Private purpose inter-regional travels

A integrator of historical inter-regional migrations

Okumura Makoto[1] and Yamaguchi Hiromichi[2]
[1]Professor, IRIDeS, Tohoku University
[2]Doctor Course Student, Tohoku University
[1]mokmr@m.tohoku.ac.jp

Abstract

The service level of inter-regional transportation will be deteriorated due to the decrease of population except the busy arteries, in case of Japan. Business and sightseeing travels have been considered as main purposes of travel to analyze the inter-regional traffic flow and to plan the transportation network improvement projects. But private purpose including family visiting is also essential purpose of travel and it's social importance will be enlarged in the coming aged society, when many people must take care their old parents living in different region, as an aftermath of their former inter-regional migrations. In this paper, private purpose inter-regional travels are analyzed, with paying attention on the distribution of annual frequency of travels and using Zero-inflated Poisson Models. We conclude that the birthplace distribution and the service level of transportation affect the frequency of private purpose travels.

Keywords: travel, inter-regional, migration, zero inflated Poisson
Acknowledgement: This paper is one of the outputs of the Research Project supported by JSPS Grant-in-aid: 25289157.

1. Introduction

Comparing with business purpose travels and sightseeing travels, private purpose travels including visiting remote families, relatives and friends have not gathered much attention of transportation planners in Japan. Possible reasons for this neglect were that private purpose travels did not occupy important percentage of total domestic travels, that they do not originate regional business such as sightseeing industries, and that they are not sensitive to transportation service. In the decreasing trend of population and total inter-regional travels in Japan after 2005, however, private purpose travels only increase its number in several purposes of inter-regional travels. According to the Net Inter-regional Trip Survey (NITS) conducted by the Ministry of Land, Infrastructure and Transportation (MLIT), Japan, 20 percent of weekday trips and 27 percent of weekend trips were private purpose in 2010.
Private purpose travels including family visiting is expected to be enlarged in the coming aged society, when many people must take care their old parents living in different region. The number of people having such travel needs may be increased as a result of former demographic circumstances; decrease number of children for old parent(s) requiring

supports, increase of emigration of the children's generation according to the increased mobility and nation-wide labor market. The increased need of private purpose travels can be understood as an aftermath of their former inter-regional migrations.

Nevertheless the growing importance of private purpose travels for people; service improvements for those travels have not yet been recognized as important policy target. Now, many travelers drive themselves on the expanded expressway network, but further aging of travelers may make self-driving more difficult and public transportation accessibility will become more important. However, transport service providers do not have incentive of service improvement for private purpose travels, because they cannot expect the demand increase, because of smaller sensitivity to the accessibility improvements. In the decreasing trend of total inter-regional trips, public transportation service is facing to the risk to drop in the quantitative shrinkage and qualitative depressions.

If the needs of private purpose inter-regional trips are the result of former inter-regional migrations, which enabled efficient nation-wide location of human resources and consequent economic growth of Japan, the national government has a responsibility to maintain the essential accessibility for the trips, in case private transportation service providers cannot get enough incentive to do it.

The purpose of this paper is to analyze the travel generation pattern of private purpose based on the Survey on Time Use and Leisure Activities (STULA), conducted by the Statistics Bureau Japan, paying attention on the distribution of annual frequency of overnight travels and using Zero-inflated Poisson Models. We conclude that the residents' birthplace distribution and the service level of public transportation affect the frequency of private purpose travels.

2. Approach

(1) Data of trip generation

In order to grasp the behavior of inter-regional travels, only small portion in the residents actually generate them, it is efficient to rely on the choice-based sampling survey, rather than home-based survey. In Japan, the largest sample survey of inter-regional trips is the Net Inter-regional Trip Survey (NITS) conducted by the MLIT every five years. It collected a response to the questionnaire sheet from about a million samples, which use inter-city trains, express buses or airlines on a weekday or a Sunday in autumn. They determine the enlargement ratio to each sample sheet based on the ratio of the annual traffic counts of the link and the number of samples using there, when we cannot know the frequency of travels of each respondent. This enlargement procedure may cause the bias for traveller's characteristics, because the survey dates are not always representing the travel behavior. Especially, we cannot get precise information of personal purpose trips, because most of them are concentrated in summer and winter holidays, rather than autumn.

