ATLANTIS PRESS
Advances in Engineering Research, volume 123
2nd International Conference on Materials Science, Machinery and Energy Engineering (MSMEE 2017)

The multiple regression model of the era "Internet +" taxi resources configuration research

Xiaocheng Gao[1,a],Yingbing Fan [2,b],Guoxian Wang[2,c],Ye Wang[2,d],Xichun Jiang[2,e],Yannan Mu*[2,f]

[1]Cpc committee office,Heihe University, 164300, Heihe City of Heilongjiang, China

[2]School of science, Heihe University, 164300, Heihe City of Heilongjiang, China

[a]270670891@qq.com, [b]276251052@qq.com,[c]251265647@qq.com,

[d]13949263@qq.com , [e]634281088@qq.com, [f]myn17@yeah.net

Keywords: The evaluation index, Regression analysis, Fuzzy comprehensive evaluation

Abstract. As taxi software increasingly hot, more and more people use a taxi booking taxi software. Based on the mobile Internet software has formed certain taxi with taxi process, showed a lot of convenient advantage, a taxi passengers every time use software taxi booking, used software company will give the driver and passengers of the corresponding subsidy, and with the upgrade of competition, the magnitude of the subsidies larger and larger. However, problems such as resource allocation and subsidies a taxi gradually leads to the difficult situation, first to collect relevant data, mathematical model is established. For problem 1: taxi "match" the supply and demand of resources, We set up three indicators, Ownership mileage effectiveness, load factor, ten thousand people through the establishment of a comprehensive evaluation model, evaluation and found only the load factor are greatly influenced by time and space, so the load factor has carried on the detailed analysis, finally come to the conclusion. For question 2,how to make a taxi software can better promotion in the society, the need for a reasonable subsidy policy with taxi. According to different total subsidy and taxi population proportion, taxi ownership of data, we established the multivariate regression model, then using SPSS software for factor into the model of multiple regression equation is obtained, thus significance analysis, the equation to satisfy different results subsidence a taxi and total population, taxi, the relationship between the ownership and subsidy scheme is given, and the error of the equation of subsidy policy accordingly reasonable analysis and sensitivity analysis and verification.

Introduction

Taxi is one of the important means of transport residents travel, "difficult to take a taxi" is a social hot spot of people. With the advent of the era of "Internet +", how a taxi companies relying on mobile Internet has established the software service platform, realize the communication between the passengers and the taxi driver, at the same time introduced a variety of taxi subsidy scheme. However, a taxi software update at the same time also appeared a lot of the problem of taking a taxi. Based on collecting relevant data, to establish mathematics model to study the following questions:

(1)Problems, according to take a taxi to take a taxi software service platform, we according to find reasonable index, analysis of different regions, time, analyses the effect of different populations of the distribution of the taxi, to see their "matching supply and demand".

(2)Asked to take a taxi to create a new optimized software service platform, to design a reasonable ease the difficult problem by subsidies, and analyze whether reasonable.

Model assumptions

(1)Assuming that the data collection and statistics, etc. There is a certain error; this paper assumes that all computing data in 5% 5% of the error range is acceptable;

(2)The assumption is not affected by environmental factors such as weather, seasons of large area;

Copyright © 2017, the Authors. Published by Atlantis Press.
This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

(3)Assume that subject only consider the city's downtown area and population more area;
(4)The hypothesis does not consider user passengers personal income, career, travel purpose, the other traffic usage of the city.

Establishment of the model and solution

Model of the fuzzy evaluation method. First of all, we selected the three evaluation indexes, hen remember mileage utilization rate y_1, load factor y_2, y_3, ten thousand respectively ownership. Mileage utilization standard value is55% ~ 60%, load factor of standard values for 70%, ten thousand ownership of no less than 20 (city); And their calculation formula:

$$\text{The mileage utilization rate} = \frac{\text{Passenger miles}}{\text{Mileage}} \times 100\% \quad (1)$$

$$\text{Load factor} = \frac{\text{Carrying}}{\text{The total number of through}} \times 100\% \quad (2)$$

$$\text{Ten thousand people ownership} = \frac{\text{Vehicle}}{\text{Population (ten thousand people)}} \quad (3)$$

Set evaluation objects for P (" demand match degree "). The evaluation index set. The evaluation index set the evaluation index set the evaluation index set. A judge object is P (“Demand match”). $U = \{y_1, y_2, y_3\}$.Judge set V={range，medium，favorable}.Judging matrix

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \text{（U,V,R）}$$

Constitute a fuzzy comprehensive evaluation model. A is the weighting matrix，

$$A = [a_1 \quad a_2 \quad a_3], \quad \sum_{i=1}^{3} a_i = 1; \quad \overline{B} = A \cdot R = (\overline{b_1}, \overline{b_2}, \cdots, \overline{b_m}), \quad B = \{b_1, b_2, \cdots, b_m\}.$$

So a judge can be set for the object P level.

