The Model of the Distribution of Taxicab Resources

Ruiqi Zhan

Institute of Electrical and Electronic Engineering of North China Electric Power University, Baoding 071000,China

739131665@qq.com

Key words: Dual Screening, Dual Prediction, Fractional Quantification, Neural Network.

Abstract. Based on the first question of problem B in the National Undergraduate Mathematical Modeling Contest of the Higher Education Cup in 2015, this model would quantify some relevant indexes to analyze the degree of supply and demand for taxicab resources.

In view of this question, this paper would estimate the distribution of taxi resource according to the analysis of the supply and demand, and establish the model of the Time and Neural Network. The assumed space-time condition is different periods and districts in one city. The supply quantity could be available through statistics, so the question is to work out the taxi demand and measure the degree of allocation of demand and supply according to the ratio of demand and actual supply (the arithmetic product of supply quantity and the rate of attendance) in the specific time and space. Firstly, to prove selecting indexes relevant to the demand for taxicabs through relativity analysis. The dimensionality of the variable could be reduced by filtering the main factors through principal component analysis. Secondly, to predict the data of principal components during 2009~2015 according to Time Series Analysis. Since the indexes would increase to the upper limit, the three Exponential Smoothing is more adoptable. The Neural Network prediction model would be established due to its high precision, and then the principal index during 2009~2015 could be input into it and the demand would be output. Finally, to estimate the demand in specific time and space and calculate the allocation degree of demand and supply on the ground of the population ratio in different periods and districts. The relativity analysis proves selected indexes relevant to the demand quantity. The principal component analysis filtered the population in downtown and operating distance. The time sequence predicts the principal component data, and the Neural Network forecasts the demand quantity. the distribution degree of demand and supply is a equilibrium value at 70%. it would be not difficult for people to catch taxis if it is within 60% to 75%. The selected indexes were based on dual analysis, and the prediction dual prediction, which reflects the high precision of this model.

Introduction

The taxi-taking software has emerged in taxi industry in such a era of Internet+, however, it still remains unknown what effects on earth it would act upon taxi drivers and the difficulties in taking taxis for the whole society. Our model is mainly to analyze the degree of supply and demand of taxicab resources and the effect degree of the subsidy scheme launched by software companies.

The Neural Network model based on Time Sequence

Problem analysis. the comprehension about the view of time and space: the time is the different period within one day, and the space is the different districts(commercial district, residential area, industrial area,etc.) in one city. And the problem should be working out the demand for taxicabs from a view of time and space in certain district, and measuring the allocation degree of demand and supply with the ratio of demand quantity and supply quantity.

The selected indexes could be proved to be relevant to the demand for taxis through the SPSS relativity analysis. The main indexes that influence the demand quantity in Nanchang city[1] could be figured out by principal component analysis. And the indexes of the relevant principal component data could be worked out through the time sequence model, and eventually, the demand for taxis during 2009~2015 would be predicted by the Neural Network. And then we take the time and space

into consideration. Due to the assumption that the population proportion of different districts in a city remain unchanged[2] while the demand quantity in this district and its population proportion are in direct proportion, the demand for taxis in different districts could be worked out, which is the demonstration of the space view. Because of the assumption that the proportion of pedestrians in different periods remains unchanged[3], the demand in different periods could be worked out, which is the demonstration of the time view. The final data in the period and space during 2009~2015 would be chosen as a model to explain. The maintenance ratio of taxis and taxis' attendance should be taken into consideration, so the formula should be rectified: adding the attendance index. To estimate the allocation degree of supply and demand should be based on the ratio of the demand quantity in certain place or certain period and the actual supply quantity in that place or that period.

Model establishment.①the data related to the demand for taxis in different cities with the means of SPSS Relativity Analysis would be shown in the following Table 1. The relativity is analyzed through SPSS Relativity Analysis.

