Proceedings of the Eastern Asia Society for Transportation Studies, Vol. 5, pp. 1132 - 1143, 2005

A STUDY ON A SHOPPING FREQUENCY MODEL WITH THE CONSIDERATION OF INDIVIDUAL DIFFERENCE

Wei CHEN
Ph.D Student
Department of Urban and Environmental Engineering
Kyushu University
Hakozaki 6-10-1, Higashi-ku, Fukuoka, 812-8581 Japan
Fax: +81-92-642-3275
E-mail: lilychen@civil.doc.kyushu-u.ac.jp

Yoshinao OOEDA
Associate Professor
Department of Urban and Environmental Engineering
Kyushu University
Hakozaki 6-10-1, Higashi-ku, Fukuoka, 812-8581 Japan
Fax: +81-92-642-3275
E-mail: oeda@civil.doc.kyushu-u.ac.jp

Tomonori SUMI
Professor
Department of Urban and Environmental Engineering
Kyushu University
Hakozaki 6-10-1, Higashi-ku, Fukuoka, 812-8581 Japan
Fax: +81-92-642-3275
E-mail: sumi@civil.doc.kyushu-u.ac.jp

Abstract: On the basic assumption that people would like to maximize the utility by minimizing total cost and maximizing their benefit to meet their daily demand, this paper proposes a model to indicate the relationship between shopping frequency and travel distance for daily shopping. Considering the different discrimination ability and different attitude people have even for the same condition such as the same transportation mode and the same travel distance, the model specifies some distribution for one parameter to represent individual difference. The model is applied to the data obtained from a questionnaire that was held in September 2002. The result shows that the model well expresses the trend that the daily shopping frequency decreases with the distance increases.

Key Words: Travel distance, Daily shopping, Shopping frequency, Individual difference

1. INTRODUCTION

In previous studies on activity-based travel behavior, most attention was paid to commuting trip as the consequence that the commute periods always represent the most congested time of the day. In recent years, however, the contribution to the increasing travel demand, travel congestion, and air pollution of non-work trip induce more and more attention to non-work trip.

Shopping is one of people's important activities. Hereinto, Daily shopping is of importance for its frequency and necessity for daily life. Different from recreational travel, travel for daily shopping is a kind of compulsory activity accompanying with some pain. Also, travel for daily shopping is different from commuting trip, which has explicit temporal constrains. For travel behavior for daily shopping, it is relatively free for people to decide their departure time, stay time, back time, destination, travel mode, frequency etc. However, the characteristic of being relatively free of several dimension choices in daily shopping trip decision makes it of challenge for modeling daily shoppers' behavior.

Shopping frequency is an important aspect for understanding shopping behavior. Some scholars have done some related researches. They provide some insights for understanding shoppers' behavior by highlighting the relationship between shopping frequency and such factors as shoppers' demographic characteristics, the possibility to combine shopping trip with other trip, and shopping motivations (Kim B. and Park K., 1997; Bacon R.M., 1995; Jacobson J., 1982; Roy A., 1994; Bawa K., and Ghosh A., 1999). Shopping frequency, what's more, is an important aspect for understanding other shopping behavior. Based on shopping frequency analysis some study tried to forecast sales of retail establishment (Ishibashi K., Saito S., Kumata Y., 1998) and distinguish the characteristics of shopping center (Dejima T., 2004). Dejima's work implies some relationship between shopping frequency and destination choice. This relationship is apparently represented in Yoshida's work, in which the nested choice structure of shopping behavior is assumed to consist of three-dimensional decisions with frequency choice at top level, and then destination choice and parking choice (Yoshida A., and Harata N., 1990).

