TOPIC 9
ADVANCED TRAVELLER
INFORMATION SYSTEMS

A STUDY ON ANALYZING METHODS OF INFORMATION EFFECTS TO TRAFFIC BEHAVIOR

SHIGERU MORICHI
Tokyo Institute of Technology
Tokyo, JAPAN

KOICHIRO MEGURO
Mitsubishi Research Institute
Tokyo, JAPAN

KEIICHI OGAWA
Tokyo Institute of Technology
Tokyo, JAPAN

HIDEKATSU HAMAOKA
Tokyo Institute of Technology
Tokyo, JAPAN

Abstract

In this age of information technology, many traffic information systems such as travel time information and parking guidance had become readily available. These systems influence the travel behaviour of the drivers of vehicles on the roads and eventually the traffic flow of the road network. The usual models could not incorporate the mentioned effects such that there is a need to develop the methodology to take into account the influence of information.

INTRODUCTION

In the advent of information technology, many traffic information systems such as travel time information and parking information for road traffic had become readily available not only through radio broadcasts, but also in street on-board systems and vehicle navigation systems using satellites. These information systems influence the travel behavior of the drivers of vehicles on the roads and eventually the traffic flow of the road network. However, difficulties were encountered by researchers because the usual models could not incorporate the mentioned effects. Given this situation, there is a need to develop the methodology to take into account the influences of information.

When information is provided to the individual, the structure of decision-making is modified to take into account this factor. In the decision-making process of choosing an alternative, the concept of utility function is used. The usual method of calculating utility is the weighted sum of the attributes which includes travel time, cost and many other quantities. However, individual choice is not that simple wherein he may on choose on the basis of the most significant attribute with the maximum value or with the minimum value. Decision structure may also be based on the mean. These complexities associated with decision-making are conveniently represented by the concept of generalized mean which is used in this study. The existence of various types of decision-making structure illustrates the need for grouping of individuals into segments.

THE DIFFICULTIES IN THE APPLICATION OF THE ORDINARY BEHAVIORAL MODELS

In the ordinary traffic behavioral models, an individual should have perfect information about the choice set. However, in the case of travel behavior analysis with given traffic information, a trip-maker need not obtain the perfect information for each of the alternatives. A model which can consider the uncertainty about the choice set should be made in this area of the study.

In the travel behavior analysis aimed in incorporating and at the same time estimating the effect of information systems, there are two assumptions which must be relaxed. These are: (1) the systematic part of the utility function is the weighted sum of the attributes, and (2) all individuals have the same way of choosing alternatives. In the following, the problem to adopt the ordinary utility function to the travel behavior analysis under new information is discussed.

The systematic part of utility function is the weighted sum of the attributes

When an individual is given several alternatives, an option may be selected on the basis of the maximum or minimum value of the significant attribute according to self-assignment of importance of certain attributes.

Another scheme would be choosing the alternative based on the mean of the attributes. The most common used expression for the systematic part of utility function which is the weighted sum may not be enough to accommodate the above-stated complexities in individual decision-making. The concept of generalized mean conveniently accommodates the mentioned complexities in decision-making. This concept is explained fully in the following section.

All individuals have the same structure of decision-making

In the classic travel behavior model, individuals must have the same way of choosing alternatives. In the road network, the traffic flows in the roads are not essentially in equilibrium because there are already fixed preferences in choosing a certain road even though that road is already too congested. There is a need to conduct segmentation that shows the preferences to certain roads or specific variables which consists the utility function.

From the discussion of the difficulty of using the ordinary behavioral models in this research, the aim of this study is now summarized. The main objective of this study is to include in the individual travel behavior the individual differences in decision-making towards choosing the alternatives. In the following section, the generalized mean is used to overcome this difficulty.