Another way to grasp the travel generation through a year is to conduct a home-based survey. The Statistics Bureau Japan conducts the largest home-based trip generation survey in the Survey on Time Use and Leisure Activities (STULA), every five years. It gathers survey sheets about daily time use, travels and leisure activities in the former one year of 200 thousands people over ten years old in the 83 thousand sampled households. In the survey, number of day trips and travels with overnight stay are asked. The overnight travels in Japan are further classified into three different purposes: a) sightseeing, b) personal visiting, and c) business and educational. This study uses the number distribution of overnight travels for each purpose aggregated in 47 Prefectures in 2011.

(2) Zero-inflated model

While normal distribution is widely used for statistical analysis of survey data, it does not theoretically match to the count data, which does not include negative counts. Most popular alternate distribution is Poisson distribution having the variance identical to the mean value that is derived from the random event occurring with constant ratio per interval. However, in many home-based count data includes much larger number of zero than the Poisson distribution. Such data is called as "zero-inflated count data."

Mullahy (1986) proposed two models to handle such zero-inflated data: hurdle model and with-zero model. In hurdle model, each sample faces two step decision; a binary choice whether he/she does consider to take the action or not in the first step, and if he/she considers the action, the number of actions are realized in the second step. For the positive counts distribution, a truncated Poisson distribution is usually applied. When the factors determine the binary choice are same to the factors determine the average frequency in the Poisson model, the mixed distribution can be described by a negative binary distribution. The with-zero model add an exceptional zero probability onto an ordinal Poisson distribution. In this model, two parts compose the probability of zero count; intimate zero by inactive samples and zero by chance resulted from stochastic process of countable actions. When the action is modeled with Poisson model, the With-zero model is called as "zero-inflated Poisson (ZIP) model." Mullahy (1986) showed that the ordinal Poisson model can be considered as a constrained case of both a negative binary distribution model or of a ZIP model, then Housman test can be used for the significance test of zero-inflation.

In ZIP model, the percentage of additional zero is considered as an unknown parameter. We will call it as "inactive ratio" in this paper. Lambert (1992) applied EM algorithm to estimate this inactive ratio and the average frequency of active samples, the only unknown parameter for the ordinal Poisson model. Haab & McConnell (1996) showed that a home based sample recreation activity counts data was zero-inflated data and applied a ZIP model. Consumer surplus evaluation through ZIP model was also proposed for welfare analysis of recreational facility development project. On the other hand, Gurmu & Trivedi (1996) criticized the difficulty of behavioral reasoning of the additional zero in ZIP model, and strengthen the theoretical superiority of hurdle models. They applied a Zero Infrated Negative Binomial (ZINB) model to the recreation activity counts data. Ridout *et al.* (2001) compared numerical characteristics of several zero-inflated models and show the possible serious bias of ZIP model for the over dispersed data having average of active samples is larger than one. They proposed ZIPB model in such data.

(3) Zero inflated Poisson (ZIP) model

As shown in later, overnight travel count data by the STULA is zero-inflated data. It is aggregated data including different type of people having different possibility of overnight travels. For example, samples from each prefecture includes different occupations from university professors frequently travels for conferences and academic meetings, to local bank workers responsible to their local business territories having no possibility of overnight business trips. For private purpose visiting travels, people living apart from their parents after migration from the birth place feels the need of visiting travels, while people living with their parents do not. We can add an understandable meaning to the inactive ratio: the percentage of inactive samples, relating to the demographical and other socio-economic conditions.

In ordinal Poisson (POI) model, probability of that an individual makes *k* times of overnight travels for purpose *m* in a year is formulated as:

$$p_m(k|\lambda_m) = \begin{cases} e^{-\lambda_m} & (k = 0) \\ \frac{\lambda_m{}^k e^{-\lambda_m}}{k!} & (k > 1) \end{cases} \tag{1}$$

$\lambda_m$is the only parameter to be estimated of this POI model, showing the average of travels for purpose *m* in a year.

ZIP model add an additional probability ω_m onto zero count. Then, probability of that an individual makes *k* times of overnight travels for purpose *m* in a year is formulated as:

$$p_m(k|\lambda_m, \omega_m) = \begin{cases} \omega_m + (1 - \omega_m)e^{-\lambda_m} & (k = 0) \\ \frac{(1-\omega_m)\lambda_m{}^k e^{-\lambda_m}}{k!} & (k > 1) \end{cases} \tag{2}$$

$\lambda_m$and ω_m are the parameters to be estimated. Needless to say, POI model is the special case of ZIP model when $\omega_m = 0$. Then the model selection test can be done through the t-statistics for ω_m estimator.

