

A bi-variate Kernel Estimation Model for Travel Time and Activity Duration

Yunkyung BAE
Graduate Student
Dept. of Urban Planning & Engineering
Yonsei University
134, Shinchon-Dong, Sudaemoon-Gu,
Seoul, KOREA, 120-749
Fax : +82-2-393-6298
E-mail: honghong@yonsei.ac.kr

Ju-Yeon LEE
Ph. D. Candidate
Dept. of Urban Planning & Engineering
Yonsei University
134, Shinchon-Dong, Sudaemoon-Gu,
Seoul, KOREA, 120-749
Fax : +82-2-393-6298
E-mail: ljourney@yonsei.ac.kr

Jin-Hyuk CHUNG
Associate Professor
Dept. of Urban Planning & Engineering
Yonsei University
134, Shinchon-Dong, Sudaemoon-Gu,
Seoul, KOREA, 120-749
Fax : +82-2-393-6298
E-mail: jinchung@yonsei.ac.kr

Hyungjin KIM
Professor
Dept. of Urban Planning & Engineering
Yonsei University
134, Shinchon-Dong, Sudaemoon-Gu,
Seoul, KOREA, 120-749
Fax : +82-2-393-6298
E-mail: hjkim@yonsei.ac.kr

Abstract: The activity duration in activity-based approach typically has been analyzed by various types of regression models, which manifest relationships between socio-economic variables for independent variables and activity duration for dependent variable. Among various approaches, the most frequently adopted model is the hazard-based model, which is a parametric approach because it assumes the probability distribution of the dependent variable prior to model estimation. Since the distribution is not usually aware of, the assumption of distribution function prior to estimation is sometimes very strong restriction. Especially, when the true distribution has a unique pattern (for example, bimodality shape), we have difficulty in choosing the relevant probability distribution functions. This study aims to develop the activity duration models using kernel density estimator (hereafter 'KDE'). KDE is a type of nonparametric estimation methods and can construct the probability distribution including some special features, which parametric methods hardly describe. In addition, relationships between travel time and activity duration are also investigated by the bi-variate KDE using travel diary survey data in Seoul, Korea.

Key Words: *Activity duration, Travel time, Hazard-based model, Bi-variate Kernel Density Estimator.*

1. INTRODUCTION

Trip is derived from the need to participate socio-economic activities distributed in space and time. Hence, it is desirable to focus on the analysis of socio-economic activities rather than trip derived from activities. Because of the theoretical soundness, the conventional trip-base approach is being replaced by activity-based approach. The activity pattern analysis, the one of prevailing issues in activity analysis, mainly intends to account for the allocation of time, the sequence of activity participation, and their duration in a day.

The activity duration in activity-based approach typically has been analyzed by various types of regression models, which manifest relationships between socio-economic variables for

independent variables and activity duration for dependent variable. Among various approaches, the most frequently adopted model is the hazard-based model, which is a parametric approach because it assumes the probability distribution of the dependent variable prior to model estimation. Since the distribution is not usually aware of, the assumption of distribution function prior to estimation is sometimes very strong restriction. Especially, when the true distribution has a unique pattern (for example, bimodality shape), we have difficulty in choosing the relevant probability distribution functions.

This study aims to develop the activity duration models using kernel density estimator (hereafter 'KDE'). KDE is a type of nonparametric estimation methods and can construct the probability distribution including some special features, which parametric methods hardly capture. In addition, The effect of travel time on activity duration has been studied by many researchers because it give an insight for destination choice behavior. Duration of activity (i.e., a proxy variable of utility from participating on activity) is significantly affected by travel time (i.e., a proxy variable of disutility to participating on activity). The relationships are also investigated by the bivariate KDE using travel data in Seoul, Korea.

The structure of the paper is as follows. Section 2 of paper provides a brief review of the existing literatures on activity duration and KDE related studies. Section3 provides an overview of statistical methods of hazard based modeling and KDE. Section 4 discusses the data that were used to test the model empirically. In section 5, we describe the estimated distribution using nonparametric KDE and apply this technique to travel time and duration modeling. Finally, section 6 summarizes our conclusions and a discussion of future research directions.

2. PREVIOUS WORKS

Beginning with its primary introduction into travel demand literature with Hamed and Mannering's (1993) study of post-work activity durations, the analysis techniques as well as findings on the relationship between activity duration and travel demand have gather a great deal of attention. Researches on the duration of shopping activities (Bhat, 1996a; Niemeier and Morita, 1996), the allocation of time to specific types of activities (Purvis, 1989; Kitamura, 1990), and the estimation of the current activity durations and the choice of the subsequent activity (Ettema et al., 1995) have been conducted to understand various travel patterns in a daily life. Previous research identified variables of interest in activity durations model such as the time expenditure for each activity type, the frequency of activity types, travel time, travel cost, speed, trip frequency, and so forth.

In terms of methodology, the hazard-based model is the most frequently in use for the analysis of activity duration (Kraan, 1996; Golob, 1996; Ettema and Timmermans, 1997; Bhat et al, 2004). Golob et al. (1987) have used the lognormal distribution for the baseline hazard to analyze accident durations in relation to trucks. Jones et al. (1991) modeled the accident duration of the highway in Seattle using log-logistic distribution. Also, Cox (1972) has used a partial maximum likelihood approach, which is unable to account for unobserved heterogeneity. Bhat (1996) has modeled the shopping activity duration as the nonparametric estimation of both the baseline hazard and the unobserved heterogeneity.

A few studies in various fields have compared advantages of the parametric methods with those of the nonparametric methods using KDE. Adamowski(1985, 1996) showed that the

nonparametric method is highly competitive with parametric counterparts for estimating floods and low-flow quantiles. Golob et al. (1996) developed a nonparametric kernel estimation model for estimating low-flow quantile. Their results indicated that the nonparametric method had smaller bias and root-mean-square error in low-flow quantile estimates. However, very few studies can be found for comparing the parametric approach with the nonparametric approach in transportation field. Graham (2008) has compared the parametric estimation methods with the nonparametric estimation methods for productivity and efficiency of city railways. Kharoufeh (2002) has applied KDE to activity duration modeling to investigate covariates and inspect heterogeneity. His data are segmented into a number of sets representing various life cycle stages to alternate the effect of covariates and the KS-test is used to determine any differences of the estimated duration densities.

