follows:

$$\emptyset(\mathrm{B})\nabla^d(Y_t-\mu)=\Theta(\mathrm{B})\varepsilon_t \quad (1)$$

Where Y_t is the underlying time series, ε_t is the white noise with zero mean and variance σ^2, ∇=(1−B), B is the backward shift operator, d is the integer number of regular differencing and $\emptyset(\mathrm{B})= 1-\sum_{i=1}^{p}\varphi_1\mathrm{B}^{\mathrm{i}}$, $\Theta(\mathrm{B})= 1-\sum_{j=1}^{q}\theta_j B^j$ are polynomials in B, p and q are integers orders of the model.

The ARIMA modeling method includes three steps: model identification, parameter estimation, diagnostic checking. Stationary is necessary for an ARIMA model to predict. Data transformation is required to generate the stationarity of these time series. The first step in model identification is that if a time series is generated from an ARIMA process, it should have some autocorrelation properties. It is likely toidentify one or more feasible (可行) models for the given time series. The temporal correlation structure of the sample data is proposed to use the autocorrelation function (ACF) and partial autocorrelation function (PACF) to identify the order of the ARIMA model1. The model that gets the minimum Akaike Information Criterion (AIC) is chosen as the optimal model. After the functions of the ARIMA model have been specified, estimation of the model parameters is forward. When the fitting model is chosen and its parameters are estimated, the Box-Jenkins methodology requires to examine the residuals of the model is minimized. It can be achieved using a nonlinear optimization process. The last step is diagnostic checking of the model. Several tests are operated for diagnostic check to determine whether the residuals of the ARIMA models from the ACF and PACF graphs are independent and identically distributed.

Transfer functions are commonly used in the analysis of systems such as single-input single-output filters, typically within the fields of signal processing, communication theory, and control theory. The term is often used exclusively to refer to linear, time-invariantsystems (LTI), as covered in this article. Most real systems have non-linear input/output characteristics, but many systems, when operated within nominal parameters (not

"over-driven") have behavior that is close enough to linear that LTI system theory is an acceptable representation of the input/output behavior.

As a good prediction model, the residuals are used to examine the goodness of fit of the model that meets the requirements of a white noise process. If the model is not suitable, a new model should be identified. The steps of parameter estimation and diagnostic checking are repeated many times until an optimal model is selected. The last selected model is used to forecast the value.

2.2 Integrated Procedures for Modeling Transfer Function

Step 1: Find the order for the autoregressive order

Step 2: Decide the order of time lag for impact factors

Step 3: Parameter estimation

Step 4: Model forecasting

2.3 Measurement of Prediction

MAE approach is frequently adopted as the measurement criteria of prediction accuracy in a fitted time series. MAE is mainly used to measure the percentage of unexplained part of a model constructed. Therefore, the smaller the MAE value obtained may indicate that more accuracy of the model will be. Also, it means that a better match exists between the historical data and the estimation result of the forecasting model. MAE equation is shown as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}| F_t - A_t |$$

3. Empirical Study

Figure 1 is a plot of the students' test score for the Chinese course from a high school during 2010-2013. The last one data (at size 18) is the score of University entrance exam which we want to forecast.