In the STULA, each sample chooses his/her annual times of overnight travels for purpose m from the following 9 items; {0, 1, 2, 3, 4, 5, 6-7, 8-9, 10 and more}. Here, we denote the number of responses to each item as:

$$\boldsymbol{N_m} \equiv \{N_{m,0}, N_{m,1}, N_{m,2}, N_{m,3}, N_{m,4}, N_{m,5}, N_{m,\{6,7\}}, N_{m,\{8,9\}}, N_{m,\{10-\}}\} \tag{3}$$

The log likelihood function for the ZIP model is formulated as follows,

$$\begin{aligned} \ln L(\lambda_m, \omega_m|\boldsymbol{N_m}) \equiv \sum_{k=0}^{5} N_{m,k} \ln p_m(k|\lambda_m, \omega_m) \\ +N_{m,\{6,7\}} \ln\{ p_m(6|\lambda_m, \omega_m) + p_m(7|\lambda_m, \omega_m)\} \\ +N_{m,\{8,9\}} \ln\{ p_m(8|\lambda_m, \omega_m) + p_m(9|\lambda_m, \omega_m)\} \\ +N_{m,\{10-\}} \ln\{1 - \textstyle\sum_{k=0}^{9} p_m(k|\lambda_m, \omega_m)\} \end{aligned} \tag{4}$$

We estimate the parameters $\lambda_m$and ω_m through maximization of eq.(4) by MATLAB solver.

3. Application of POI and ZIP models to the STULA data

(1) Nation wide travel count data for each purpose

At first, POI and ZIP models are compared through allocation to the nation wide travel count distribution data of the STULA 2011, for 3 purposes. The estimated result is summarized in table 1 and figure 1. From figure 1, we can find that for all 3 travel purposes, the observed distribution is zero-inflated distribution, which has much larger ratio of zero counts, but smaller ratio of 1 and 2 times counts than the ordinal Poisson distribution. Introduction of excess zero parameter in ZIP resolved the problem and improved the fitting, remarkably. Table 1 shows all parameter estimates are significantly positive, and affirms ZIP rather than POI.

Comparison of the estimates among the purposes gives interesting findings. In POI

Table 1 Parameter estimates for nation wide travel frequency data

Travel Purpose	Model		estimate	Standard Error	p-value	-log L	AIC
a) sightseeing	POI	λs	1.21	2.13x10-3	0.00	332865	665733
	ZIP	λs	2.46	0.0954	0.00		
		ωs	0.507	2.97□10−4	0.00	277803	55611
b) private	POI	λp	0.563	0.0175	0.00	289147	578296
	ZIP	λp	2.99	2.85x10-3	0.00		
		ωp	0.751	4.16x10-4	0.00	190394	380792
c) business	POI	λb	0.247	2.17x10-6	0.00	237242	474487
	ZIP	λb	3.76	0.0237	0.00		
		ωb	0.87	3.19x10-3	0.00	127177	254357