Shopping frequency may be influenced by many factors such as the household's socioeconomic characteristics, the free time available to shoppers and by the performance of travel mode and the built environment. It is impossible to conclude all factors in a model. Even if we can do so, the interdependence and interaction between those factors may make it is impossible to clarify how each specific factor affects the travelers' behavior (Sumi T. et al., 1990). This paper, thus, tries to grasp the most effective factor's influence on the shopping frequency. Based on the assumption that people seek to minimize the travel cost associated with shopping and the cost of holding goods in inventory and maximizing their benefit to meet their daily demand, this paper proposes a shopping frequency model to indicate the relationship between shopping frequency and travel distance for daily shopping. It is worthy to note that this basic assumption is almost the same to Bawa's assumption (Bawa and Ghosh 1999). The difference is that Bawa studied the relationship between shopping frequency choice and the household characteristics that may influence travel cost, inventory cost, and consumption need, while this paper quantifies the travel cost and storage (similar to inventory) cost to catch the mechanism of frequency choice behavior. What's more, the different relationships for different transportation mode are studied in the paper. In addition, the individual difference is added into the model by presenting the parameter of the distance variable A' in a lognormal distribution to represent the different discrimination ability and different attitude people have even for the same condition such as distance. The relation between frequency and distance is represented clearly by quantifying. The paper will contribute to the understanding of the physical environment's influence on shopping behaviors.

The rest of this paper is organized as follows. Section 2 develops the structure and presents the estimation procedure for the frequency model for daily shopping. Section 3 discusses the data and the empirical results. The final section provides a summary of the research.

2. MODEL STRUCTURE AND ESTIMATION

2.1 Assumptions

Daily shopping means the shopping related to daily life and usually happened in the shops nearby the residence areas. We assume that the most goods are grocery including fresh foods, dry foods, cosmetic, medicine etc. For daily shopping, it is reasonable to expect that the need for a person in a certain period is constant. On the basis of this, this model assumes that for a certain person total shopping associated cost, produced by all cost a person spends related to shopping subtracting the money a person spends in purchasing goods directly, includes travel cost and storage cost.

Hereinto, travel cost includes three parts. One is the cost influenced by transportation mode, parking etc. One is the cost spent on the travel from the residence to the shop. The third one is the cost spent on the travel from the shop to the residence while the shopper must take the goods with him or her. The travel distance and the amount of goods, which is measured by the weight of goods that can be measured by spent money, influence the later two parts. Storage cost means the cost spent for storing goods. It includes the cost caused by the quality reduction of goods, especially fresh foods, refrigeration and storage space etc. Storage cost has two parts, one is the cost independent of the frequency, and the other is the part that is dependent on the frequency. On the basic principle that the shoppers are willing to spend the least total shopping associated cost, it is possible to get the relationship between frequency and distance.

Furthermore, considering the different attitudes and distinguish ability of different individual person should holds for even the same situation, there should be a parameter representing the individual difference’s influence on shopping frequency choice.

2.2 Modeling

The model is shown as below:

$$C = C_t + C_s \tag{1}$$

Where,
C: Total cost for daily shopping during a certain period (for example: one day);
C_t: Travel cost during a certain period (for example: one day);
C_s: Storage cost during a certain period (for example: one day);

$$C_t = (c \times Y \times L + g \times L + c_0) / d \tag{2}$$

Where,
c, g: Non-negative parameter;
Y: Money spent each shopping;
$d = 1/n$
d: Interval;
n: Frequency;
L: Distance
c_0: Other travel cost, for example parking, gas etc; it is a constant for a certain transportation mode and certain O-D;
$g \times L$: Represents the travel cost spent on the travel from the residence to the shop (includes time consumption);

$c \times Y \times L$: Represents the travel cost spent on the travel from the shop to the residence while the shopper must take the goods with him or her, and Y represents the influence of the goods' weight on the travel cost;
$c \times Y \times L + g \times L + c_0$: Represents travel cost for one shopping;

$$Y = a + b \times d^{\beta} \tag{3}$$

Where,
β : Parameter;
a: The part of spent money that is independent of d ; the part of money that the shopper will spend once he or she goes shopping;
b: The part of spent money which is dependent of d ; the part of money that the shopper spends according to the frequency and the need; and the need is assumed to be constant for a certain person in a certain period;

$$C_s = (e + f \times d^{\alpha}) / d \tag{4}$$

Where,
e : The part of storage cost which is independent of d;
f : The part of storage cost which is dependent on d;
α : Parameter;