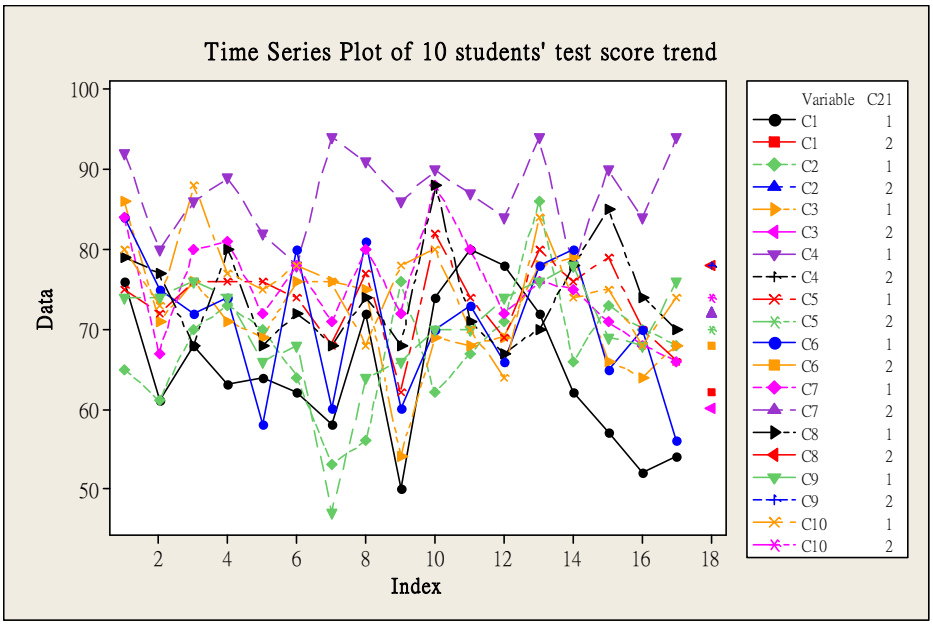

Figure 1 Time Series Plot of 10 Students' Test Score Trend

First we use the AR(1) and ARMA(1,1) to do the forecasting work. The comparison result was shown at the Tables 1-3 (male), Tables 4-6 (female).

Table 1 Comparison the Forecasting Performance with Two Models (Chinese)

Observation	1	2	3	4	5	6	7	8	9	10	MAE
Real value	62	72	60	78	70	68	72	78	78	74	
Forecasting with AR(1)	62	68	71	83	77	78	77	75	73	76	5.2
Forecasting with ARMA(1,1)	61	68	71	83	77	78	76	72	72	76	5.6

Table 2 Comparison the Forecasting Performance with Two Models (English)

Observation	1	2	3	4	5	6	7	8	9	10	MAE
Real value	86	28	52	98	72	62	46	72	70	60	
Forecasting with AR(1)	83	36	72	94	63	75	64	66	54	62	9.9
Forecasting with ARMA(1,1)	82	50	69	94	67	71	59	67	46	55	10.8

Table 3 Comparison the Forecasting Performance with Two Models (Mathamatics)

Observation	1	2	3	4	5	6	7	8	9	10	MAE
Real value	64	64	60	80	72	60	52	72	44	76	
Forecasting with AR(1)	86	64	75	91	73	74	51	61	60	67	10.0
Forecasting with ARMA(1,1)	82	50	69	94	67	71	59	67	46	55	10.6

Table 4 Comparison the Forecasting Performance with Two Models (Chinese)

Observation	1	2	3	4	5	6	7	8	9	10	MAE
Real value	60	72	66	48	92	68	68	72	70	52	
Forecasting with AR(1)	72	74	81	72	86	75	70	81	71	55	8.1
Forecasting with ARMA(1,1)	48	61	89	54	83	72	69	77	59	42	8.2

Table 5 Comparison the Forecasting Performance with Two Models (English)

Observation	1	2	3	4	5	6	7	8	9	10	MAE
Real value	70	68	82	40	82	88	78	86	76	50	
Forecasting with AR(1)	65	82	91	43	84	95	79	84	76	58	5.1
Forecasting with ARMA(1,1)	51	68	95	65	83	95	73	76	77	46	8.5

Table 6 Comparison the Forecasting Performance with Two Models (Mathematics)

Observation	1	2	3	4	5	6	7	8	9	10	MAE
Real value	68	72	60	40	52	92	76	76	80	24	
Forecasting with AR(1)	68	60	66	28	66	81	68	79	67	28	7.5
Forecasting with ARMA(1,1)	62	65	80	20	75	89	81	79	51	29	12.1