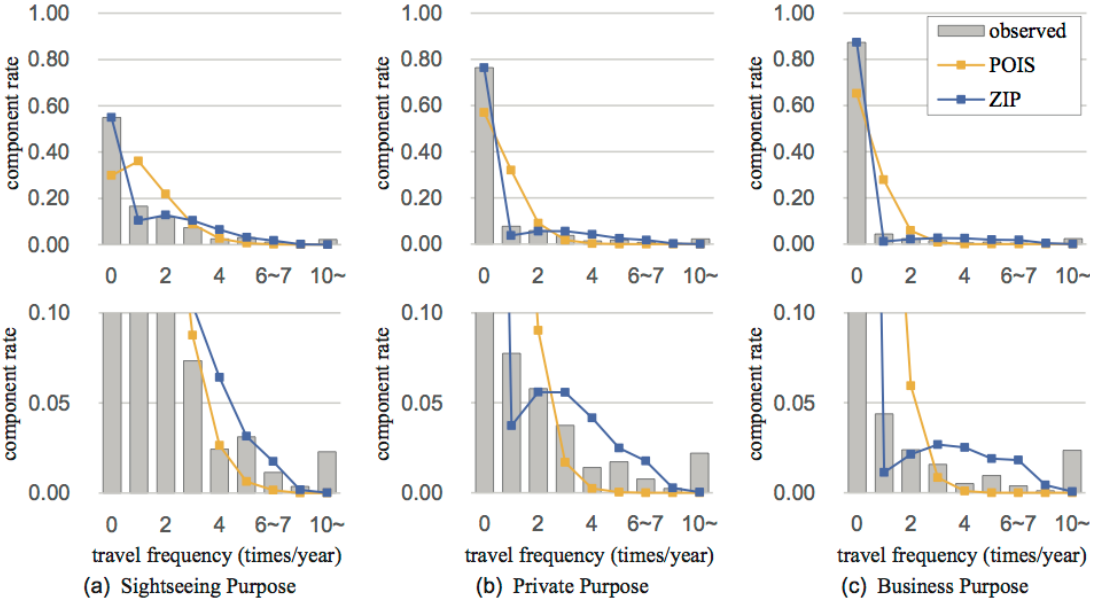

Figure 1 Travel frequency distributions and fitting of models

model, the average frequency parameter estimates have the order of a) sightseeing > b) private > c) business ($\lambda_s > \lambda_p > \lambda_b$), but this order is a result of inactive ratio difference: $\omega_s < \omega_p < \omega_b$. If we compare the average frequency of active samples in ZIP, the order is totally opposite: $\lambda_s < \lambda_p < \lambda_b$. We find that about half of the population would join the overnight sightseeing travels with average rate of 2.46 times a year, while only 13% of respondents are active business travellers with much higher average travel frequency of 3.76.

(2) Regional difference of the estimated parameter value

In the STULA 2011, we can get the overnight travel frequency distribution data aggregated in 47 Prefectures level, while the survey was not conducted in heavily tsunami-affected areas in Miyagi and Fukushima Prefectures. We applied ZIP model to the travel frequency distribution data of the three purposes in each Prefecture. Figure 2 shows the distribution of the two estimated parameters among the prefectures.

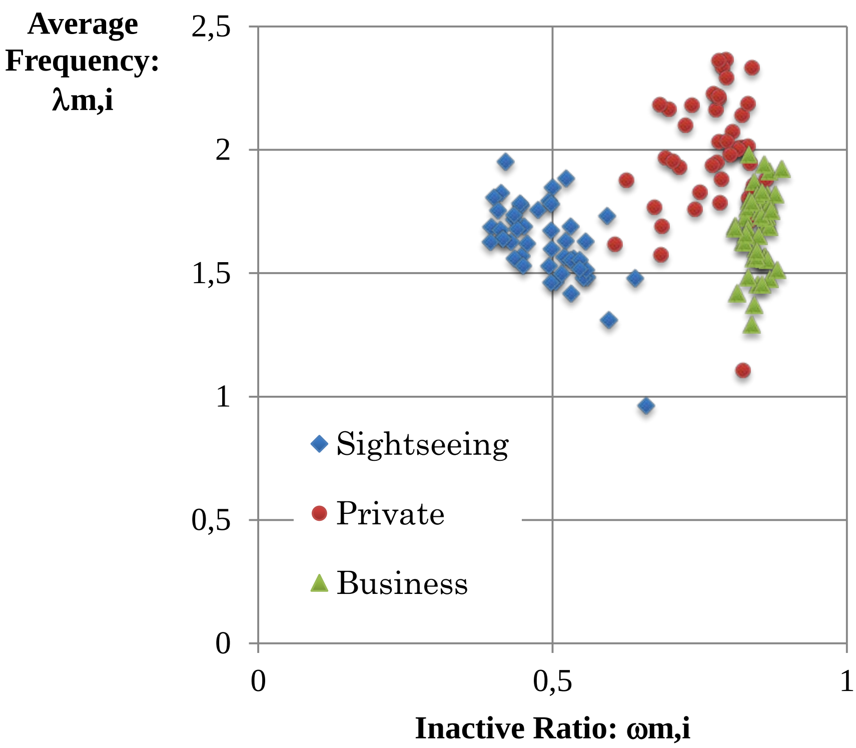

Figure 2 Distribution of the estimated parameters

At a glance, the inactive ratio of business purpose travels concentrate to a similar value, while other six parameters have wider range of distribution. It means the percentage of working people with possibility of overnight business travels in the total population takes similar value, regardless the prefectural locations or transportation accessibility. Both in the sightseeing and private, one plot is having especially low average frequency than the others; that plot is for Okinawa Prefecture, far apart from the Mainland Japan. We can find a negative correlation between the two parameters for sightseeing purpose, while private purpose parameters have positive correlation if we neglect Okinawa Prefecture.