Equation (1) is rewritten by using equation (2), (3) and (4), then

$$C = (c \times a \times L + g \times L + c_0 + e) / d + c \times b \times L \times d^{\beta - 1} + f \times d^{\alpha - 1} \tag{5}$$

When $\beta = 1$

$$C = (c \times a \times L + g \times L + c_0 + e) \times d^{-1} + f \times d^{\alpha - 1} + c \times b \times L \tag{6}$$

Let $dC / dd = 0$, then

$$\begin{aligned} d &= \sqrt[\alpha]{\frac{c \times a \times L + g \times L + c_0 + e}{(\alpha - 1) \times f}} \\ &= \sqrt[\alpha]{\frac{A \times L + B}{\alpha - 1}} \end{aligned} \tag{7}$$

Hereinto,

$$A = (c \times a + g) / f \tag{8}$$
$$B = (c_0 + e) / f \tag{9}$$

The equation (7) can be changed to below equation:

$$(\alpha - 1) d^{\alpha} = A \times L + B \tag{10}$$

Where, d, L can be got from the questionnaire.

To eliminate the influence among parameters the equation (10) is changed to equation (11) as below:

$$d^{\alpha} = A' \times L + B' \tag{11}$$

Where,

$$A' = \frac{A}{\alpha - 1} \tag{12}$$

$$B' = \frac{B}{\alpha - 1} \tag{13}$$

Here, denote parameter A' as the individual difference, whose distribution is assumed to be some traditional classical distribution.

2.3 Estimation

Linear regression parameters estimation method and Least Squares estimation method are used to estimate the values of parameters. Different from the previous paper (see Chen et al. (2004)), this paper uses the linear regression method to get the values of the parameters A' and B' by replacing d^{α} by d' with any possible value of α. (α can be evaluated as any real number except 0 and 1) A''s value is limited to meet the significance of linear regression. For bicycle data, the significance level is 90% as well as the significance level 95% for other data groups. The values of α, A', and B' are fixed by least square method to ensure the least difference between the observed interval and the calculated interval, which is represented as below equation:

$$R = \sum_{i} (M_{oi} - M_{ci})^2 \tag{14}$$

Where,
M_{oi} : Observed interval of the ith O-D group;
M_{ci} : Calculated interval of the ith O-D group;

After the estimation of parameters, the values are used to estimate parameter A''s distribution. One classical distribution is assumed to be parameter A''s distribution. Then the probability distribution of calculated interval could be calculated based on A''s distribution as below equation.

$$\phi_d(d) = \phi_{A'}(A') \frac{dA'}{dd} \tag{15}$$

Where,
$\phi_d(d)$: The probability density function of the probability distribution of the calculated interval;
$\phi_{A'}(A')$: The probability density function of the probability distribution of the parameter A'.

The distribution parameters for the classical distribution of the parameter A' are fixed on the principle of getting the highest correlation between the possibility distribution of the calculated interval and the observed interval. If the correlation is high enough, the assumed distribution of the parameter A' should be accepted.

3. APPLICATION

3.1 Data

The data used in the study are drawn from the survey held in the suburb of Fukuoka in 2002. The survey was held to help researchers understand the shoppers' behavior. The contents of the survey include the attributes of the answers, for example, residence zone, age, family, gender, license ownership, available transportation mode for shopping, and traffic tickets ownership, and the data related to daily shopping, for example, the shopping zone, transportation mode, travel time, stay time, frequency, money spent for each time, and the data related to non-daily shopping. The research is done on the basic of analyzing the attribute data and the daily shopping data that got from the survey. The feedback rate of the survey is 25.1% with total distribution number 1932 and return number 485. Hereinto, 327 samples are qualified for daily shopping analysis so that they were used in this study.