To solve the model. Now in some areas, for example, by online data shows it is 66646 in downtown Beijing taxi ownership, and the main population of Beijing is 19.72 million, amount of private cars in Beijing is 63%, and travel at a rate of 34%, bus route 948, an average of 13.65 million passengers every day.

(1) First analyze the mileage utilization, Beijing taxi 2012 mileage data into the Equation. 1, get the following data:

table 1 Beijing taxi mileage data

	2012year	2013year	2014year	2015year（up）
Average monthly mileage	7480km	8650km	9968km	4921km
Average monthly passenger miles	5088km	5890km	6543km	3015km
The mileage utilization	68.02%	68.09%	65.64%	31.27%

(2)The load factor:

By the Equation. 2, it can be seen that load factor affected by the capacity and the taxi number, according to online data shows that load factor is bounded by 70%, less than 70%, then the area is saturated, the number of the taxi, on the other hand, more than 70%, should right amount send a taxi. But the Equation. 2 the results is not very accurate, because there is time, space, and the influence of such factors as different, so "the matching of supply and demand will be in order to achieve maximum

(3) Time factor：

table 2 Time factor data

	Go to work/school	after work/school is over
Students to and from school	8:30/14:00	11:50/17:40
Go to work or after work	7:00-8:30	16:30-18:00

By online data shows table 2. In the densely populated area, in this time nature also have a high level of the matching of supply and demand. So, the taxi passenger basic subject to the standard normal distribution probability density is $X \sim N(\mu, \sigma^2)$.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

x is time, σ、μ is constant, $\sigma = 0.45, \mu = (7.5, 11.83, 13.5, 17)$.

During the morning rush in, taxi passenger probability density is greater than 0.5, and the corresponding period, can enter the command in MATLAB and the results: the corresponding time period is 7 ~ 8.Similarly get in afternoon peak (school), the taxi passenger probability density is greater than 0.5, and the corresponding time period is 11:20 ~ 12:20.Peak in the afternoon (school), the taxi passenger probability density is greater than 0.5, and the corresponding time period is13:00 ~ 14:00. In the afternoon rush hour taxi passenger probability density is greater than 0.5, and the corresponding time period is16:30 ~ 17:30. Results: the above data can be concluded that "prime time" is7:00~8：00，11:20~12:20，13:00 ~ 14:00，16:30 ~ 17:30，In the above time period.Cab be used f will significantly increase .So $f = Ff_1(x)$.

(4)the space factors:

In different space, cab be used are also different. You can see from figure 1, in different locations, the demand of taxi is obviously different. To the demand of taking a taxi can be roughly in figure 1 will be Beijing cut into parts, each part is based on a certain point as the center, is radioactive divergence to the surrounding.

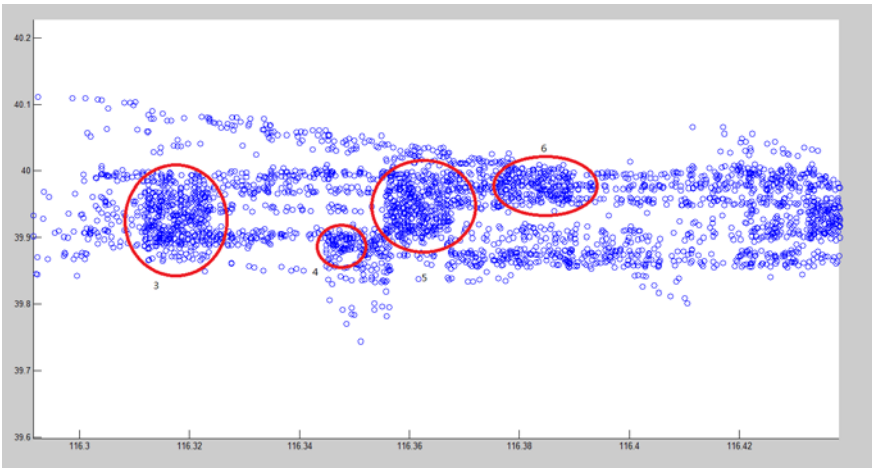

Fig.1 The scatter plot

Abscissa screen shot for longitude, ordinate for latitude. Fitting the scatter plot Figure 1.Screen shots can be seen from the above data, demand of taxi did comply with the principle of center divergence. So we can determine the capacity is also influenced by space, nearer the center distance, the greater the probability of a taxi passenger，d is the car rental and center distance, the impact index is $\frac{1}{d}$,