3. STATISTICAL METHOD

3.1 Hazard Based Models

As mentioned in the previous section, the hazard-based model is the most prevailing method to analyze the activity duration. The hazard function expresses the rate at which we expect an activity to end given that an individual has been participating in the activity up to a certain time, t . The hazard function can be expressed as a function of the probability density function $f(t)$ and the cumulative distribution function $F(t)$, as shown in equation (1).

$$h(t) = \frac{f(t)}{1-F(t)} \quad (1)$$

The hazard, $h(t)$, gives the rate at which events are occurring at time t , given the event has not occurred up to time t . $F(t)$, the cumulative distribution function can be described as follows:

$$F(t) = P(T < t) \quad (2)$$

where T is a continuous random variable and t is some specific time.

The hazard function takes the one of three different shapes such as monotonically increasing, monotonically decreasing, and constant over time. Hence, the functional forms are assumed to be Exponential, Weibull, and Gamma. The basic hazard base model is called as the baseline hazard function (denoted $h(t,0)$) defined as a hazard function not considering covariates effects. Alternatively proposed proportional hazard function to include some covariates is defined as follows:

$$h(t, x) = h(t,0)\exp(\beta'x + \omega) \quad (3)$$

where, X the covariate vector, β' the vector of parameters to be estimated, ω unobserved heterogeneity and $\exp(\beta'x)$ represents effects of covariates. Most authors used parametric distributions for the estimation of baseline duration distributions and the distribution of the unobserved heterogeneity.

3.2 Kernel Density Estimator[KDE]

Continuous random variable X_1, \dots, X_n , denote n independent, identically distributed observations of X . Denote by f , the probability density function for X . Probability density function can be estimated by attempts to either parametrically or nonparametrically

approximate. In parametric density estimation, this approach requires the analysts to determine the appropriate distribution function prior to parameter estimations. However, the nonparametric KDE is able to represent the multimodal characteristics (e.g., bimodal shape) of the observations better than parametric methods, since it is not bounded by a fixed functional form (Silverman, 1986). The nonparametric KDE in this study is an alternative tool to the parametric method, since it provides local estimates using weighted moving averages of points of estimation (Lall and Bosworth, 1994). The representative nonparametric methods are the well-known histogram and kernel density estimator (KDE), and KDE is an active research topic in economic, statistics and hydrology.

3.2.1 Univariate Kernel Density Estimator

Given a set of observations x_1, \dots, x_n , a mathematical expression of a univariate kernel probability density estimator is

$$\hat{f}_x(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \tag{4}$$

where K is the kernel function and h is bandwidth that controls the variance of the kernel function. Table 1 shows examples of kernel functions typically used. The kernel function, K , is usually a symmetric probability density estimation. We used the Gaussian kernel function which is a symmetric density. The choice of the bandwidth (smoothing parameter), h , is an important issue in estimating the probability density function, since the kernel estimator is very sensitive to bandwidth. Silverman (1986) gave the appropriate bandwidth method which is adopted in this study in Eq. (5);

$$h_{opt} = 0.9 \min\{\sigma, \text{interquartile range}/1.34\}n^{-1/5}, \tag{5}$$

where σ is the standard deviation

Table 1 Examples of Kernel Functions

Kernel	K(t)
Espanechnikov	$K(t)=0.75(1-t^2), t \leq 1$ $K(t)=0, \text{ Otherwise}$
Triangular	$K(t)=1- t , t \leq 1$ $K(t)=0, \text{ Otherwise}$
Gaussian(Normal)	$K(t)=(2\pi)^{-1/2} \exp\left(-\frac{t^2}{2}\right)$
Rectangular	$K(t)=0.5, t \leq 1$ $K(t)=0, \text{ Otherwise}$

3.2.2 Bivariate Kernel density estimator

Bivariate kernel density estimator can be constructed in a similar manner. Bivariate KDE is simply expressed by a product of univariate functions with a set of bandwidths, (h_x, h_y) . The bivariate KDE is in Eq. (6).

$$\hat{f}_{xy}(x,y) = \frac{1}{nh_x h_y} \sum_{i=1}^n \left\{ K\left(\frac{x-x_i}{h_x}\right) K\left(\frac{y-y_i}{h_y}\right) \right\}, \tag{6}$$

$$\int_{\mathbb{R}^2} K(x) dx = 1, K(x) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}x^T x\right), \tag{7}$$

where n is number of observations (x_i, y_i) and h_x and h_y are bandwidths for x and y direction, respectively.

4. DATA IN USE

The data used in this research come from the Household Travel Diary of Seoul Metropolitan Area of 2006. This survey is conducted to reveal the changes of travel patterns in Seoul metropolitan area at an interval of 5 years. The sample surveyed is about 96,727 households and 330,424 persons in Seoul and its suburban cities such as Suwon, Koyang, Incheon and Sungnam. The job proportion of persons is composed of workers 43.3%, students 25.1%, housekeepers and the unwaged 24.1% and others 6.5%. Also, the sex ratio of surveyed persons is male 49% and female 51%. Survey contains person trip diary information for every household member, travel behaviors, socio-economic characteristics and weekend travel information. Each individual trip is described the trip purpose, mode, starting and ending time of each trip, location and other related attributes. Trip purpose is comprised of commuting, going to school, work-related, academy, returning to home, post-work, picking someone on, shopping and leisure. Among trip purposes surveyed, we used commute, going to school, shopping and leisure trip purpose for analysis to define the difference in travel patterns. Commute trip means working trips and destination of the trip is working place. University students are not taken into account in analysis because their travel patterns are completely different to students from elementary to high school students. Hence “going to school” trips only made by students from elementary to high school. To more detail analysis, different level of students should be treated separately, which remains for the future studies.

We estimate the distribution of activity duration and travel time using the nonparametric KDE approach and investigate correlation between activity duration and travel time. Thus, only single trips are taken into account in this analysis because sequential trips have more spatio-temporal constraints.