Table 1:the indexes Table of major domestic cities[4]

	the Area	the Population	Per Capito GDP	Bus Number	the Bus Freight Mileage	Price	Taxi Number
Shijiangzhuang	455.8	200	25476.06	1865	727	9.2	6553
Hohhot	2504	85	11788.62	423	327	9	6959
Changchun	3583	292.8	21335.72	3024	739	8.1	12537
Shenyang	3495	185	19406.97	3590	1553	10.34	16735
Harbin	1637	329	18244.04	3165	721	11.8	11556
Nanjing	976	264.9	27128.08	4362	7296	10.2	8956
Hangzhou	683	128.76	38246.9	2038	656	12	6500
Fuzhou	224	70	31582.03	1463	367	11.6	3473

To analyze each index data with SPSS software and find out the Relativity Matrix. The Formula of the pearson relativity coefficient is:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

②Principal Component Analysis

To utilize Principal Component Analysis based on the data of Nanchang city from 1997 to 2008 and reduce the variable dimension.

Table 2: the Data of Nanchang city from 1997 to 2008

Year	Urban Population	the Mileage of Urban Operation Road	the Expense of Residents' Trips and Communication	GDP	the Number of Operating Buses	the Freight
1997	155.74	1.0495	0.0304	3752067	604	2429
1998	159.56	1.081	0.0393	3992606	68	2380
1999	164.21	1.096	4.1801	4237630	735	2338
2000	168.26	1.2385	0.0499	4651411	867	2300
2001	172.45	1.2812	0.0719	5245868	952	2144
2002	180.99	1.3503	0.0817	6019950	1160	2132
2003	196.37	1.4194	0.0936	7054437	447	2120
2004	203.73	1.9682	0.1033	85111066	1650	2110
2005	214.52	2.517	0.1276	10077025	1824	2100
2006	221.32	2.668	0.1347	11838973	2229	2091
2007	222.79	2.819	0.2337	13898920	2448	2083
2008	223.09	2.97	0.3028	16600847	2518	2076

If the contribution rate of the analyzed component rises to 80%, it would be the principal component which could replace all the other components.

③Time Sequence Model

Due to the assumption of different upper limits of each index, we adopt the cubic exponential smoothing method of Time Sequence to predict the parameters of the principal components during 2009~2015.

④Neural Network

On account of the high precision of Neural Network, it is wise to adopt BP Neural Network algorithm to figure out the demand for taxicabs during 2009~2015.

The condition,the parameter $\alpha > 0$ in the activation function $\varphi(v) = \frac{1}{1+\exp(-\alpha v)}$ should be satisfied so as to control the slope. The input signal is the main indicator through principal component analysis, and the output signal is the demand for taxicabs in this city.

⑤Provided that the population proportion remains unchanged, and the population proportion in residential, commercial, and industrial districts is respectively 50%, 5%, 10%, the demand for taxicabs in different districts should be: $N_I = n_i * D$

Provided that the peak hours of different proportions in sequence are:7：30-9：00、16：30-19：00、21：00-22：00. And the normal hours is: the proportion of the extra hours and the residents' trips on the same day

0.15+0.09+0.04+0.095+0.085+0.02=0.48

Therefore, the demand for taxicabs in different districts should be: $M_I = m_i * D$

⑥the supply quantity of taxicabs: $G = S \times \alpha$; "S" stands for the actual amount of taxicabs the city owned, and α should be 80%, then the degree of supply and demand could be estimated on the ground of the passenger load factor. $\beta = \frac{D}{G}$

It is a best allocation degree for β to float in the vicinity of 0.7. If it ranges within 60-75%, it is a demonstration that the difficulty in taking taxis is not severe, and if it is below 60%, it indicates that the difficulty in taking taxis doesn't exist and the the supply quantity of taxis exceeds the demand quantity. If it is above 75%, that means the problem of taking taxis is quite severe.

Model Solution.