4. Determinants for the inter-regional differences of private purpose travels

(1) Need of private purpose travels

Need of private purpose visiting travels may differ people by people according to the living location. People who are living apart from their parents after migration from the birth place; they may strongly feel the need of visiting travels to their parents, while people living with their parents in a same house do not. Okumura *et al.*(2013) estimated the birthplace distribution of people living in each Prefecture in Japan, based on the pre-residing Prefecture data in Japanese Population Census. Based on the estimated result, we can use the ratio of migrators who had been born in different Prefecture from the present residing Prefecture. Even in case of migrators, he/she can easily make a day trip to the near Prefectures such as neighboring one. Based on the same estimated result and inter-prefectural distance data, we can estimate average distance to their birthplace for the migrators in each Prefecture. Those variables were prepared for year 2005 and 2010, to be used for regression of the data in 2006 and 2011, respectively. These variables are considered as the accumulated result of the former inter-regional migration flows.

Table 2 Logistic regression model for inactive ratio

	private			
variables	estimate	S.D.	t-value	p-value
intersect	-4.843	0.542	-8.933	5.7x10-14
elderly ratio (>70yr)	-0.859	0.113	-7.597	3.1x10-11
migrators ratio	0.449	0.114	3.931	1.7x10-14
average distance to birthplace	0.423	0.075	5.674	1.8x10-7
n=91	R2	0.693	F	64.59
	R2Adj	0.679	P	3.6x10-22

(2) Explanation of the inactive ratio of private purpose overnight travels

Demographic variables defined in (1) and other socio-economic variables such as average income, ratio of tertiary industries in total employment are used as the candidate explanatory variables for logistic regression analysis of the Prefectural value of the inactive ratio parameter $\omega_{p,i}$. Considering the ease of straight understanding of the positive estimates as cause for active travelling behavior, we use the following regression formula;

$$\ln\left(\frac{1-\omega_{p,i}}{\omega_{p,i}}\right) = a_0 + \sum_k a_k X_{ki} \quad (5)$$

where, a_0:constant, and X_{ki}:value of *i*th prefecture of *k*th explanatory variable.

Table 2 shows the estimated result for the pooled data of 2006 and 2011. The migrators ratio has strong positive effect on the active join to private travels, and so for the average distance to birthplace. On the contrary, ratio of elderly people over 70 years old gives negative effect to travel generation, because such elderly people seldom have their parents in different place and they feel more difficult to prepare the monetary and physical cost for overnight travels. It is noteworthy to mention that economical conditions such as average income level do not have significant effect to the private purpose travels.

(3) Accessibility index for overnight travels

As shown by the Okinawa case in figure 2, accessibility may affect on the travel behavior. Because the STULA data is home based and not include any information of the travel destinations, we may produce an accessibility index aggregated for each origin Prefecture. When people evaluate the transportation service, several factors such as travel distance, travel time, fare, frequency and so on are integrated. In order to know the relative importance of those factors in evaluation, a gravity model of inter-regional trips is estimated, as follows;

$$\ln R_{ij} = b_0 + b_1 \ln Y_{ij} + b_2 \ln Time_{ij} + b_3 \ln Freq_{ij} \quad (6)$$

Table 3 Log linear model of private purpose trips including day trip

variables	estimate	S.D.	t-value	P-value
intersect	2.855	0.385	7.411	1.8x10-13
ln (Birth place)	1.003	0.017	59.213	0
ln (Time)	-0.923	0.056	-16.370	1.2x10-56
minimum freq.	0.056	0.019	2.860	4.3x10-3
n=2014	R2	0.819	F	3040.0
	R2Adj	0.819	P-value	0

