A Proposal of a Differential Equation Model for Analyzing Restaurant User Trend in Business District

Seiya Omonishi
Xiao Xiao
Hisashi Yamamoto
Faculty of System Design
Tokyo Metropolitan University, Tokyo, Japan
Tel: +81-42-585-8674, +81-42-585-8462, Email: omonishi-seiya@ed.tmu.ac.jp, xiaoxiao@tmu.ac.jp, yamamoto@tmu.ac.jp

Yukio Maruyama
Nozomi Oomiya
Faculty of Engineering
Nippon Institute of Technology, Saitama, Japan
Tel: +81-480-33-7682, +81-480-33-7483, Email: maru.y.3@nit.ac.jp, n.oomiya@nit.ac.jp

Abstract. Since the maintenance of a staff canteen is very costly, it is hard for small and medium-sized businesses to own one. Company A developed an online lunch-support and health-management system that enables the employees of these businesses to use nearby restaurants, cafes or lunchers as if they are staff canteen. Purpose of this study, is to propose a model to analyze the trend of people who have lunch in business district, by using the usage history data collected from the online lunch-support and health-management system of company A. In this paper, we firstly analyze both the choice factors and the change of consumer trends when the number of available restaurants, cafes or lunchers increases. From these analyses, factors such as location, price, and novelty are led by multiple regression analysis. Then we verify the validity of our proposal by applying actual usage history data to the proposed model.

Keywords: Differential equation model, Multiple regression analysis, Choice factor, Consumer trends, Actual usage history data.

1. INTRODUCTION

When running restaurant, manager has to consider its schedule of employees, inventory of food or logistics and reduce wastes. Therefore it is necessary to understand how many people will use. By environment and preference of users, they may use not only restaurant but also lunch-delivery. Which restaurants users use is determined by factors such as price and location. Therefore, if we find influences of factors, it will be helpful in the effective forecast of number of users.

With the cooperation of the company A, we obtained usage history data. In this data, users are employees and use some restaurants and lunch-deliveries around company A.

It is found that the total number of users increase when number of available restaurants increase from preliminary analysis. In this paper, we associate the increase of number of users with a boom, so one of our proposal is based on a differential equation model of booms (Nakagiri and Kurita, 2004).

There are many studies about choice factor or consumer trend under a static situation. Up to now, however, not many reports about choice factor or consumer trend under a dynamic situation were presented (e.g., Fujino *et al* (2008), Bucklin and Latin (1992), Urbany *et al* (1996)). One of the dynamic situation is, for example, the situation that the number of available stores or restaurants continues to be changed with time.

In this paper, we aim to analyze the trend of people who have lunch in business district, by using the usage history data collected from the online lunch-support and health-management system of company A. We reveal the choice factors of restaurants, cafes or lunchers by multiple regression analysis.

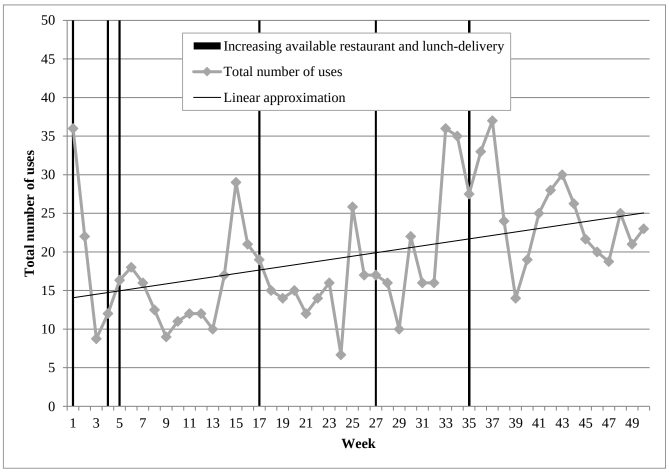

Figure 1: Changes of the total number of uses per week

Table 1: A summary of available restaurants and lunch-deliveries

	Style	Available start date	Distance	The number of signals	Price
A1	Restaurant	11/4/2014	350	2	¥789
A2	Restaurant	11/25/2014	350	1	¥989
A3	Restaurant	12/4/2014	230	2	¥1,055
A4	Restaurant	2/26/2015	600	5	¥914
B1	Lunch-delivery	5/8/2015	0	0	¥760
B2	Lunch-delivery	7/1/2015	0	0	¥430

2. USED DATA AND PRELIMINARY ANALYSIS

With the cooperation of the company A, we obtained usage history data (from 8/19/2014 to 11/5/2015). In addition, we also obtained weather data from Japan Meteorological Agency (URL: http://www.jma.go.jp/jma/index.html) and map data from Google map (URL: https://www.google.co.jp/maps). In this paper, weeks including Year-end and New Year holidays are excluded from analysis as they are outliers, because their values are much lower than other values.

Table 1 shows a summary of available restaurants and lunch-deliveries. The each style of A1, A2, A3, and A4 is restaurant and the each style of B1 and B2 is lunch-delivery. Column “Available start date” of Table 1 means the first date that each restaurant and lunch delivery is available. Column “Distance” of Table 1 means distance from company A which customers is working at to each restaurant and lunch-delivery. Column “The number of signals” of Table 1 means the number of signals located between company A and each of restaurants or lunch-deliveries. Column “Price” of Table 1 means the average of tax-including spending per customer, which is an index of price range of each restaurant and lunch-delivery.