The taxi capacity

$$f = \frac{1}{d} Ff(x), \tag{4}$$

The Equation.4 into Equation.2，So

$$y_2 = \frac{\frac{1}{d} Ff(x)}{F} = \frac{1}{d} f(x) \tag{5}$$

Ten thousand people ownership： $y_3 = \frac{66646}{1972} = 33.79 \approx 34$

Results analysis：

$$R = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.3 & 0.3 & 0.4 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}, \quad A = \begin{bmatrix} 0.2 & 0.6 & 0.2 \end{bmatrix}$$

$\overline{B} = A \cdot R$ =[0.32 0.34 0.34],In the MATLAB command input so W=1.

Using normalized concluded w=1， The biggest influence that y_2 demands for "match", y_2 index of tiny change, will change "demand match" result. y_2 Standard of 70%, is the actual value of y_2 closer to 70%, it is concluded that the matching "demand", the better $\frac{1}{\mathrm{d}} f(x)$， 70% can make the optimal results.

Set Model. We use multiple linear regression analysis and processing data subsidies reward function is reasonable, first assumes that the total cost of the subsidy （p） proportion (a_1)， population by taking a taxi, a taxi ownership (a_2)， the influence of available models:

$$p = \beta_0 + \beta_1 a_1 + \beta_2 a_2 \tag{6}$$

table 3 Total cost of the subsidies in 2013-2015

Time	Total cost of the subsidy (yuan)	The taxi ownership (car)	Take a taxi population
January 2013	1870	896	80.16%
June 2013	3760	972	81.72%
July 2013	5478	989	84.35%
February 2014	8912	1050	85.82%
May 2014	10820	1360	86.57%
August 2014	18840	1820	87.46%
March 2015	21600	1972	87.65%
July 2015	28900	1990	87.89%
September 2015	37860	2016	88.32%

Solve the model. Use SPSS to find the solution for multiple regression and analysis:So according to the multiple regression analysis:

$$p = \beta_0 + \beta_1 a_1 + \beta_2 a_2$$

The regression equation is obtained:

$$\hat{y} = -36391.8 + 0.053 a_1 + 0.885 a_2$$

Predictive value of the standard deviation of $\hat{y}$ standard residual mean square estimation available:

$$s^2 = \sqrt{2.78} = 1.67 .$$

Final conclusion. Through the significance test of regression equation: from the table to know: 19.370 F statistics result data accord with the expression. Has a certain practical significance in the practical application, is quite reasonable. The design scheme is table.4.

table.4 The model results

Time	A taxi passenger cashback	The driver subsidies	Successful vehicle number	Number of issued a red envelope
2015/9/20	11	10	228	12
2015/9/30	5	3	786	5
2015/10/15	9	7	156	11
2015/10/29	10	8	78	15
2015/11/4	6	0	65	11
2015/11/28	5	3	82	13
2015/12/16	9	-3	345	6
2015/1/1	6	5	675	21

Model assessment and promotion

Model advantages. Using SPSS software to analyze the data given, found that the comprehensive evaluation model is a good way to play to the resource configuration and reasonable ways, and multiple regression analysis also obtained the regression equation of relation between the three variables. You can use to multiple variables in real life.

Model of faults. Collecting data is overmuch, cause data processing time is too long, so we only choose some data, analyzes the part of the different time points, cause analysis is not comprehensive, the result is not universal. We can put the model in this paper using commodity subsidies to the practical life, shopping subsidies, such as can also have a good effect.

Acknowledgments

This work was financially supported by the University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (No.2015114) and the Teaching Reform Research Project of Heihe University (xjg1502).

References

[1] Z.W. Zhang, J.N. Wang: Crane Design Manual (China Railway Press, China 1998), p.683-685. (In Chinese)

[2] Q.Y. Jiang,J.X Xie, and Y. J., Mathematical model. Beijing: Higher education press, 2012.(In Chinese)

[3] Z. Kai, J.Q. Song and X.J Wu. Introduction to mathematical modeling competition and the enhancement, Hangzhou: Zhejiang University press, 2012.(In Chinese)

[4] S.T.Lou, R.Y Yao and J.X. Shen. MATLAB programming language, Xi 'an: Xi 'an University of electronic science and technology press, 2012.(In Chinese)