The distance alternatives include 16 values obtained from the distance of the permutation and combination between the residence zone set and the shopping zone set. Ideally, the distance should be determined according to the detail residential and shopping address and the shopper's shopping travel route. However, it is impossible to get the accurate shopping travel distances of the answerers because of the lacking of those detail data. Therefore, this study use the average distance, which is the distance from the center of the residence zone to the center of the shopping area, as components of the distance alternatives set. They are computed by multiplying map distance by a parameter such as 1.4. The values of average distances of each O-D groups are shown as Table1.

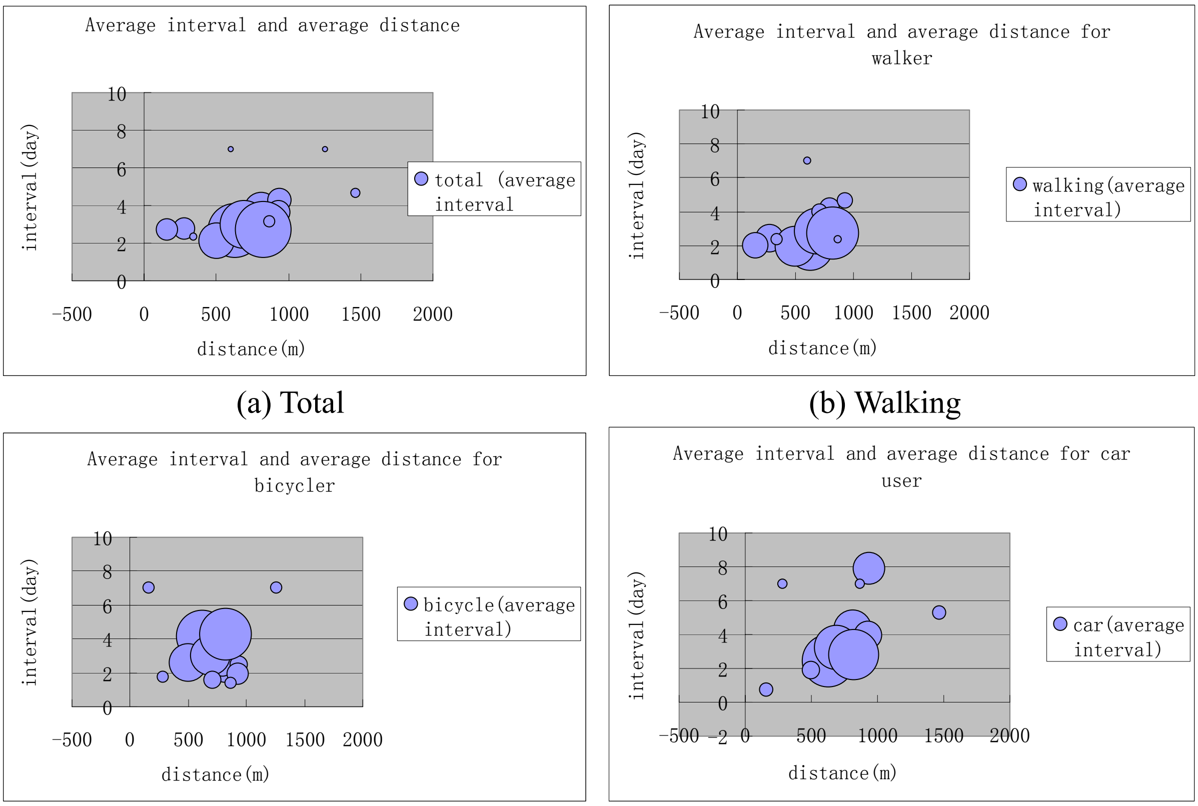
Accordingly, this study uses average frequency of each O-D group to get the value of average interval. The sample used in this paper comprises 327 household-based person-shopping records obtained from the overall travel survey sample. The mode choice shares in the sample are as follows: walking (43.1%), bicycle (19.6%), car (32.1%), other (5.2%). The distance distribution in the sample is as follows: 158m (3.7%), 283m (3.4%), 341m (0.6%), 503m (9.2%), 606m (0.3%), 629m (20.5%), 694m (16.2%), 709m (1.8%), 802m (2.4%), 825m (22.6%), 868m (1.2%), 932m (4.0%), 938m (4.6%), 1256m (0.3%), 1468m (0.9%). The interval distribution in the sample is as follow: 1 day (4.6%), 1 to 2 days (37.3%), 2 to 4 days (42.8%), 4 to 8 days (13.5%), more than 8 days (1.8%).