Table 2 Descriptive statistics of Trip data

Group	Activity types	Single trips (n)	Trip chains (n)	Single trips/ Trip Chains (%)
1	Commute	77,020	22,073	77.8
2	School	58,052	11,341	83.7
3	Shopping	6,861	1,002	87.3
4	Leisure	6,240	2,344	72.7

As shown in Table 2, the ratios of single trips to trip chains for different activity types are approximately 70~90%. Table3 shows the differences of average travel times and average activity durations for different activity types in single trips.

Table 3 Average travel times and average activity durations for activity types in single trips

Activity types	Average travel times(min)	Standard deviation	Average activity duration(min)	Standard deviation	Travel time ratio
Commute	40	25	629	139.8	0.06
School	23	21.08	433	148.7	0.05
Shopping	27	20.6	147	147.8	0.16
Leisure	39	37.7	255	184	0.13

Average travel times do not differ significantly among activity types. However, some variation between average activity duration in each activity type can be seen. Table 3 indicates that the average activity duration of discretionary activity such as shopping and leisure is between 2 and 5 hours. On the contrary, the average activity duration of mandatory activity such as commute and school is between 7 and 11 hours. Thus, activity duration of mandatory activity is much longer than its discretionary activity.

Dijst (2000) proposed the travel-time ratio concept to explain the relationship between activity duration and travel time. The travel-time ratio is the relative ratio of between activity duration and travel time, which accounts for the range of distance to activity places. People consider travel time and activity duration prior to decide when, where, what to do. Within the previously perceived trade-off concept, people have choice sets considering activity duration and travel time (Dijst, Vidakovic, 2000; Tim Schwanen, 2002). The travel-time ratio is defined as follows:

$$\tau = \frac{T_t}{T_t + T_s} \quad (8)$$

where τ indicates the travel-time ratio, T_t stands for the travel time, T_s denotes the activity duration. In this study, the travel-time ratio of mandatory activity such as commute and school was lower rather than that of discretionary activity. This result is in accordance with the findings of Dijst (1995) in Amsterdam. When we compared the travel-time ratio for single trips in this result with that of Dijst (1995), there were some differences in segmented activity types, however, we can conclude that the travel-time ratio of commute activity is lower than that of shopping activity. This analysis of travel-time ratio deserves association travel time with activity duration. Further, it has to be extended scrutinized activity types and the implication of trip chains.

5. ACTIVITY DURATION MODELS USING KERNEL DENSITY ESTIMATORS

5.1 Univariate KDE for Activity Duration

Utility obtained from economic activity highly depends on its duration. The longer activity duration the higher individual utility gains, with decreasing marginal utility. Reaching to expired time, individuals tend to end their activities due to time-spatial constraints. Hence, activity duration certainly follow a unique pattern dependent on various factors such as type of economic activity, locations, personal characteristics, etc. We employed KDE to estimate the PDF of activity duration. Eq. (9) shows the KDE density function for activity duration. Let J denote the all activities in which individuals participate. Define T_{jk} as the continuous random variable denoting the amount of activity duration individuals $j \in J$. Define $\hat{f}_j(t)$ as the PDF of T_{jk} . The kernel density estimate is defined as follows:

$$\hat{f}_j(t) = \frac{1}{n_j h_j} \sum_{k=1}^{n_j} K\left(\frac{t - T_{jk}}{h_j}\right) \quad (9)$$

where n_j is the number of observations for T_{jk} , h_j is bandwidth, and T_{jk} is the k th observations of activity duration participating in activity j .

Fig.1 shows the comparison of histogram, hazard-based model with gamma function and the

kernel density estimate for activity types. The distributions of activity duration for different purposes have different shapes. Since the mandatory activities such as school and commute are normally distributed, little probability is revealed for the extremely short activity duration. In contrast, the discretionary activities such as leisure and shopping revealed the bimodality distribution which we mainly had hypothesized.

