A repeated web survey for inter-regional passenger transportation demand

Hiroshima Univ. Student Member ○Yushi Sugahara
Hiroshima Univ. Regular Member Makoto Tsukai

1. Purpose

Net Passenger Transportation Survey[1] (hereafter, NPTS) is conducted once at every five years to estimate inter-regional trip demand. However, there are several limitations[2] in NPTS, such as long time to publish the data, lack in seasonal demand change and no information in trip frequency of each record. In order to overcome above limitations, long-term continuous observation using web survey is proposed in our study.

This study models the number of trip frequency and the modal choice weighted by trip frequency, using data set from web survey. Our web survey conducted for every three months, repeated four times to cover a whole year.

2. Web survey

The respondents in the web survey were those who voluntarily applied as a monitor to Intage Co. Ltd, which is a marketing research company. The main item in this survey is the inter-regional trip records made in the latest three months, which recorded up to three different destinations with its frequency. Table 1 shows the outline of the survey. In NPTS, the amount of inter-regional passenger flow is estimated with expansion coefficient for a day for each sample. In this survey, a weight by trip frequency on each person is used instead of expansion coefficient. A weight is defined in eq. (1)

$$w_s = \frac{t_s/90_{(days)}}{\sum_s (t_s/90_{(days)})} \times N = \frac{t_s}{\sum_i t_s} \times N \qquad (1)$$

Where, N is number of samples, w_s is weight for sample s, and t_s is trip frequency on sample.

$$f_{(X=y)} = \frac{\Gamma(y+r)}{y!\,\Gamma(r)} \left(\frac{r}{\mu+r}\right)^r \left(\frac{\mu}{\mu+r}\right)^y \qquad (2)$$

Where, r is a parameter, y is a trip frequency, μ is exp($\boldsymbol{x\beta}$), $\boldsymbol{x}$ is an explanatory variable vector and $\boldsymbol{\beta}$ is a parameter vector.

Table 1 The outline of the web survey

Survey type	Panel survey
Target of season : Survey month	Spring : August 2015 Summer : October 2015 Autumn : January 2016 Winter : April 2016
Target of respondents	Over 20 who live in Tokyo, Chiba, Saitama or Kanagawa
The number of respondents	34,534 in all seasons
Target of trip purpose	Business, sightseeing, private or others (excluding commuting and schooling)
Items (individual attribute)	Gender, age, profession, income, family, hometown and so on
Items (trip)	Frequency, departure area, visited area, all the modes used in trip and so on

Table 2 Parameter range

Season	Spring		Summer		Autumn		Winter	
Category	Range	Rank	Range	Rank	Range	Rank	Range	Rank
Gender	0.196	4	0.229	3	0.226	5	0.240	3
Age segments	0.075	7	0.186	5	0.124	8	0.133	6
Profession	0.774	1	0.371	2	0.557	1	0.435	1
Income	0.157	6	0.195	4	0.307	2	0.392	2
Marry	0.073	8	0.105	7	0.165	7	0.065	8
Child	0.275	3	0.097	8	0.297	3	0.130	7
Family	0.174	5	0.138	6	0.278	4	0.228	4
Hometown	0.435	2	0.416	1	0.212	6	0.199	5

3. Trip frequency model

Trip frequency is modeled by negative binomial model defined in eq. (2). The result of trip frequency model is shown in table 3. The trip frequency seasonally changes on each individual attribute item. Table 2 shows the parameter ranges between the maximum item and minimum item in each category of individual attribute. The larger the range of individual category is, the stronger the individual attribute has influence on trip frequency.

Comparing in seasonal tendency, some of the ranks are quite similar in spring and autumn, except income and hometown which are swapped in these seasons.

キーワード　trip frequency, modal choice, weight
連絡先　〒739-8527　東広島市鏡山 1-4-1　広島大学大学院工学研究科　社会基盤環境工学専攻　事務室
TEL：082-424-7819・7828