THE ALTERNATIVE MODEL TO CONSIDER UNCERTAINTY

The concept of generalized mean

There are many kind of calculations to get the representative. The main concept to calculate the representative is shown in Equation 1.

$$h: [x_1, x_2, \cdots x_n] \to x \tag{1}$$

The calculation is subject to satisfy the condition,

$$\min(x_1, x_2, \cdots x_n) \le h\,(x_1, x_2, \cdots x_n) \le \max(x_1, x_2, \cdots x_n)$$

is called as a kind of mean. There is a concept of generalized mean (Equation 2) that summarizes the concept mentioned earlier:

$$x = \left\{\frac{x_1^{\alpha} + x_2^{\alpha} + \cdots + x_n^{\alpha}}{n}\right\}^{\frac{1}{\alpha}} \tag{2}$$

The α ($\alpha \neq 0$) is the parameter to distinguish the representatives such as the maximum, the minimum and some kinds of means. The representatives are shown below as α changes, and the concept of α is shown in Figure 1.

- maximum ($\alpha \to +\infty$) $$x = \max\{x_1, x_2, \cdots, x_n\} \tag{3}$$
- minimum ($\alpha \to -\infty$) $$x = \min\{x_1, x_2, \cdots, x_n\} \tag{4}$$
- arithmetic mean ($\alpha=1$) $$x = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{5}$$
- geometric mean ($\alpha=0$) $$x = \{x_1 \cdot x_2 \cdots x_n\}^{\frac{1}{n}} \tag{6}$$
- harmonic mean ($\alpha=-1$) $$x = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}} \tag{7}$$

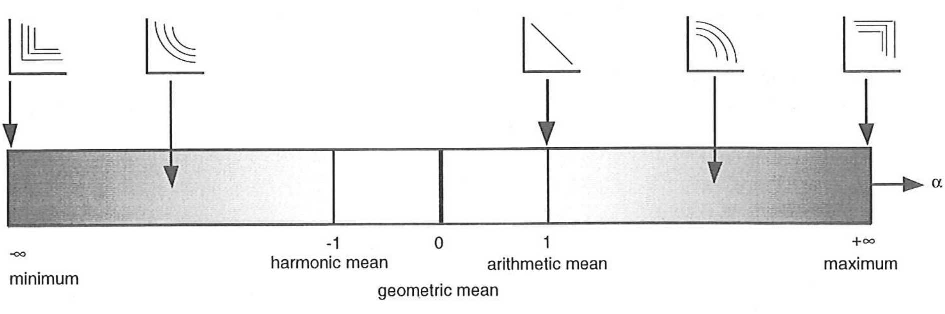

Figure 1 The indifference curves of generalized means

The parameter α shows the substitution for each variables. If there are two variables, x_1 and x_2 which make up the utility function, the indifference curve is shown in Figure 1. Below, k is specified to be:

$$k = \frac{1}{\alpha} \tag{8}$$

where k indicates the ratio of substitution. The meaning of parameter k=0 ($\alpha=\pm\infty$) shows the relationship between the variables with regard to the substitution. The parameter k=1 (α=1) shows the perfect substitution, that is when x_1 increases 1 unit, x_2 decreases by 1 unit. This function is basically the constant elasticity of substitution (CES).

Another idea is to include the weight w_i into the generalized mean, shown below.

$$x = \left\{ w_1 x_1^{\alpha} + w_2 x_2^{\alpha} + \cdots + w_n x_n^{\alpha} \right\}^{\frac{1}{\alpha}}$$
$$\sum_{i=1}^{n} w_i = 1 \tag{9}$$

These ideas of the generalized means are discussed by Dyckhoff and Pedrycz (1984), and they have been used as ideas of aggregation operations of fuzzy sets which cover the range of averaging operations. General description about fuzzy aggregation operations is done by Klir and Folger (1988).

In this paper, by utilizing the equation of generalized mean and weighted generalized mean, the travel time estimation model of homecoming trips and the destination choice model for shopping trips are estimated. Formulation of these models will be discussed in the following sections.