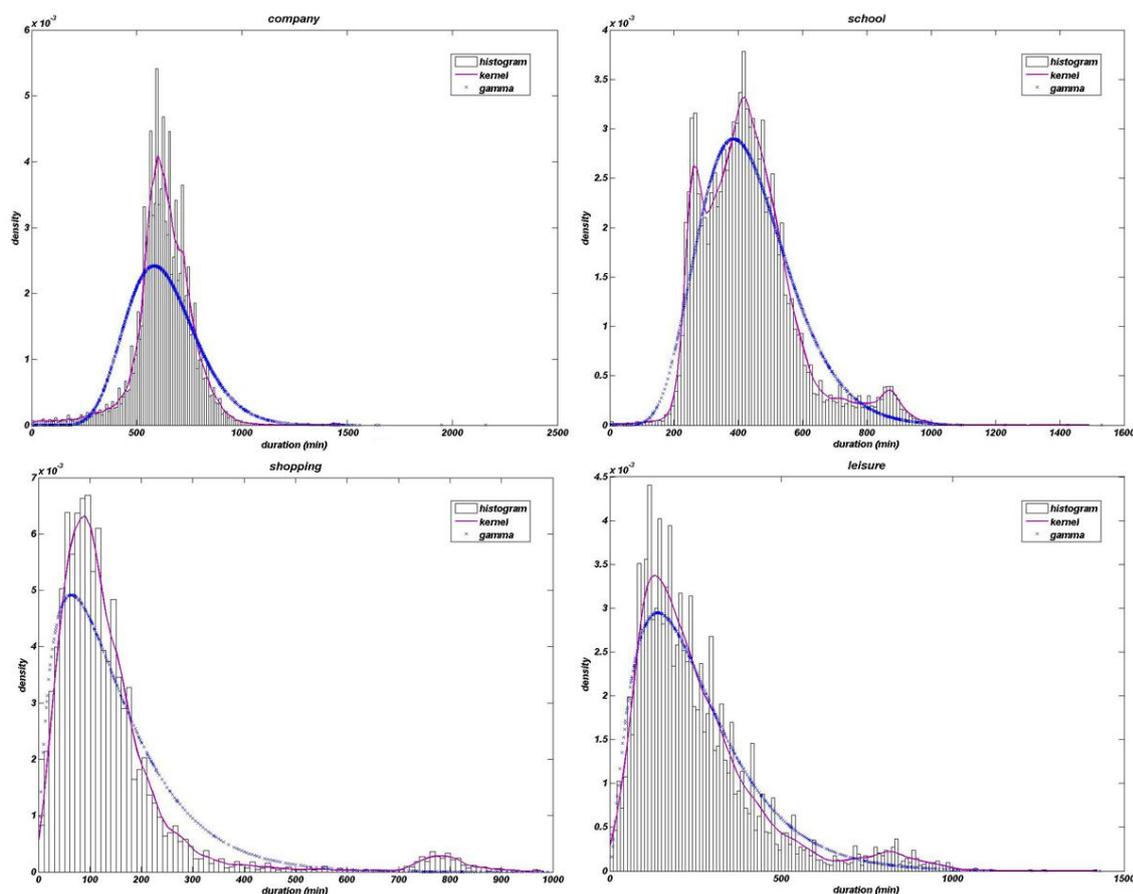


Figure 1 Comparison of probability density function using parametric methods and nonparametric methods for activity duration

Parametric approaches may have difficulty in accounting for the distinct shape of distribution as the bimodality characteristic. The reason that revealed these features can be considered as following. First, in cities, facilities attracting the influx of people are clustered and activity duration is more likely to extend in these locations. For example, in bimodality shape for shopping activity, we compared the feature of activity locations of the first mode with those of the second mode. The number of employees and total areas of wholesale/retail sale service are used to investigate the different property of each activity location attracting activity pattern/duration. Table 4 shows that mean value of the number of employees and total area for second mode are 2,233 persons and 16,364(m^2), respectively. The two values for second mode are much higher than the mean of first mode. It indicates that the activity places located in second mode are more concentrated on commercial facility than first mode. Thus, clustered locations in first mode attract individuals to stay and activity duration become longer. Second reason is the difference of individual utility for activity types. Individuals assume to maximize the utility they obtain by undertaking activities. Yet, decreasing their marginal utility, individuals finish activity due to the time-space constraints. However, according to activity types, people stay longer in the activity who consider utilities more important than constraints

or/and do not have constraints of other activities. Thus, as we have shown, density estimates of activity participation in second mode increase.

Table 4 Analysis of bimodality shape for shopping activity

	Index	N	Mean	Std. Deviation	Std. Error Mean
The number of employees	First mode	387	1,354.6	2,149.8	109.2
	Second mode	125	2,233.5	3,425.9	306.4
The total area (m ²)	First mode	387	11,283.6	16,799.6	853.9
	Second mode	128	16,364.2	19,139.7	1,711.9

In this section, we compared KDE with parametric estimates for activity duration in segmented activity types. As the aforementioned, nonparametric methods can be revealed scrutinizingly the distribution corresponded to data more than parametric methods (Silverman, 1986). Values of Histogram are used to compare parametric with nonparametric values. The optimal size of histogram is computed by Eq. (10).

$$W = 3.49\sigma N^{-1/5} \tag{10}$$

where σ is standard deviation.

Table 5 Comparison of goodness of fit between parametric methods and nonparametric methods for activity duration

Variable	Analysis methods	Commute		school		shopping		leisure	
		RMSE	SARB	RMSE	SARB	RMSE	SARB	RMSE	SARB
Activity duration	Histogram	-	-	-	-	-	-	-	-
	Gamma	0.0019	0.2095	0.0136	0.0276	0.0296	0.0458	0.0137	0.0360
	Kernel	0.0006	0.0308	0.0089	0.0156	0.0001	0.0002	0.0127	0.0312

RMSE : root mean squared error, SARB : sum of the absolute relative bias

Gamma is the distribution which prescribed in hazard-based model.

The results in Table 5 show that the PDF estimated by the nonparametric KDE has smaller RMSE and SARB than the parametric gamma function for all activity types. Especially, in case shopping activity expressed definitely the bimodality, RMSE of activity duration distribution which estimated by KDE and gamma distribution are 0.0001 and 0.0296, respectively. The fitness of KDE is highly competitive with the gamma function.

5.2 Bi-variate KDE for Travel Time and Activity Duration

5.2.1 Travel Time

Individuals have a choice to maximize their utility by activity participation while they tend to minimize travel time (disutility). The chance is higher to find a more attractive activity place on a longer travel distance, eventually activity participants determine activity locations to compare the utility gained from activity participation with the travel time (disutility) (Kraan, 1996; Kitamura et al., 1997). This situation exists clearly in discretionary activity such as leisure and shopping activity.

We analyzed nonparametric methods the relative ratio between travel time and the number of trips (activities) for activity purposes. The impedance functions can be expressed by the relative ratio between travel time and trips using the exponential, gamma, and power function. In this study, we compared the nonparametric KDE with gamma distribution frequently used as the impedance function. Eq. (11) shows the density estimate of travel time. Let I denote the single trips in which individuals travel. Define T_{ik} as the continuous random variable

denoting the amount of travel time individuals $i \in I$. Define $\tilde{f}_i(t)$ as the PDF of T_{ik} . The kernel density estimate is defined as follows:

$$\tilde{f}_i(t) = \frac{1}{n_i h_i} \sum_{k=1}^{n_i} K\left(\frac{t - T_{ik}}{h_i}\right) \tag{11}$$

where n_i is the number of observations for T_{ik} , h_i is bandwidth, and T_{ik} is the k th observations of travel time pursuing in trip i .

The distribution analyzed in this study is the travel time distribution of individual who already determine for traveling to do activities. In other words, it means that the probability of activity participation with respect to travel time. People would have the different distribution of activity behavior for activity types depending on travel time. In the case of mandatory activities such as commute and school, individual's choice is limited regardless of disutility from travel time. However, in the case of discretionary activities such as shopping and leisure, utility higher than disutility enables individual to travel to attractive places even if they are located at a long distance. As shown in Fig 2, travel time distribution analysis using KDE is expressed in detail more than gamma distribution.