Table 3 The result of trip frequency model

Season	Spring		Summer		Autumn		Winter	
item	Coefficient		Coefficient		Coefficient		Coefficient	
Constant	0.297	***	0.197	***	-0.088		0.035	
Male	0.196	***	0.229	***	0.226	***	0.240	***
Female								
20-39	-0.075	***	-0.137	***	-0.124	***	-0.132	***
40-49								
50-59	-0.004		-0.034		-0.067	**	0.001	
Over 60	-0.031		-0.186	***	-0.092	**	-0.050	
Manager	0.183	***	0.253	***	0.357	***	0.214	***
Salaryman	0.090	***	0.118	***	0.193	***	0.143	***
Woreker								
Student	0.661	***	0.156		0.092		-0.069	
Part time	-0.113	**	-0.118	*	-0.086	.	-0.074	
Non-woreker	-0.089	**	-0.076	*	-0.028		-0.127	**
Other	0.049		0.010		0.471	***	-0.221	.
Under 1 million	0.070		0.187		0.286	*	0.323	*
1-5 million	-0.050	*	-0.008		-0.021		-0.069	*
5-10 million								
Over10 million	0.107	***	0.029		0.182	***	0.130	***
No answer	0.027		0.016		0.032		0.140	***
Married	0.073	*	0.105	**	0.165	***	0.065	
Non-maried								
No child	0.053		0.013		0.004		0.085	
One								
Two	0.007		0.032		0.047		0.081	*
Over three	0.275	***	0.097	*	0.297	***	0.130	*
Single	0.088	.	0.082		0.219	***	0.070	
With partner	-0.087	.	-0.055		-0.059		-0.158	**
Familly								
With parents	-0.041		-0.037		0.126	*	-0.150	*
Other	-0.017		0.084		0.048		-0.038	
Gunma	0.426	***	0.416	***	0.212	.	0.196	
Saitama	0.010		0.047		0.075	*	-0.003	
Chiba								
Tokyo	-0.009		0.011		0.063	*	0.079	**
Kanagawa	0.067	**	0.057	*	0.081	*	0.118	***
r	4.354	***	3.153	***	2.841	***	2.612	***
-2 log L	4.196.E+04		4.027.E+04		3.703.E+04		3.585.E+04	
R^2_{DEV}	0.077		0.067		0.089		0.070	
Samples	12499		12091		11280		10962	

Significant level : '***' 0.1% '**' 1% '*' 5% '.' 10%

4. Modal choice model

This study models the modal choice between air and train defined in eq. (3). The parameters are estimated by maximum likelihood method. Log of likelihood function with sample weight is specified in eq. (4).

$$P_m = \frac{\exp(V_m)}{\sum_{j \in m} \exp(V_j)} \qquad (3)$$

where, m is representative mode on the trip, P_m is probability to select mode m, and V_m is utility function of each mode

$$log(L) = \sum_{s} w_s \times log(P_s(i)) \qquad (4)$$

where, s is sample, w_s is a weight for s and $P_s(i)$ is the probability to choose mode i by s

Table 4 shows the result of three modal choice models, such as NPTS, weighted web data by frequency and unweighted web data. After weighting, significance level of parameters and likelihood ratio are improved. Therefore,

Table 4 The result of modal choice model

		NPTS		Weighted		Unweighted	
	item	estimate		estimate		estimate	
	Constant (railway use)	-1.886	**	-0.301	*	-0.949	**
	time(/60min)	-0.200	**	-0.112	**	-0.156	**
	fare(/1000yen)	-0.058	**	-0.032	**	-0.029	**
Individual attribute	Male	-0.499	**	-0.659	**	-0.580	**
	20 to 39	0.863	**	0.451	**	0.262	*
	50 to 59	0.147	**	0.348	**	0.229	*
	Over 60	0.000		-0.021		0.102	
	Manager	0.875	**	-0.273	*	-0.118	
	Salaryman	-0.273	**	-0.443	**	-0.034	
Trip attribute	Business	0.527	**	0.542	**	0.424	**
	Private	0.539	**	0.522	**	0.142	
	Other	0.824	**	-0.096		0.645	
	Likelihood ratio	0.585		0.343		0.323	
	Fixed likelihood ratio	0.585		0.340		0.320	
	Amount of train trip	156047		4197		3595	
	Amount of air trip	31699		1498		1447	
	Time value (yen/h)	3431	**	3499	**	5440	**

Significant level : '***' 0.1% '**' 1% '*' 5% '.' 10%

weighted model is better than unweighted model in model performance.

In table 4, amount of train / air trip are calculated multiply the number of trip and, expand coefficient in NPTS or weight in web survey. Comparing to NPTS, parameter in Manager is different in its sign, which requires further studies to be confined. On the other hands, most of other parameters become similar in NPTS and weighted.

5. Conclusion

The trip frequency model clarifies the difference in the contribution of individual attributes by seasons. The weighted modal choice model shows an improvement in the likelihood ratio compared the unweighted model.

In the future work, both trip frequency model and modal choice model has to be improved to introduce novel explanatory variables.

Reference

1) 国土交通省：全国幹線旅客純流動調査，2010

2) 奥村 誠：全国幹線旅客純流動調査の意義と課題，運輸と経済，第 72 巻，pp.21-30，2012