①SPSS Relativity Analysis

Table 3:the Correlation Coefficients of each Index

The Area	The Population	Per Capito GDP	The Number of Buses	The Mileage of Public Transport Freight	Price
0.736	0.612	0.662	0.574	0.575	0.665

②Principal Component Analysis

The contribution rate of each index:

Table 4: The Contribution Rate of Each Index

Urban Population	The Mileage of Urban Operation Road	The Expense of Residents' Trips and Communication	GDP	The Number of Operating Buses	Freight
0.636	0.158	0.1428	0.0485	0.0136	0.0011

Then the selected indexes are: population in downtown, The mileage of Urban operation road

③Time Sequence Model

The data during 2009~2015 figured out with the cubic exponential smoothing method respectively are:(year, urban population/ten thousand, the mileage of urban operating road/km)2009，223.04，3159.2）（2010，220.5，3111.6）（2011，216.52，3446.4）（2012，211.08，3563.8）（2013，204.2，3663.6）（2014，195.86，3745.9）（2015，186.07，3810.7）

④Neural Network

The demand for taxicabs during 2009~2015 worked out with Neural Network is:(year, the demand for taxis)（2009，3739.7）（2010，3730.4）（2011，3903.7）（2012，3984.2）（2013，3980.9）（2014，3354.1）（2015，3986.5）

⑤the Time-Space view

The demand for taxicabs in different districts based on time-space view. We just take the data during 2009~2015 as an example, the demand for taxicabs based on time-space view is:

Table 5: the demand for Taxicabs in Different Districts

Year	Commercial District	Industrial District	Residential District
2014	177.705	355.41	1777.05
2015	154.325	308.65	1543.25

The demand for taxicabs in different periods. We just take the data during 2009~2015 as an example, the demand for taxicabs based on time-space view is:

Table 6:The Demand for Taxicabs in Different Periods

Year	In Peak Hours	In Normal Hours
2014	1705.968	1848.132
2015	1481.52	1604.98

⑥The supply quantity in each district and each period could be calculated out with the help of specific observation instruments, and the computational Formula could be derived from the Formula 9, therefore, we could estimate the specific allocation degree of the demand and supply of taxicabs in different periods and different districts and the exact severity of the problem of taking taxis in society.

Summary

According to the SPSS Relativity Analysis, the correlation coefficients of different indexes are below 0.5, and all the selected indexes are related to the demand for taxicabs. It could be learned from Principal Component Analysis that the cumulative contribution rate of urban population and the mileage of urban operation road amounts to over 80% with the downtown and the operation mileage taken as the principal components. The index data is predicted by Time Sequence. Based on the Neural Network, the demand for taxicabs could be figured out from the data of two major principal components during

2009~2015. The actual value of Neural Network is identical with the predictive value, namely, the actual demand could be estimated accurately according to the demand for taxicabs worked out with Neural Network. The time-space view is indicated from the different districts and different periods in one city, and the specific value of in specific periods and districts is correspondingly worked out with the help of population proportion, when it is 70%, that means balance between supply and demand. When it is within 60-75%, that means the problem of taking taxicabs is not severe.

References

[1] Yang Jianrong，the Analysis of Features of Taxi Industries in Domestic Cities and the Model Establishing of the Demand Prediction-Taking Nanchang City as An Example[D], a master dissertation of Nanchang University，2009，63.

[2]Information on https://www.haosou.com/s?q=%E5%9F%8E%E5%B8%82%E5%95%86%E4%B8%9A%E5%8C%BA%E6%AF%94%E4%BE%8B&src=srp&fr=lm&psid=0f46df90aeee96e53a9284ed4a5ffa03&ls=n36a7f6a197. 2015/9/12.

[3] Che lan，the Situation of the Taxi Development in Cities and the Demand Prediction [J], Shanxi Science and Technology，6，89-91，2006.

[4] Yang Changtao，the Research on the Inventory of Urban Taxicabs[D]，a master's dissertation of South China University of Technology,32-36，2011.