Formulation of the model for estimating the travel time for homecoming trips

The travel times of homecoming trips are not equivalent because of the heavy traffic congestion during those seasons. In those periods, the day and the time of departure vary from individual to individual based on the traffic information of road congestion. This study assumes that the travel time is estimated by the equation below (Equation 10).

$$T = \left\{ \frac{t_1^{\alpha} + t_2^{\alpha} + \cdots + t_n^{\alpha}}{n} \right\}^{\frac{1}{\alpha}} \tag{10}$$

In the above model, the expected travel time is estimated by the generalized mean of the time that was experienced in the past. Using the parameter α, the characteristics of the respondents could be determined. If the α is positive and large, the expected travel time is influenced by the longest time. In the risk theory, it is said as "risk avoidance". Similarly, a negative and small α indicates "risk preference".

Furthermore, the weighted generalized mean can be applied estimate the expected travel time (Equation 11).

$$T = \left\{ w_1 t_1^{\alpha} + w_2 t_2^{\alpha} + \cdots + w_n t_n^{\alpha} \right\}^{\frac{1}{\alpha}} \tag{11}$$

The parameter w_i is defined as the weight which places relative importance to experienced travel times in the past which are components of expected travel time. Using this weight, the decreasing of the weight of the experienced travel time in longer time ago will be considered. Also, the data errors such as time delay due to traffic accident can be eliminated.

Formulation of the destination choice model for shopping trips

Formulation of the model

There are many factors which influence the choice of shopping place. Here, it is considered that the choice is made based on the factors such as the travel time to the shopping place, the waiting time to park the car and its attraction. The attraction of the shopping location is defined in terms of the price level of products sold and the product variety. In this paper, the destination choice models are formulated using the concept of weighted generalized mean in terms of the utility function incorporated in the discrete choice analysis (Equation 12).

$$P_i = \frac{e^{V_i}}{\sum_j e^{V_j}}$$

$$V_i = \left\{w_1 x_1^{\alpha} + w_2 x_2^{\alpha} + \cdots + w_n x_n^{\alpha}\right\}^{\frac{1}{\alpha}} \tag{12}$$

It is possible to compare both to see the relative magnitude in selecting the alternatives, because the variables x_i are standardized to have the same mean and variance. The parameter α indicates the way of choosing or decision-making method of an individual, that is to choose the maximum or minimum of the variables. Also, the α shows the effectiveness of the information in the decision-making of the individual. If an individual has a large and positive α, and when variables are increased by the new information, the total utility is also increased. If an individual has a small and negative α, and when variables are decreased by the new information, the total utility is also decreased. Furthermore, another variation of utility function which uses the fuzzy integral formula can be considered. Ordinary fuzzy integral formula is shown below (Equation 13).

$$x = \sup_{F' \subset F}\left[\inf_{x \in F'} h(x_k) \wedge g(F')\right] \tag{13}$$

where h(x) is the evaluation value of each variable and g(F') is a weight parameter of the set of variables F'. These values are standardized in the range [0,1] and g(F') has to satisfy the condition below.

$$g(F') = \frac{1}{\lambda}\left[\prod_{x_k \in F'} (1+\lambda g(x_k)) - 1\right]$$

$$g(F) = 1 \tag{14}$$

The new utility function is the combination of the fuzzy integral and the generalized mean. In this function, the maximization part of the fuzzy integral is changed to the generalized mean. Formulation of the model is shown in Equation 15.

$$P_i = \frac{e^{V_i}}{\sum_j e^{V_j}}$$

$$V_i = \left\{\frac{\sum_{F' \subset F}\left[\inf_{x_k \in F'} h(x_k) \wedge g(F')\right]^{\alpha}}{\#\text{ of } F'}\right\}^{\frac{1}{\alpha}} \tag{15}$$

In this utility function, the structure of decision-making of the individual is changed from the weighted sum to the fuzzy integral, which means the relaxation of the additive property of utility function. The parameter α indicates the decision-making method and the effectiveness of the information in decision-making same as the case of weighted generalized mean model.