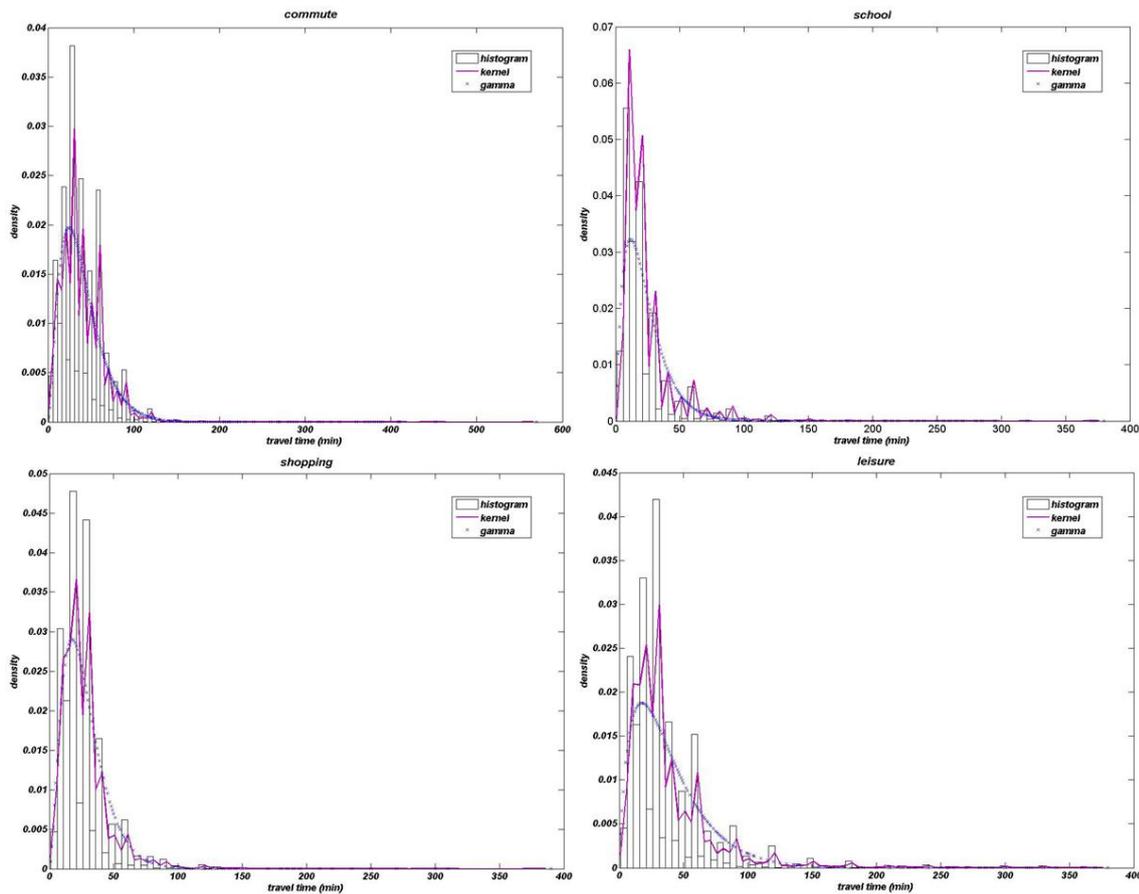


Figure 2 Comparison of probability density function using parametric methods and nonparametric methods for travel time

In this study, a bivariate kernel estimator is also employed to represent the bivariate behavior of travel time and activity duration as analyzing conditional probability function and joint probability function. First of all, to investigate the relationship between travel time and

activity duration, we constructed the conditional probability density function (CPDF) for activity duration given travel time.

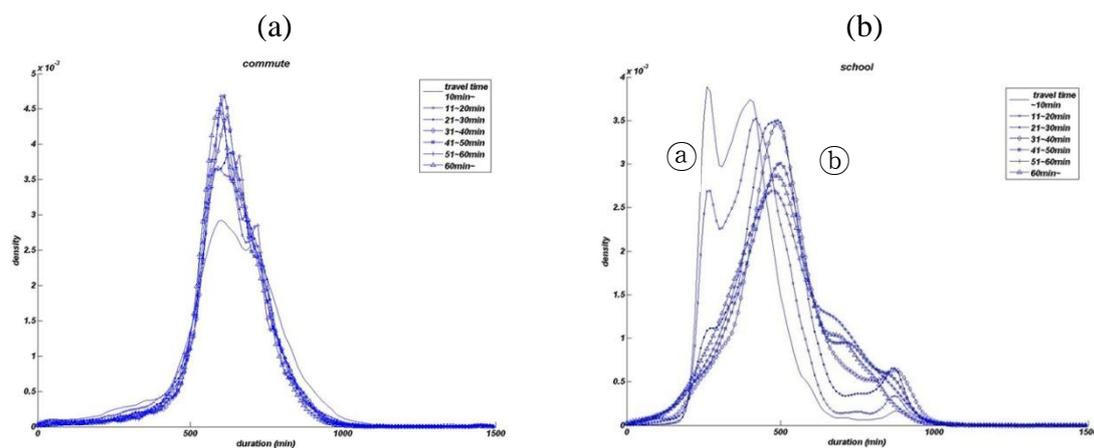
5.2.2 Conditional Probability Function of Activity Duration given Travel Time

Analyzing conditional probability density function, we can examine the interrelationship between travel time and activity duration for various activity types. Activity duration distribution given travel time 10 mins is shown in Figure 3. In the case of commute activity (see Fig 3(a)), the variation of activity duration is very small due to the characteristics of mandatory activities simply irrelevant to travel time. In other words, it indicates that there is no relationship between travel time and activity duration for commute activity. In the distribution of school activity, Fig 3(b)Ⓐ(given travel time less than 20min) and Fig 3(b)Ⓑ(given travel time more than 20min) differ significantly from each other. Features of distribution (Ⓐ) are consisted of people who went to school within are 20 min. Since most people in distribution (Ⓐ) are elementary students, they attend to school in a close distance and have little activity duration. The distribution (Ⓐ) did not reveal the shape of bimodality while the other CPDFs for school activities had bimodality. People in the distribution (Ⓑ) mainly attended to high school and had fixed travel time.