Table1: Average Distance and Symbols of O-D Groups:

Shopping area / Residence	1		2		3		4	
	D (m)	Symbol	D (m)	Symbol	D (m)	Symbol	D (m)	Symbol
1	283	1	628	2	813	3	802	4
2	158	5	938	6	932	7	503	8
3	341	9	694	10	1256	11	709	12
4	606	13	825	14	1468	15	868	16

(Note: D means Distance)

Figure 1 shows the relationship between travel distance and daily shopping interval (the size of the circle represents the sample's number). The data were also divided according to the three main transportation mode, walking, bicycle and car. The observed relationship between the travel distance and the interval of each of them is shown in Figure1 (b), (c), (d) respectively.



(a) Total
(b) Walking
(c) Bicycle
(d) Car

Figure1 Observed Relationship between Travel Distance and Shopping Interval

3.2 Model variables and empirical results

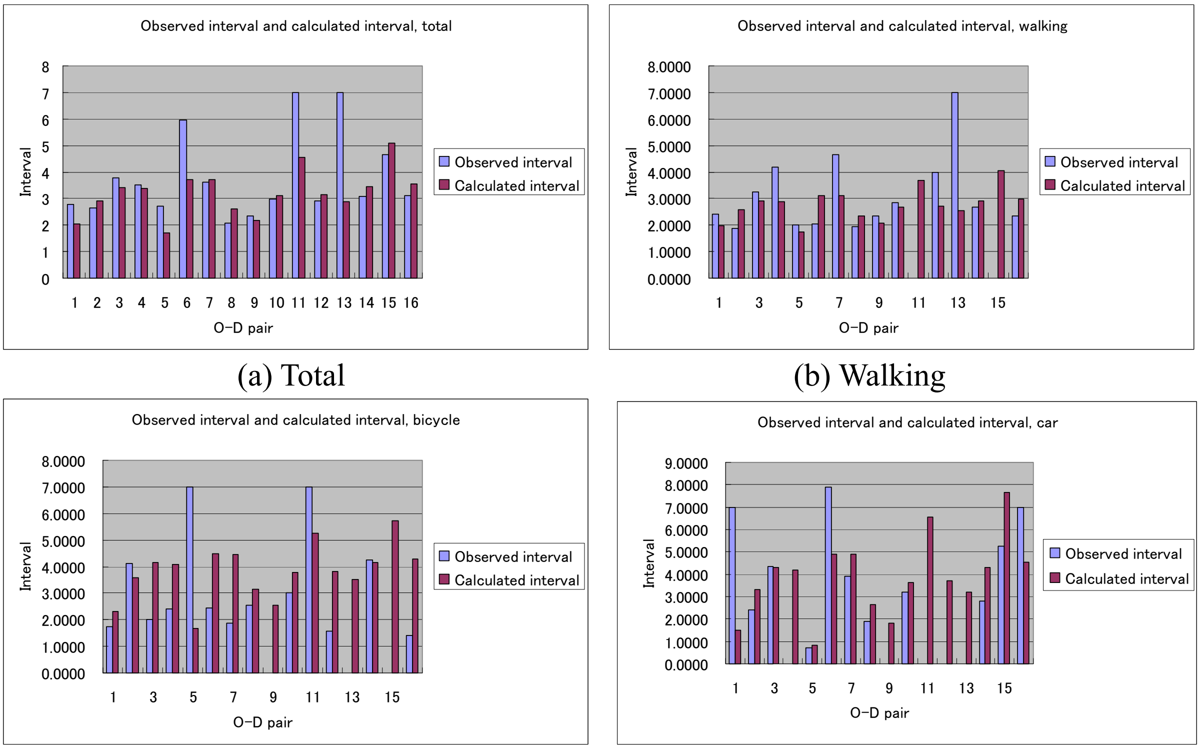
There are totally three parameters in this model with one of them has been assumed to belong to some classical distribution. As discussed in estimation part, the linear regression parameter estimation method is used in this procedure. The results of parameters calculation are shown in Table2, and the result of calculated interval that got by applying the parameter into the model is shown in Figure2. The relation among the symbols of O-D groups, shopping areas and residences is shown in Table1.