Segmentation of individuals

In these new utility functions (Equation 12, 15), as discussed above, the parameter α is considered to represent the individual's decision-making method and effectiveness of the information. But it is considered that this evaluation of each individual is different due to the individuals. So in this paper, to consider the difference in the way of choosing the alternatives among individuals, the groupings are made using the parameter w_i and α within the utility function.

The method of segmentation is the iteration of convergence of segmentation group and estimation of parameter of each group, which is proposed by Katahira (1987). In the first step, the number of segmentation groups and the initial values of parameters of each group are assumed. In the second step, likelihood of each individual using initial parameters of each group is calculated, and each individual belongs to the group where the likelihood used its parameters is maximum. In the third step, the parameters of each group are estimated by the ordinary maximum likelihood method. After the third step, estimated parameters are replaced with initial parameters of each group, and the second step and the third step are iterated until convergence. The procedure of segmentation is shown in Figure 2.

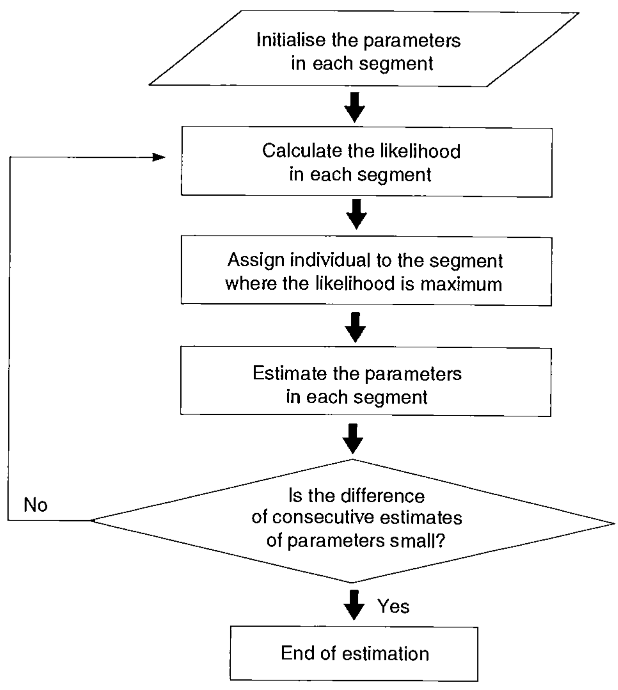

Figure 2 Flow of the method of segmentation

APPLICATION TO SURVEY DATA

The survey data

The data were collected at a survey conducted at Yokohama City in 1993. There are two kinds of data, that is, the home-based survey in the residential area, and the survey at parking areas in the commercial district. The number of data collected are 207 in former and 209 in the latter.

The survey questionnaire is divided into two parts where the first part inquires about the expected travel time of the respondent's homecoming trip and the second part asks the choice of shopping place. In the part of the homecoming trip, the respondents are asked about their usual mode and the experienced travel times of that mode for the last 5 times. And also asked is the expected travel time when the travel time information is obtained. In the part of the shopping trip, the respondents are asked to name places they go frequently by car, indicate the travel time to the shopping place, indicate the waiting time for parking, and the attraction for the shopping place. Also asked if whether the travel time or parking information is shown at the time of the trip.

Based on these data, expected travel time model for homecoming trips and the destination choice model for shopping trips are estimated using the generalized mean concept.

Application to the estimation of travel time for homecoming trips

In this section, the expected travel time is estimated by the generalized mean, the weighted generalized mean and also by the linear estimation. The models shown below are estimated and the corresponding parameters and performance are compared:

$$T = \sum_{i=1}^{5} \theta_i t_i + \theta_{ord} t_{ord} + \theta_{max} t_{max} + \text{const.} \tag{16}$$

$$T = A.\left(\frac{\sum_{i=1}^{5} t_j^{\alpha} + t_{ord}^{\alpha} + t_{max}^{\alpha}}{7} \right)^{\frac{1}{\alpha}} + \text{const.} \tag{17}$$

$$T = A.\left\{ \sum_{i=1}^{5} w_i t_j^{\alpha} + w_{ord} t_{ord}^{\alpha} + w_{max} t_{max}^{\alpha} \right\}^{\frac{1}{\alpha}} + \text{const.} \tag{18}$$