Since we feel that the Math and English achievement might impact the Chinese study achievement. We will take these two courses into consideration on the model construction. In order to see the forecasting performance, let's start by adding Math&Eng test into considerations, then use mean value of Math + English test score into considerations. They construct the model. That is we use

First model is Chinese X_{t+1} vs Chinese X_t and English Y_t

$$X_{t+1}=\Phi_1 X_t + \Phi_2 Y_t + \varepsilon_{1t}$$

Secondly, Chinese X_{t+1} vs Chinese X_t and Math Z_t

$$X_{t+1}=\Phi_1 X_t + \Phi_2 Z_t + \varepsilon_{1t}$$

Third, Chinese X_{t+1} vs Chinese X_t and mean of (Eng Y_t +Math Z_t)

$$X_{t+1}=\Phi_1 X_t + \Phi_2 (Y_t + Z_t) + \varepsilon_{1t}$$

Table 7 Comparison the forecasting performance with three models.

Table 7 Comparison the Forecasting Performance with Three Models

Observation	1	2	3	4	5	6	7	8	9	10	MAE
Real value	62	72	60	78	70	68	72	78	78	74	
Forecasting with X_{t+1} vs X_{t+1} and Y_t	61	69	64	83	76	77	76	74	72	75	4.3
Forecasting with X_{t+1} vs X_{t+1} and Z_t	60	68	71	82	78	77	76	75	74	76	5.1
Forecasting with X_{t+1} vs X_{t+1} and Y_t and Z_t	60	64	66	70	69	69	61	65	58	60	8.4

4. Conclusion

The results reveal that the single variable ARIMA models based on data obtained from past 17 tests. For single variable models, the ARIMA models show stability and accuracy across all three disciplines, while Transfer function models show more "good" accuracy. Therefore, the transfer function models will be a better model to predict students' academic performance of college entrance exam.

When AR(1) and ARIMA are adopted to predict the grades, the results from AR(1) performs well both in gender and subjects. As for subjects, the result would approach the real date by using AR(1) to predict male students' performance on Chinese and female students' performance on English.

However, when using transfer function models to predict the grades, the results would be better if English grades rather than Chinese grade are included in calculation. The results are proved ineffective when both mathematic and English grades are used. It is also a reason which we must consider when adopting transfer function models to predict the college entrance exams grades of students

This research provides a new method to predict students' academic performance. Using normal grades as time series and comparing the predicted grades with real grades to exam the effect of prediction model, this method could provide teachers and students to predict the performance of college entrance exam in advance, which can also provide suggestion to improve academic performance and analyze potential college departments.

Further studies are needed to use more variables with various research methods trying to compare the prediction accuracy made by different approaches for teachers and students as references

References:

Lee C. C. and Chen S. T. (2007). "Non-linearity in the defense expenditure—Economic growth relationship in Taiwan", *J. Defence and Peace Economics*, Vol. 18, No. 6, p. 537.

Mishra A. K. and Desai V. R. (2005). "Drought forecasting using stochastic models", *J. Stochastic Environmental Research Risk Assessment*, Vol. 19, p. 326.

Pisoni E., Farina M., Carnevale C., and Piroddi L. (2009). "Forecasting peak air pollution levels using NARX models", *J. Engineering Applications of Artificial Intelligence*, Vol. 22, p. 593.

Sousa S. I. V., Martins F. G., Pereira M. C., and Alvim-Ferraz M. C. M. (2006). "Prediction of ozone concentrations in Oporto City with statistical approaches", *J. Chemosphere*, Vol. 64, p. 1141.

Shahwan T. and Odening M. (2007). *Computational Intelligence in Economics and Finance*, Springer, Berlin Heidelberg, New York, p. 63.

Schlink U., Herbarth O., Richter M., Dorling S., Nunnari G., Cawley G. and Pelikan E. (2006). "Statistical models to assess the health effects and to forecast ground-level ozone", *J. Environmental Modelling & Software*, Vol. 21, p. 547.