Table2 Parameters' Calculation Results

	α	A'	B'	M (Least square)	Sample size
Total	1.01	0.002649	1.2878	144.6784	327
Walking	1.01	0.001783	1.4754	63.3836	141
Bicycle	1.80	0.015783	0.0001	112.8275	64
Car	1.01	0.005316	0.0001	223.2242	105

The correlation of the observed interval and the calculated interval is calculated with the result turns to be 0.60. The mean and the standard deviation of the observed interval are 3.11

and 0.84 respectively. In addition, the mean and the standard deviation of the calculated interval turn to be 3.11 and 0.51 respectively.



(a) Total (b) Walking

(c) Bicycle (d) Car

Figure2 Observed Interval and Calculated Interval

3.3 Individual difference

It shows that the statistical result is not very good. But it is understandable because here the average travel distance is used and the individual difference of shopping behavior is neglected. But there should be some variances existing in not only the travel distance of the people who have the same O-D choice but also in people's distinguish ability and attitude to even the same accurate travel distance. Those differences will induce people to have different reaction to same conditions, for example, the same travel distance and same transport mode.

For those reasons, this paper assumes that even when the other variables such as the travel mode, the O-D choice are the same still different person perceives the certain distance in different way. Parameter A', the constraints coefficient of the travel distance L, is selected to represent the individual difference among samples. Finally the lognormal distribution is found to fix the data well in this study.

$$f(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma x}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, & x>0\\ 0, & others\end{cases} \tag{16}$$

The parameters of the lognormal distribution for total data and each mode data are got and shown in Table3.

The χ^2 test was used to test the reasonability of above distribution. The P-values of the distribution of walking and car data turn to be 0.0002 and 0.001. The distribution of the total data doesn't pass the test. As to bicycle data, being lacking of data the χ^2 test is infeasible.

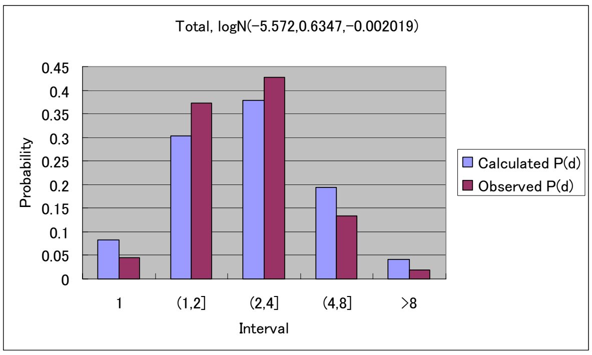
Table3 Parameters of A''s Lognormal Distribution

	Location	Scale	Threshold
Total	-5.572	0.6347	-0.002019
Walking	-5.632	0.5670	-0.002375
Bicycle	-5.212	1.563	0.001005
Car	-6.106	0.9963	0.001295

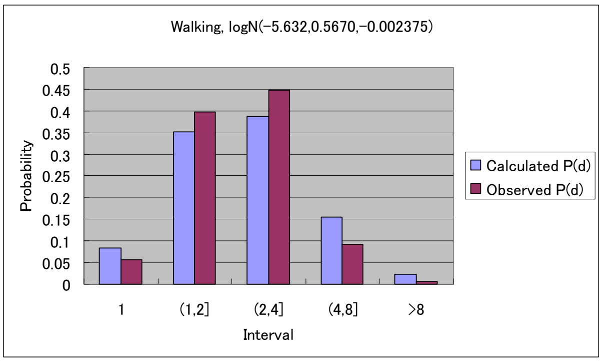
3.4 Empirical results with the consideration of individual difference