The results of these models are compared in Table 1. This table shows that the result of generalized mean model has a lower R-squared value than the linear model because of the smaller number of parameters, but the weighted generalized mean model has a larger R-squared value which indicates a better solution. The parameter α for each model is estimated close to 0 or 1 which means that the individual's expected travel time is close to the average travel time in the past. And comparing the parameters w_i, the parameter for the ordinary time is the largest. This result means that their expectation of travel time is neither "risk avoidance" nor "risk preference" in the homecoming trip.

Table 1 Results of estimated parameters of linear, generalized mean and weighted generalized mean model

	Linear model	Generalised mean model	Weighted generalised mean model
α		1.67 (260)	0.727 (253)
travel time (t-1)	0.146 (48)		0.113 (37.8)
travel time (t-2)	0.327 (67)		0.289 (123)
travel time (t-3)	0.0414 (6.98)		0.00636 (-22.1)
travel time (t-4)	-0.0184 (-3.62)		0.0284 (-42.7)
travel time (t-5)	-0.0783 (-52.3)		0.000304 (2.91)
ordinary time	0.497 (76.9)		0.515 (136)
longest time	0.0267 (12.3)		0.0571 (112)
scale parameter		1.00 (734)	-0.994 (-700)
constant	-5.40 (-10.6)	-48.3 (-125)	-11.4 (-6680)
R-squared	0.966	0.916	0.972

To investigate the effect of the traffic congestion information, the parameters of each of the 4 levels of traffic congestion length information are computed and compared, and then the expected travel time is estimated by using the generalized mean. The contents of information on the highway are no congestion, 15km congestion, 30km congestion and 60km congestion. In this estimation, scale parameter and constant term are fixed to 1 and 0 to know the effect only of the parameter α. The parameters are shown in Table 2. Analyzing the results, the α increases by the increase of congestion length indicated by traffic information, and this strongly shows the trend that expected travel time is tend to be influenced by the longest time as the congestion length as indicated by traffic information increases.

Table 2 Results of the expected travel time model

	No information	With information			
		0 km	15 km	30 km	60 km
α	1.12 (295)	-117 (-3590)	-0.212 (-1301)	2.93 (283)	93.1 (42.3)
R-squared	0.903	0.958	0.860	0.832	0.854

Application to the destination choice model for shopping trips

Parameter estimation

The formations of the destination choice models for shopping are utility functions of generalized mean mentioned earlier. To validate these models, the model having the ordinary utility function is also made.

Table 3 compares the results of the models having the generalized mean utility function, fuzzy integral utility function and the ordinary linear utility function. The destination choice model is a binary logit model where the 1st alternative is the most frequented shopping place and the 2nd alternative indicates the second most frequented shopping following the 1st alternative. In these models, the formations of utility function are basically as follows:

$$V_i = w_1x_1 + w_2x_2 + w_3x_3 + \theta_t t + \theta_1 d_1 + \theta_2 d_2 + \text{const.} \tag{19}$$

$$V_i = A.\left\{w_1x_1^{\alpha} + w_2x_2^{\alpha} + w_3x_3^{\alpha}\right\}^{\frac{1}{\alpha}} + \theta_t t + \theta_1 d_1 + \theta_2 d_2 + \text{const.} \tag{20}$$

$$V_i = A.\left\{\frac{\sum_{F'\subset F}\left[\inf_{x_k\in F'} h(x_k)\wedge g(F')\right]^{\alpha}}{7}\right\}^{\frac{1}{\alpha}} + \theta_t t + \theta_1 d_1 + \theta_2 d_2 + \text{const.} \tag{21}$$

In generalized mean model and fuzzy integral model, the variables x_1, x_2 and x_3 are standardized and signs of variables x_1 and x_2 are changed to compare the importance. The variable t is the difference of the travel time expected by the individual and that obtained from information, which indicates the change of the expected travel time by the obtained information. The dummy variables, d_1 and d_2, are the mode-specific variables to indicate the existence of traffic information for each alternative. Observing these dummy variables, the sign of both parameters are positive which shows the traffic information influences individual trip-maker to change his destination. And the parameter for the difference of the time has a positive value and enough effectiveness which means the destination choice behavior under the traffic information is influenced by the expected travel time before obtained information, not only by the obtained travel time from the information.