Table 4 Logistic model of overnight travelling

variables	estimate	S.D.	t-value	p-value
intersect	-7.210	0.255	-28.252	2.7x10-145
ln(dist)	1.300	0.040	32.640	2.1x10-183
n=1775	R2	0.375	F-value	1065.4
	R2Adj	0.375	p-value	2.1x10-183

where, R_{ij}: weekly inter-prefectural private purpose trips including day trips obtained by the NITS 2005, Y_{ij}: number of migrators born in *j*th Prefecture living in *i*th Prefecture in 2005, $Time_{ij}$: travel time from *i*th to *j*th Prefecture using the fastest public transportation mode in 2005, and $Freq_{ij}$: the minimum daily transport frequency of the links included in the fastest route. Table 3 shows the estimated result of this gravity model, and shows the relative importance of travel time rather than the frequency.
Based on the estimated parameters, we define an accessibility index for overnight private purpose travels originated from ith Prefecture, as follows;

$$ACC_i \equiv \frac{\Sigma_j (Y_{ij} r_{ij}) {Time_{ij}}^{b_2} {Freq_{ij}}^{b_3}}{\Sigma_j (Y_{ij} r_{ij})} \quad (7)$$

where, r_{ij}: the probability of overnight travels in all private purpose travels from *i*th to *j*th Prefecture. It is also estimated through the following logistic model based on the NITS data.

$$\ln\left(\frac{r_{ij}}{1-r_{ij}}\right) = c_0 + c_1 Dist_{ij} \quad (8)$$

where, $Dist_{ij}$: travel distance from *i*th to *j*th Prefecture by railway. Table 4 shows the estimation result of equation (8).

(4) Explanation of the average frequency difference
At last, we estimate a linear regression model of the average frequency of private purpose overnight travels of each Prefecture, using the explanatory variables prepared in (3) and a dummy variable for 12 metropolitan Prefectures (Saitama, Chiba, Tokyo, Kanagawa, Gifu, Aichi, Mie, Shiga, Kyoto, Osaka, Hyogo and Nara).
Table 5 shows the result of estimation, where accessibility and ratio of 20-24 year old inhabitants have positive effects on average frequency of travels, while Metropolitan dummy affected negatively. We can conclude that transportation accessibility give a

Table 5 Linear regresion of average frequency of private travels

variables	estimate	S.D.	t-value	p-value
intersect	2.231	0.398	5.608	1.5x10-06
accessibility	0.143	0.061	2.346	0.024
20-24yr ratio	0.288	0.129	2.233	0.031
metropolis dummy	-0.246	0.037	-6.648	4.6x10-8
n=46	R2	0.520	F-value	15.2
	R2Adj	0.486	p-value	7.9x10-7

significant effect on the travel generation. It means that in case of decline of the accessibility, part of private purpose travels cannot be realized even though the needs are increasing.

5. Conclusion

In this paper, Zero-inflated Poisson (ZIP) models were used to analyze the travel frequency of inter-regional overnight travels in Japan. We show the different patterns are observed for 3 different purposes; sightseeing, private, and business. We conducted further analysis for the Prefectural parameter values of ZIP model, and conclude that the birthplace distribution and the service level of transportation affect the frequency of private purpose travels.

In the decreasing trend of total inter-regional trips, public transportation service is facing to the risk to drop in the quantitative shrinkage and qualitative depressions. If the needs of private purpose inter-regional trips are the result of former inter-regional migrations, which enabled efficient nation-wide location of human resources and consequent economic growth of Japan, the national government has a responsibility to maintain the essential accessibility for the trips, in case private transportation service providers cannot get enough incentive to do it.

References

1) Mullahy, J. (1986) Specification and testing of some modified count data models, *Journal of Econometrics*, 33:341-365.
2) Lambert, D. (1992) Zero-inflated Poisson regression with an application to defects in manufacturing, *Technometrics*, 34 (1): 1–14.
3) Haab, T.C., and McConnell, K.E. (1996) Count Data Models and the Problem of Zeros in Recreation Demand Analysis, *American Journal of Agricultural Economics*, 78: 89-102.
4) Gurmu, S., and Trivedi, P.K. (1996) Excess Zeros in Count Models for recreational Trips, *Journal of Business & Economic Statistics*, 14: 469-477.
5) Ridout, M., Hinde, J. and Demetrio, C.G.B. (2001) A score test for a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives, *Biometrics*, 57: 219-233.
6) Okumura.M., Kim, J. and Okubo, K. (2013) Estimation of Residential Experience, based on the Pre-residing Prefectural Data in Japanese Population Census, *Journal of the City Planning Institute of Japan*, 48(3): 933-938. (in Japanese)