In discretionary activities such as shopping and leisure, the longer given travel time the longer activity duration. As travel time is longer, the mode of distribution is shifted to the right gradually. The each distribution of CPDF given travel time for shopping and leisure also had the bimodality shape. The analyses of CPDF show that travel time and activity duration are positively correlated, while variations for activity types exist. As the travel time and activity duration are not independent each other in the case of shopping and leisure, the each distribution of activity duration given travel time separately has different figure. The conditional probability density function is described in Eq. (12).

$$P[A_d/A_t] = \frac{P[A_d \cap A_t]}{P[A_t]} \tag{12}$$

where A_d denoted activity duration, A_t denoted travel time.



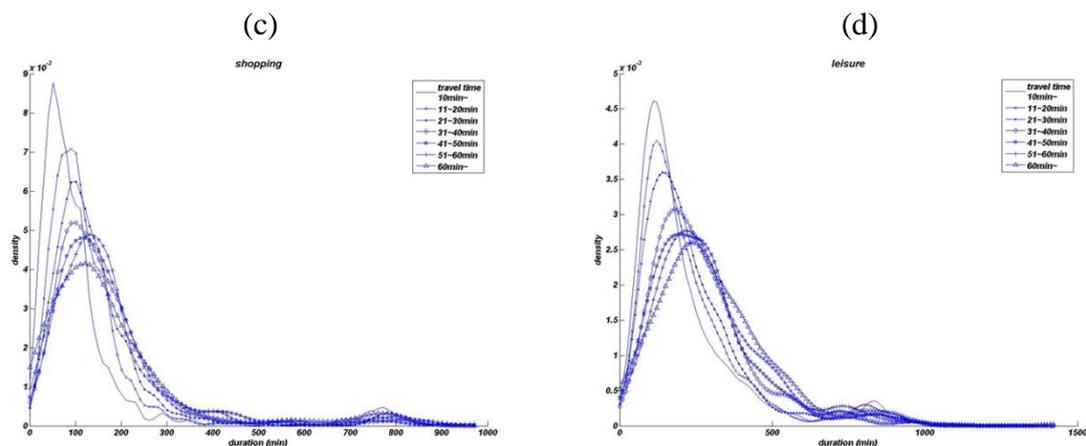


Figure 3 Conditional probability density function for: (a) work duration given travel time; (b) school duration given travel time; (c) shopping duration given travel time; (d) leisure duration given travel time

5.2.3 Bivariate KDE for Travel Time and Activity Duration

The CPDF analysis and the former researchers have proven that travel time and activity duration are interdependent (Kitamura et al., 1990, 1998; Levinson, 1999). Especially, in the case of discretionary activity such as shopping and leisure, joint probability density function has the implications which consider two variables as travel time and activity duration are closely correlated.

Eq. (13) shows the joint density estimate of activity duration and travel time. Let I denote the single trips in which individuals travel and J denote the activities in which individuals participate. Define T_{ik} as the continuous random variable denoting the amount of travel time individuals $i \in I$ and define T_{jk} as the continuous random variable denoting the amount of activity duration individuals $j \in J$. Define $\tilde{f}_i(t)$ as the joint PDF of T_{ik} and T_{jk} . The joint kernel density estimate is defined as follows:

$$\hat{f}_{ij}(t_i, t_j) = \frac{1}{n_{ij} h_i h_j} \sum_{k=1}^{n_{ij}} K\left(\frac{t_i - T_{ik}}{h_i}\right) K\left(\frac{t_j - T_{jk}}{h_j}\right) \quad (13)$$

where n_i is the number of observations for T_{ik} and T_{jk} , h_i and h_j are bandwidths, and T_{ik} is the k th observations of travel time pursuing in trip i and T_{jk} is the k th observations of activity duration participating in activity j .

