Mathematical Forecasting of the Tourism Activity Importance in Chinese Economy based on Holt's Exponential Smoothing Model

Juxing Wang[1, a]

[1] Southwestern University of Finance and Economics, Chengdu, China

[a] edixingw@gmail.com

Keyword: Tourism Activities, Forecasting of Economy, RMSE, Holt's Model

Abstract: As economy grows, the tourism shares a more and more important part in the development. This paper aims to forecast the importance of the tourism activities in Chinese economy. We got the initial values of forecast by linear regression based on the data on percentage of tourism industry in terms of employment from 1988 to 2015, and used Holt's exponential smoothing model to get the optimal RMSE. Finally, we get the predicted values from 2016 to 2018 with minimized RMSE.

Introduction

Background: Tourism, the exploration of body and soul, has been the phenomenal growth industry in the economy. According to Chinese State Information Center, the tourism industry contributes to 10.1% growth of GDP in 2015. The driving effect to consumption, investment, import and export is increasing year by year, tourism has become an important engine of the troika.

Tourism industry promotes a new growth point of China's reform of the supply side. In 2015, China's tourism gained more than 4 trillion yuan revenues, and direct investment exceeded one trillion for the first time, which reached 1.0072 trillion yuan. In this year, 27.98 million people employ for the tourism directly, 79.11 million people is related to the indirect employment in China, which accounts for 10.2% of the total employment in the country. According to statistic, tourism employment of rural labor force accounted for more than 60% of the tourism employment, rural labor force accounted for more than 69% of the new tourism employment. Tourism plays an important role in the transfer of rural surplus labor force employment, and it has become an important industry to promote consumption, adjust structure and benefit the people in China.

Literature review: Tourism modelling and forecasting methods have developed in recent years. Time series analysis is a kind of dynamic data processing method of statistics. The method based on random process theory and mathematical statistical methods, complying with the statistics law of random data sequence. Box and Jenkins (1970) proposed integrated autoregressive moving-average models (ARIMAs) to deal with the tourism forecasting. But Goh and Law (2002) thought ARIMA model did not perform well, and thought SARIMA models worked better in some condition. Chan et al (2005) used GARCH models to analyze tourism and some shocks in the process. Song and Turner (2006) divided the quantitative tourism forecasting methods into two categories, non-causal time series models and the causal econometric approaches.

Our Work: Forecasting is an essential analytical tool in tourism policy and planning. This paper

provides expertise on projections of importance of the tourism activities in the economy. We forecast the percentage of importance of employment from 2016 to 2018 by Holt's method, which bases on the data on percentage of tourism industry in terms of employment from 1988 to 2015.

Notations

Symbol	Definition
	Smoothing constant for the average
	Smoothing constant for the trend
	Time period, the time between the currentyear and the base year (1988)
	Exponentially smoothed forecast
	Exponentially smoothed trend
	The actual percentage of tourism industry in terms of employment
	The predicted percentage of tourism industry in terms of employment
	Value(intercept) of $\hat{y}_t$ at t=0
	Slope of the regression line
	The root mean square error

Forecasting

Data analysis: We use MATLAB to plot the scatter diagram (Figure 1(a)) of the actual percentage of tourism industry in terms of employment, and add the actual line to it, we can see that Y increases as time pass by. The plot reveals the existence and nature of the trend whose direction is positive.

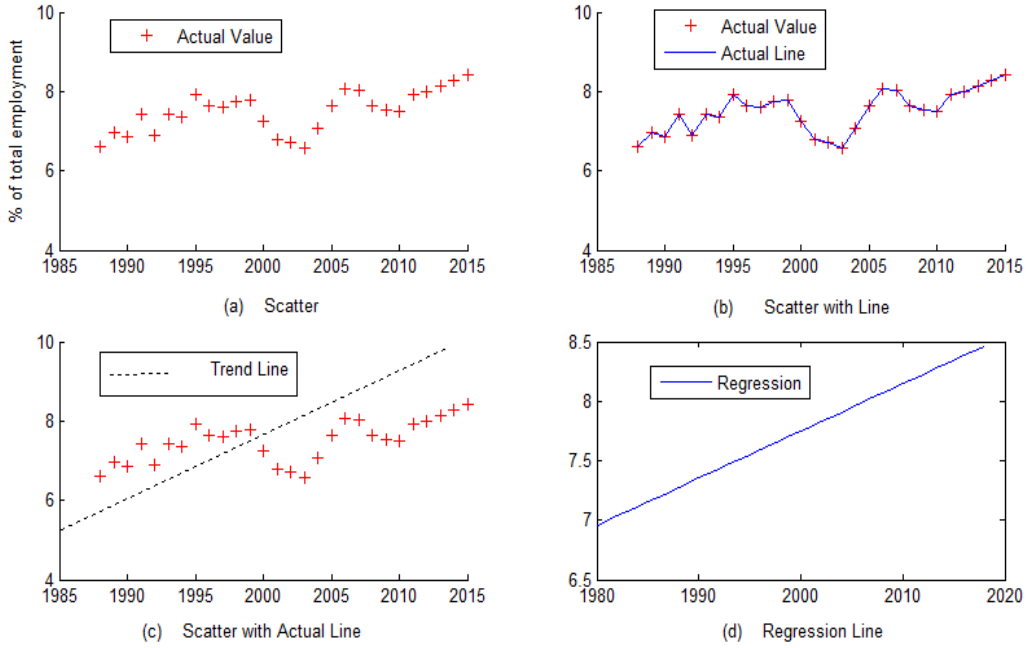

Fig. 1 The data analysis

- Trend projection

We use Excel to analyze the regression, independent variable is the time period and the dependent variable is the actual observed value in the time series.