Table 3 Results of parameter estimation

	Linear model	Weighted generalized mean model	Fuzzy integral model
α		2.04	2.30
		(15.2)	(11.0)
travel time (w_1)	-0.181	0.161	0.139
	(-1.26)	(5.73)	(1.19)
waiting time (w_2)	-0.839	0.662	1.00
	(-20.4)	(14.0)	(16.6)
attraction (w_3)	0.214	0.178	0.309
	(5.73)	(6.80)	(2.42)
λ			-0.865
			(-6.64)
scale parameter		1.31	9.52
		(19.0)	(25.7)
difference of time (expected-informed)	1.66	1.52	1.50
	(10.0)	(14.3)	(5.28)
information for 1st destination	0.395	0.402	0.614
	(3.51)	(4.95)	(6.25)
information for 2nd destination	0.860	0.853	0.833
	(7.34)	(10.4)	(8.06)
constant	0.567	0.556	0.587
	(4.37)	(6.63)	(5.23)
roh	0.192	0.193	0.198

To compare the results of these three models, the roh in the weighted generalized mean model and fuzzy integral model are increased a little than that in the ordinary linear utility function model. Therefore, in summary, the estimates of the model utilizing the generalized mean for its utility function provide better solution than that of the model using the ordinary utility function.

Segmentation of individuals

The method grouping individuals as mentioned before is to distinguish the differences in selecting certain choice sets. In this study, the parameters w_1, w_2 and w_3, and the parameter α of the weighted generalized mean model are used to make the groups or segments.

Figure 3 shows the parameters estimation for w_1, w_2 and w_3, for the individuals as mentioned in the procedure earlier. This figure reiterates that the weights of the variables in choosing the alternatives are different from individual to individual. It is found out that the parameters are distributed, so it is not a good way to make a model with pooled data. In this study, 6 particular segments are considered. In each group, the way of choosing or the decision-making method towards an alternative is considered as follows:

1. Assign importance to the travel time to the shopping place
2. Assign importance to the waiting time for parking
3. Assign importance to the attraction of the shopping place
4. Assign importance to both the travel time to and waiting time for parking
5. Assign importance to both the waiting time for parking and the attraction of the shopping place
6. Assign importance to both the travel time to the shopping place and the attraction of the shopping place

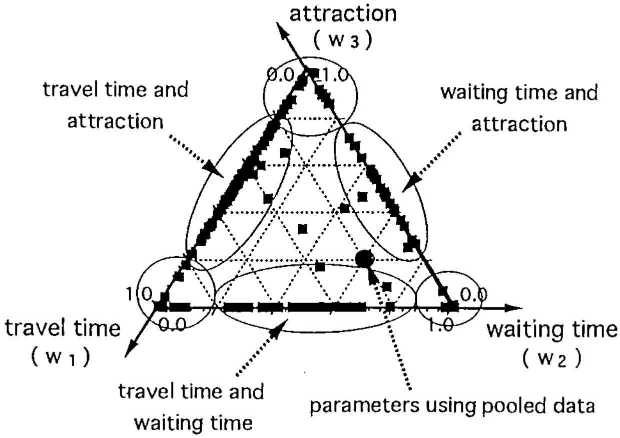

Figure 3 Distribution of individual parameters

Table 4 shows the parameters within each segment. From these results, the parameters for the weight w_i are different per group so the difference in the structure of choice is identified. Based on the obtained values of the likelihood ratio, roh, greater likelihood ratio are obtained for the segmented data compared to the original pooled data. This indicates that the model performance becomes better with segmentation.