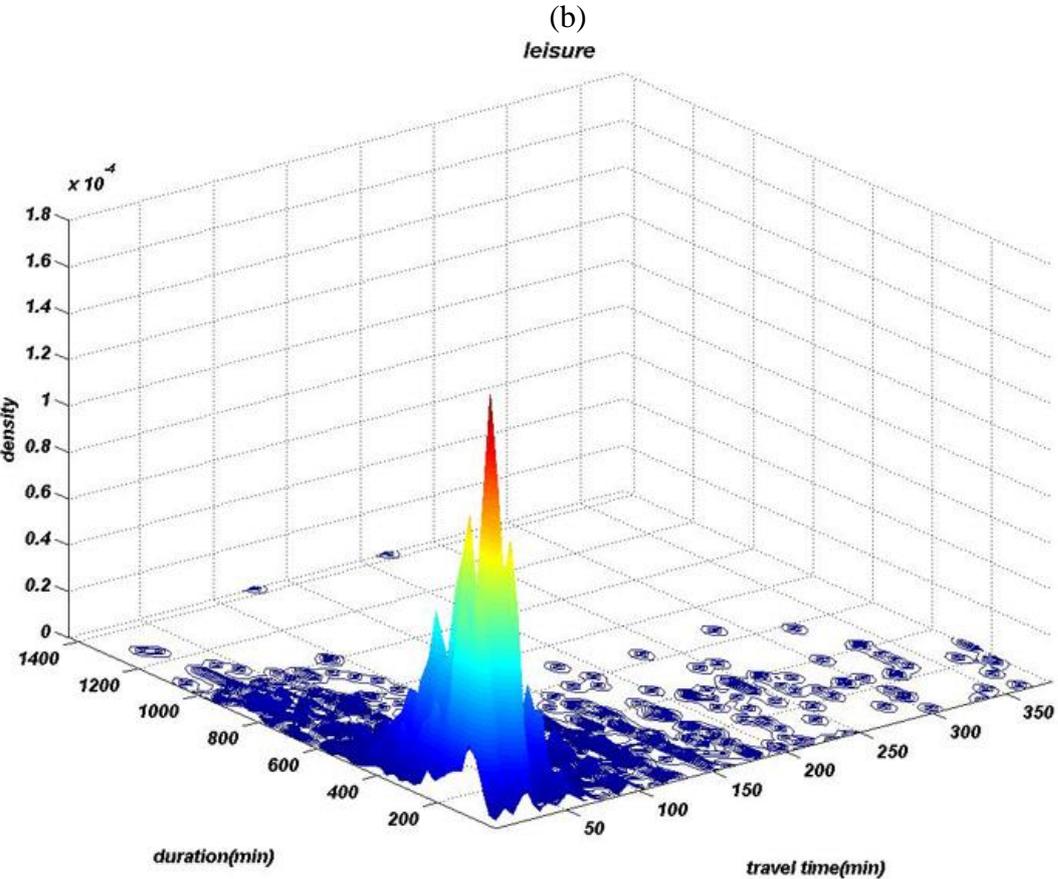
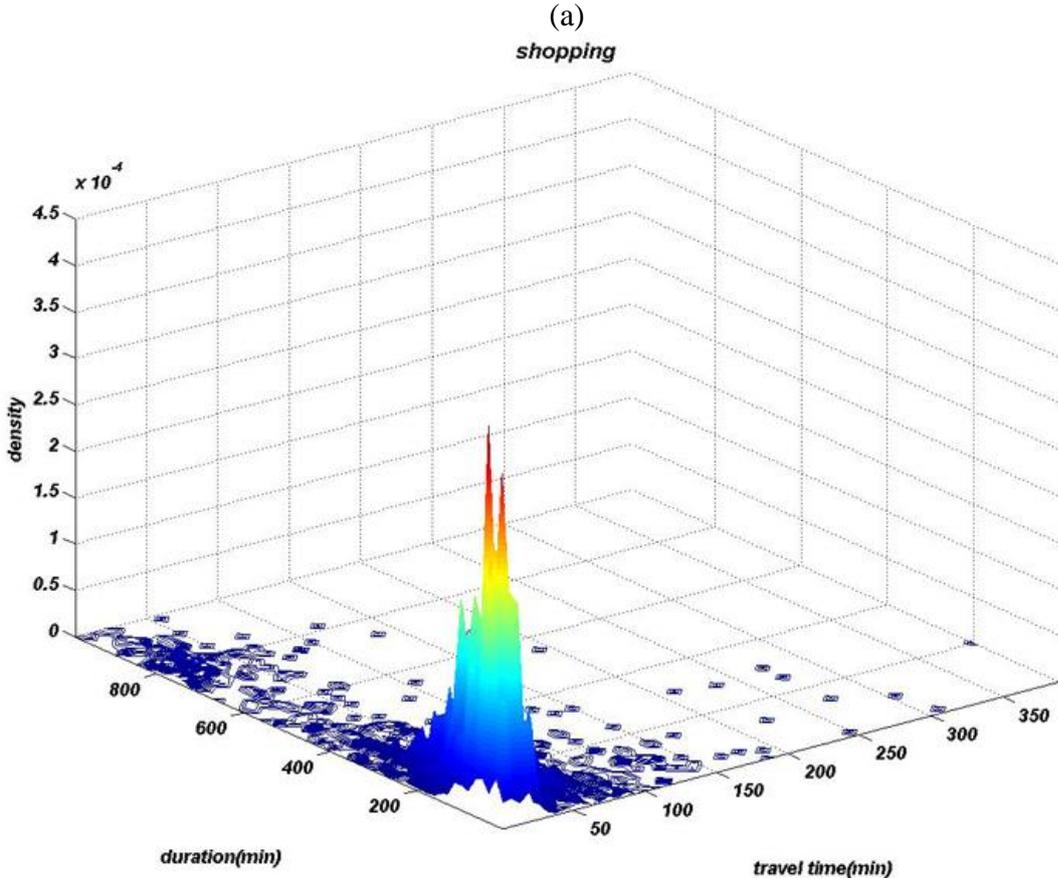


Figure 4 Joint probability density function for: (a) travel time-duration of shopping; (b) travel time-duration of leisure

As seen in Fig 4, joint distributions of travel time and activity duration for activity types had distinct variation ranges. For example, we compared the joint probability density estimate for shopping and leisure activity in equal range which travels between 10 min and 25 min corresponding activity duration between 50 min and 150 min. As the probability of shopping and leisure are 28.3% and 13%, respectively, this result varies among activity types. The distribution shows that activity duration tends to rise with travel time for leisure, but for shopping, activity duration tend to increase in spite of short travel time.

Furthermore, joint distribution reflects travel time and activity duration simultaneously. The joint distribution for shopping and leisure had the diversified choice sets of travel time and activity duration. Since shopping and leisure are discretionary activities, individuals choose not fixed choice sets travel time and activity duration, but various choice sets. Also, it is estimated that bimodality shape affect on the range of choice sets.

6. SUMMARY AND CONCLUSION

In order to forecast the travel demand precisely and reliably, it is important to understand the spatio-temporal distribution of trip and activity. The activity-based approach has significantly advanced and the study of activity duration and covariates among activity-based approaches has been developed and applied to various fields. In this study, we presented the analysis of activity participation based on the relationship of activity duration and travel time. For the better understanding of the relationship between travel demand and activity participation, distributions between activity duration and activity participation, travel time and activity participation, travel time and activity duration have examined.

The analyses were estimated using nonparametric KDE. The nonparametric KDE enables us to construct the PDF as an alternative tool to the parametric methods, because it does not require any assumption on distribution and can fit a multimodal density function. Furthermore, the KDE was tested on activity duration for different types of activity comparing gamma distribution to show a better fit of KDE. Using conditional and joint distributions, the limitations of the univariate analysis could be overcome, and the bivariate effects of travel time and activity duration could be understood more thoroughly.

Some meaningful conclusions can be drawn from the estimated distributions. First, the KDE is superior to parametric method in terms of goodness of fit measure because KDE which does not assume distribution shape expressed distribution minutely. Second, the distribution of travel time for commute and school had tighter ranges than shopping and leisure. It may show that residential locations are usually determined by places of work and school. Third, bimodality shape which is mainly dealt with in this study is seen on activity distribution for shopping and leisure and significant differences are seen between parametric distribution and KDE. Forth, CPDFs indicate the interrelationship between travel time and activity duration and have different distributions for activity types.