Table4 Correlation between Calculated Probability Distribution and Observed Distribution of Interval

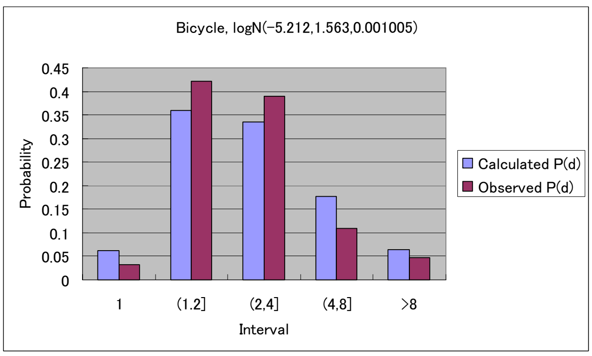
		Mean	Standard deviation	Correlation
	Observed possibility	0.2	0.19	
Total	Calculated possibility	0.2	0.14	0.98
	Observed possibility	0.2	0.21	
Walking	Calculated possibility	0.2	0.16	0.99
	Observed possibility	0.2	0.19	
Bicycle	Calculated possibility	0.2	0.14	0.98
	Observed possibility	0.2	0.17	
Car	Calculated possibility	0.2	0.14	0.98



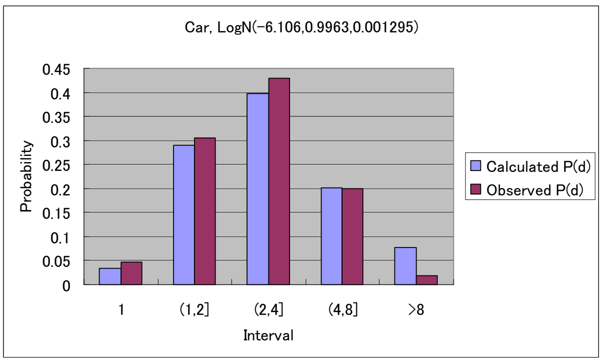
(a) Total



(b) Walking



(c) Bicycle



(d) Car

Figure3 Observed probability and calculated probability of interval with the consideration of individual difference

With the consideration of individual difference, the lognormal distribution of A' is introduced into the model. The possibility distributions of intervals were recalculated with the results turn to be good as shown in Figure3 and Table4. The correlation between calculated probability and observed probability turns to be well acceptable. It shows that individual difference explains most of the marginal influence of some unobserved variables or taste variables.

It shows the model with the consideration of individual difference provides a satisfying measure to grasp the relationship between the shopping frequency and the travel distance. If we know the travel distance, it is possible to estimate the shopping frequency and vice versa.

4. CONCLUSIONS

Shopping frequency analysis is an important component of the analysis of shoppers' behavior. It not only provides insight to the frequency choice, but also provides significant foundation of other shopping related choices.

Though there should be some random elements will influence people's travel behavior in daily shopping, it is still reasonable to expect that people would like to maximize the utility by minimizing total cost and maximizing their benefit to meet their daily demand. It is obvious that the longer the distance, the higher the person's travel cost. On the other hand, if the shopping frequency increases, storage cost will decrease and travel cost will increase. Usually, people tend to do daily shopping in some shop repeatedly so the travel distance is mostly constant for a person. Considering the influence on the total cost of travel distance and shopping frequency, it is assumed that to some extent travel distance should influence shopping frequency.

In this paper, we formulated and estimated an original model for the daily shopping frequency choice to indicate the relationship between shopping frequency and travel distance for daily shopping with the consideration of individual influence. The model is an original one with a simplified form. Linear regression estimation method is used to evaluate the parameters with the principle of the least square estimation. The estimation method is verified to be straightforward and efficient.