Similarly, the results of the segmentation by the parameter α are shown in Table 5. In the estimates, 3 segments are considered and all weight parameters and the scale parameter of each group are fixed to the same values to know the exclusive effects of the parameter α. This result shows that the way of the choosing is different among individuals.

Table 4 Result of segmentation by the parameters w_i

Group	1	2	3	4	5	6	Total
α	0.963	-3.37	1.04	0.945	0.967	0.875	1.09
	(2.60)	(-15.7)	(5.06)	(3.37)	(2.31)	(1.91)	(8.10)
w_1	0.943	0.0004	0.0000	0.558	0.0000	0.388	0.204
	(4.58)	(0.881)	(0.000)	(4.46)	(0.000)	(2.78)	(6.59)
w_2	0.0568	0.989	0.332	0.442	0.596	0.174	0.604
	(0.756)	(40.1)	(3.30)	(4.31)	(3.06)	(1.92)	(11.3)
w_3	0.0000	0.0073	0.668	0.0000	0.404	0.437	0.192
	(0.000)	(0.975)	(8.24)	(0.000)	(2.85)	(2.86)	(6.64)
scale parameter	1.53	1.58	0.715	1.54	1.59	1.54	1.20
	(7.01)	(17.9)	(5.60)	(9.17)	(6.62)	(6.14)	(18.4)
constant	0.179	-0.311	0.0997	0.195	-0.373	0.604	0.156
	(1.19)	(-4.05)	(0.807)	(2.09)	(-2.44)	(3.66)	(3.77)
roh	0.373	0.372	0.227	0.202	0.280	0.200	0.147
# of samples	51	72	21	38	15	13	176

Table 5 Result of segmentation by the parameter α

Group	1	2	3	Total
α	-27.4	-2.28	8.27	0.945
	(-2.23)	(-1.62)	(5.01)	(2.09)
constant	0.123	0.368	0.157	0.131
	(1.94)	(2.61)	(2.58)	(3.20)
roh	0.218	0.147	0.107	0.106
# of samples	83	16	77	176

Totaling these two results, it can known that this segmentation method becomes fully validated.

CONCLUSION

In this paper, to analyze travel behavior which includes the effects of traffic information systems, the concept of generalized mean is adopted to solve this general problem. Concerning the parameter estimation for the homecoming trip, the expected travel time is estimated using the generalized mean. Alternative models containing the generalized mean utility function are used which can consider the different types of weighting factors for the utility function. Based on the results of segmentation of individuals, the models which can consider individual differences in the way of choosing alternatives could be estimated. Summing up the results of these models, the effectiveness in using these models can be shown.

Moreover, other structures of decision-making should be considered to express the choice behavior under traffic information systems more exactly. Also, it is important to study methods which can classify individuals into segments before the actual estimation of the model. Since, in this study, the group where an individual belongs is known only after the model estimation. These are very important topics in future studies.

REFERENCES

Dyckhoff, H. and Pedrycz, W. (1984) Generalized means as a model of compensative connectives, *Fuzzy Sets and Systems* ,14, 143-154.

Katahira, H. (1987) Buyer Behavior Research and Logit Model, in Suzuki, Y. et al. (eds) *Quantitative Analysis of Social Science*, Chapter 7, 137-157, Tokyo University Publishing Association (in Japanese).

Klir, G.J. and Folger, T.A. (1988) *Fuzzy sets, Uncertainty and Information*, Prentice Hall, N.J.

Meguro, K. and Morichi, S. (1994) An analysis of impact of traffic information using new utility concepts, *Proceedings of the 49th Annual Conference of the JSCE* 4, 772-773 (in Japanese).

Morichi, S. et al. (1983) Comparison of various utility functions for behavioral travel demand model, *Proceedings of the 3rd World Conference on Transport Research*, 159-173.