Also, JPDFs suggest that the distribution takes travel time and activity duration simultaneously. As taken together, the analysis illustrated here can be considered an alternative to activity choice model of individuals under time-space constraints. Generally, people have the choice jointed travel time and activity duration. The presented model means that the travel time and activity duration taken as utility and disutility, space and time are

traded off against each other. Also, this analysis using segmented data can be proposed an alternative to proportional hazard-based duration model. It is useful applying the distribution free density estimate presented to the baseline hazard function. Furthermore, the analysis of expanded multivariate estimate and various data use can be useful for unobserved heterogeneity and covariates. However, these results need to further testing and study to enhance the validity of model. More research must be undertaken for the multivariate analysis.

REFERENCES

- Adamowski, K.(1985) Nonparametric kernel estimation of flood frequencies., **Water Resource Res.** **21(11)**., p1885-1890.
- Adamowski, K.(1996) Nonparametric estimation of low-flow frequencies., **J. Hydraul. Eng.** **122(1)**., p46-49.
- B.W. Silverman, 1986. Density estimation for statistics and data analysis.
- Bhat, C.(1996a) A hazard-based duration model of shopping activity with nonparametric baseline specification and nonparametric control for unobserved heterogeneity. **Transportation research part B30**, p465-480.
- Bhat, C.(1996b) A generalized multiple durations proportional hazard model with an application to activity behavior during the evening work-to-home commute. **Transportation research part B30**, p465-480.
- Bhat, C., Gossen, R.(2004) A mixed multinomial logit model analysis of weekend recreational episode type choice, **Transportation research part B38**, p767-787.
- Cox, J., 1972. Regression models and life tables. *J. R. Statist. Soc. B* 26, 186-220.
- Dijst, M., Vidakovic, V. (2000) Travel time ratio: the key factor in spatial reach. **Transportation** **27**, p179-199.
- Ettema, D., Borgers, A., Timmermans, H. (1995) A competing risk-hazard model of activity choice, timing, sequencing, and duration. **Transport. Res. Rec.** **1493**, 101-109.
- Ettema, D., Timmermans, H., 1997, Activity-based approaches to travel analysis, Pergamon.
- Ettema, D., Bastin, F., Polak, J., Ashiru, O. (2007) Modelling the joint choice of activity timing and duration, **Transportation research part A41** p827-841.
- Golob, T.F., Bradly, M.A., Polak,J.(1996) Travel and activity participation as influenced by car availability and use. **Presented at the 75th Annual Meeting of Transportation Research Board.**
- Golob, T.F., Recher, W., Leonard, J. (1987) An analysis of the severity and incident duration of truck-involved freeway accidents., **Accident Analysis and Prevention** **19(5)**, p375-395.
- Graham, D.(2008) Productivity and efficiency in urban railways: Parametric and non-parametric estimates, **Transportation research partE** **44**, p84-99.
- Hammed, M.,Mannering, F.L.(1993) Modeling travelers' postwork activity involvement: toward a new methodology. **Transportation Science****27**, p381-394
- Jeffrey P. Kharoufeh, Konstadinos G. Goulias (2002) Nonparametric identification of daily activity durations using kernel density estimations, **Transportation research part B36**, p59-82.
- Jones, B., Janssen, L., Mannering, F. (1991) Analysis of the frequency and duration of freeway accidents in Seattle., **Accident Analysis and Prevention** **23**, p239-255.
- Kim, Tae-Woong., Valdes, J., Yoo, Chulsang. (2003) Nonparametric approach for estimating return periods of droughts in arid regions.
- Kim, Tae-Woong., Yoo, Chulsang., Valdes, J. (2003) Nonparametric approach for estimating effects of ENSO on Return periods of droughts.

- Kitamura, R., Nishii, K., Goulias, K. (1990) Trip chaining behavior by central city commuters: a causal analysis of time-space constraints. Avebury, Aldershot, p 145-170.
- Kitamura, R., Van den Hoorn, T., Van Wijk, F. (1997) A comparative analysis of daily time use and the development of an activity-based traveler benefit measure.
- Kitamura, R., Chen, C., Narayanan, R. (1998) Traveler, destination choice behavior: effects of time of day, activity duration, and home location. **Transportation Research Record 1645**, p76-81.
- Kraan, M. (1996) Time to travel? A model for the allocation of time and money. Ph. D. Thesis, University of Twente, Enschede.
- Lall, U., and Bosworth, K. (1994) Multivariate kernel estimation of functions of space and time hydrologic data., **Stochastic and statistical methods in hydrology and environmental engineering, Vol. 3**, K. V. Hipel, A. I. Mcleod, U. S. Panu, and V. P. Singh, eds., Kluwer Academic, Boston, p301-315.
- Levinson, D. M. (1999) Space, money, life-stage, and the allocation of time. **Transportation 26**, p141-171.
- Lee, Ming S., McNally, M. (2003) On the structure of weekly activity/travel patterns. **Transportation research part A37**, p823-839.
- Meyer Miller (2001) Urban Transportation Planning, McGrawhill.
- Mohammadian, A., Doherty, S. (2006) Modeling activity scheduling time horizon: Duration of time between planning and execution of pre-planned activities, **Transportation research part A40**, p475-490.
- Niemeier, D., Morita, J. (1996) Duration of trip-making activities by men and women: A survival analysis, **Transportation 23(4)**, p353-371.
- Purvis, C. (1989) Survey of travel surveys, Second Conference on Application of Transportation Planning Methods, University of North Carolina at Charlotte, Charlotte, NC.
- Schwanen, Tim., Dijst, M.(2002) Travel-time ratios for visits to the workplace :the relationship between commuting time and work duration, **Transportation research part A36**, p573-592.
- Washington, S., Karlaftis, M., M.,Mannering, (2003) Statistical and econometric methods for transportation data analysis, Chapman & Hall.