The empirical analysis of the paper applied the frequency model to the estimation of frequency choice in a suburb using data drawn from a survey held in a Fukuoka suburb in September of 2002. It is found that the lognormal distribution fit the data well while representing the individual difference by using the lognormal distribution to describe the distribution of the coefficient of the travel distance. With observed characteristics of shoppers' behavior shows that the shoppers would like to reduce their daily shopping frequency with the travel distance increases, the result of this paper shows that the model well expresses this tendency. Also the estimation and statistical results turn to be satisfying. The significant difference of the parameters of the lognormal distribution for different mode shows that the individual person's attitude to the travel distance varies a lot among different modes.

Frequency is an essential factors influence travel demand. It is obvious that the more frequency, the more traffic flow. Considering the contribution of car driving to the transportation problems such as congestion, pollution, accidence, urban planners and managers are trying to reduce the travel demand of car driving. As to shopping travel, we also need to consider the measures to control the travel demand of driving and induce the travel of walking or bicycling. For the same distance, the feeling of travelers who take different transportation mode will be different. The perception may be influenced by the road condition, parking condition etc. If urban planners and managers try to provide amiable travel environment for bicycler and walker, it is possible to induce the impedance of distance for them and thus induce the travel by bicycling and walking. On the other hand, if we give less consideration or even set barriers for drivers it is possible to increase the impedance of distance for drivers and thus reduce the travel by driving. Therefore, it is important to understand the interaction between distance and frequency for different transportation mode.

For the future research, the author will attempt to investigate the relation among shopping frequency, transportation mode choice and destination choice by distinguishing the effects of the physical influencing. This study will provide reference for future research.

REFERENCES

a) Books and Books chapters

David S.M. and George P. M. (2002) **Introduction to the Practice of Statistics**. W.H. Freeman and Company, New York.

b) Journal papers

Bacon, R. W. (1995) Combined trips and the frequency of shopping, **Journal of Retailing and Consumer Services, Vol. 2, No. 3**, 175-183.

Bhat, C. R., Frusti T., Zhao, H., Schonfelder, S., and Axhausen, K. W (2004) Intershopping duration: an analysis using multiweek data, **Transportation Research Part B 38**, 39-60.

Ishibashi, K., Saito, A., and Kumata, Y (1998) A disaggregate Markov shop-around model to forecast sales of retail establishments based on the frequency of shoppers' visits—its application to city center retail environment at Kitakyushu City, **Journal of the City Planning Institute of Japan, No.33**, 349-354 (in Japanese).

Jacobson, J. (1982) Alternative specifications of the random disturbances for trip frequency models, **Transportation Research-A Vol.16A, No.3**, 219-223.

Kim, B., and Park, K. (1997) Studying patterns of consumer's grocery shopping trip, **Journal of Retailing, Vol. 73(4)**, 501-517.

Luoma, M., and Mikkonen, K., and Palomaki, M. (1993) The threshold gravity model and transport geography---How transport development influences the distance-decay parameter of the gravity model, **Journal of Transport Geography, Vol.1, No. 4**, 240-247.

Roy, A. (1994) Correlates of mall visit frequency, **Journal of Retailing, Vol. 70, No. 2**, 139-161.

Sumi, T. (1990) Departure time and route choice of commuters on mass transit systems, **Transpn. Res., Vol. 24B, No.4**,247-262.

c) Papers presented to conferences

Chen, W., Dejima, T., Ooeda, Y., and Sumi, T. (2004) A study on the model of travel behavior for daily shopping, **Proceedings of International Symposium on City Planning**, 31-40.

Yoshida, A., and Harata, N (1990) Frequency choice model for non-grocery shopping trips, as an aspect of multidimensional travel behavior. **Proceedings of JSCE** No 413, 107-116 (in

Japanese).

d) Other documents

Dejima, T. (2004) Analysis of shopping behavior from the view about frequency --- A example of a suburb residence zone of Fukuoka, **The Annual report of Japan Society for Distributive Sciences** (in Japanese).