(1)

We get the results in Table 1, and draw the regression line in Fig. 1(d).

Table 1 The results of Linear Regression

	Coefficient	Std. Error	Prob.
Intercept	,(39.9219)		0.0000
	,(3.5481)		0.0016
（Adj- ）	0.3349,(0.3083)		

(The brackets of the estimate value is the corresponding t value,***,**,*mean the significance at the level of 1%, 5% and 10%.)

- Hypotheses Testing

Knowing the sampling distribution for the standardized estimator allows us to carry out hypothesis test.

T test:

H_0:

H_a:

We can get the values of the variables, the P-value of intercept is 0.0000 and the P-value of t is 0.0016, both are less than 0.01. So, we can reject all the H_0 at the significance level of 1%. All the variables are statistically significant at the 1% level.

F test:

H_0:

H_a:

Just as with t statistics, P-value can be calculated by looking up the percentile in the appropriate F distribution. Using MATLAB, we get the P value of F, P=0.001564<0.001, so we reject H_0 when . At least one variable can explain Y.

Exponential smoothing with trend adjustment: Holt's method makes use of two different parameters and allows forecasting for series with trend. So we can use Holt's method. When a trend is present, exponential smoothing needs to be modified: Forecast including trend = Exponentially smoothed forecast + Exponentially smoothed trend.

- Smooth initial forecast

The forecast at time t is based on weighting the most recent observation with a weight and weighting the most recent forecast at time (t-1) with a weight of (1-), where is the smoothing constant.

(2)

- Smooth the trend

(3)

- Trend-adjusted forecast

, (4)

- Initialization

The initialization process for Holt's linear exponential smoothing requires two estimates. We have got the initial estimate of level and trend by running a linear regression Table 1.

(5)

- Evaluating forecast- Minimize RMSE

Root mean square error(RMSE) is one of the most commonly used measures of forecast accuracy. And we use Excel Solver to minimize RMSE by changing at the same time.

(6)

Results: We get the optimal value of when RMSE gets its minimized value 10.2336. And at this time, we can get the predicted value in 2016, 2017, 2018 without the actual value of the three years.

Table 2 The results of Forecasting

Time	Y_hat	Optimal Value			Value
2016	6.442937	Apha	0	a	0.036673
2017	7.851091	Beta	0.301279	b	6.974123
2018	6.215399	RMSE	10.32131		

Conclusion

From the analysis above, we can get the Figure 2. The actual and predicted value of percentage of tourism industry in terms of employment.

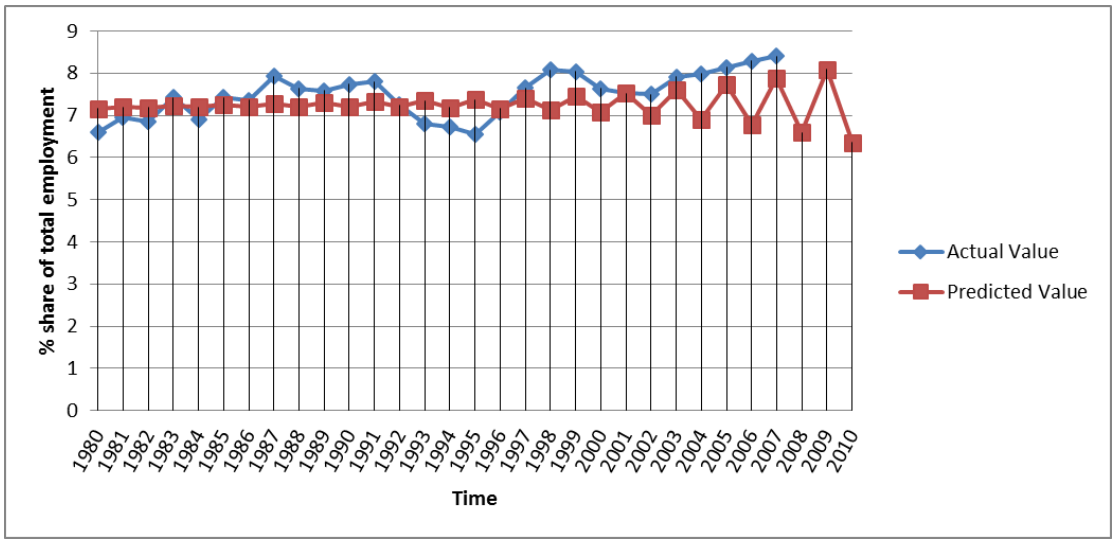

Fig. 2The actual and predicted value of Y

In 2016, it seems that tourism employment will account for 6.44% of the economy, and it reaches 7.85% in 2017. But it will fluctuates around 7 these years, and it may drop to 6.22% in 2018 as the

figure shows. For policy makers and the people concerned with Chinese tourism, the development of tourism may a little different with what they speculated, just as W. Edwards Deming said, “In God we trust, all others bring data.” So we can use the forecasting of mathematical model to make plans and policies.

References

[1] Athiyaman A, Robertson R W. Time series forecasting techniques: Short-term planning in tourism[J]. International Journal of Contemporary Hospitality Management, 1992, 4(4).

[2] Chu F L. Forecasting tourism demand in Asian-Pacific countries[J]. Annals of Tourism Research, 1998, 25(3): 597-615.

[3] Song H, Li G. Tourism demand modelling and forecasting—A review of recent research[J]. Tourism Management, 2008, 29(2): 203-220.

[4] Chen R J C, Bloomfield P, Cubbage F W. Comparing forecasting models in tourism[J]. Journal of Hospitality & Tourism Research, 2007.

[5] Opening of the first world conference on tourism development, Global travel comes into the "Chinese time", http://www.csunews.com/news/dt/34996.html