Figure 1 shows changes of the total number of uses per week. It is found that the number of uses of almost all restaurants and lunch-deliveries within two or three weeks since becoming available is larger than that of other weeks. Figure 2 shows changes of the number of uses of existing and newly available restaurants and lunch-deliveries per week. It is found that existing restaurants and lunch-deliveries is used regularly though the number of available restaurants and lunch-deliveries increases.

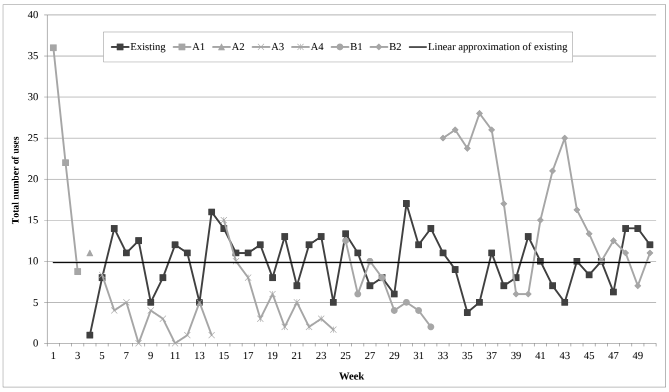

Figure 2: Changes of the number of uses of existing and newly available restaurants and lunch-deliveries per week

3. CHOICE FACTORS ANALYSIS BY MULTIPLE REGRESSION ANALYSIS

In this chapter, we analyze choice factors by multiple regression analysis. Besides, we predict the total number of uses per week by using the extracted factors in chapter 4. In section 3.1, analytic method is described. In section 3.2, analytic results are shown. In section 3.3, the results are discussed.

3.1 Analytic Method

In this section, analytic method is described. Equation (1) is multiple regression equation.

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_p x_p + e. \quad (1)$$

In equation (1), the expressions are given below.

y	: Response variable
$x_1, x_2, \ldots, x_p$	: Explanatory variable
$b_1, b_2, \ldots, b_p$	: Partial regression coefficient
b_0	: Intercept
p	: The number of explanatory variable
e	: Error

Table 2: Explanatory variables.

Characteristics or External factors
・The number of rainy days per week(Rainy days)
・Average tax-included spending per a comsumer(Price)
・Novelty of the predicted restaurant or lunch-delivery(Novelty1) (less than 3 weeks since becoming available : 0, more than 4 weeks since becoming available : 1)
・Novelty of other restaurant or lunch-delivery(Novelty2) (less than 3 weeks since becoming available : 0, more than 4 weeks since becoming available : 1)
・Distance(m)
・The number of available retaurants and lunch-deliveries(Availability)
・The number of signals(Signals)
・Style (restaurant : 0, lunch-delivery : 1)
Time factors
・(a) : The number of weeks since a restaurant or lunch-delivery
・(b) : The value which (a) divided by the number of week from week1 is
・(c) : The number of uses of the previoius week

Table 3: One of the analytic result
(time factor (b) and other characteristics and external factors are selected as explanatory variables by forced-entry method)

Explanatory variable	β	t-value	p-value
(Constant)		11.4949	0.000
Rainy days	0.0403	0.8819	0.379
Price	−0.8119	−10.6898	0.000
Novelty1	0.1260	2.3924	0.018
Novelty2	−0.1107	−2.4027	0.017
Distance	0.0658	0.4817	0.631
Availability	0.0740	1.3450	0.180
Signals	−0.9224	−8.3251	0.000
Style	−1.5138	−7.5209	0.000
(b)	−0.9609	−8.3614	0.000

We estimate partial regression coefficients and intercept by using least-squares estimation. In this paper, response variable is the number of uses of each restaurant and lunch-delivery per week. Table 2 shows explanatory variables. One of the three explanatory variables (a), (b) and (c) about the passage of time is selected by forced-entry method in order to apply multiple regression model to predict in chapter 4. Explanatory variables about characteristics of each restaurant and lunch-delivery or external factor are always selected first and then by step-down procedure.

Table 4: Tendencies of explanatory variables

Explanatory variable	By forced-entry (absolute value of β)	By step-down (selected or not)
Rainy days	very small	excluded
Price	big	selected
Novelty1	a little big	selected
Novelty2	a little big	selected
Distance	small	excluded
Availability	small	excluded
Signals	big	selected
Style	a little big	selected

3.2 Analytic Result

In this section, analytic results are shown. Table 3 shows one of the analytic results by multiple regression analysis. In Table 3, β is standardized regression coefficients. From the results, tendencies are found as shown in Table 4.

3.3 Consideration

In this section, the results are discussed. It is found that factor “Rainy days” has few influence for response variable. This is seemed because users often use any restaurant or lunch-delivery in order to have lunch. Although both “Distance” and “Signals” means physical distance, influence of “Signals” is larger than that of “Distance” within 600m at least. It is found that factor “Price” is likely to have influence. However, influence of “Price” is slightly less than that of “Signals” as shown in Table 3. It is found that factor “Novelty1” and “Novelty2” have influence for response variable.

4. PREDICTION AND EVALUATION

In this chapter, we predict the number of uses by multiple regression model, and then evaluate the prediction. In section 4.1, predictive models are described. In section 4.2, evaluative method is described. In section 4.3, the predictive results and evaluation are shown.

4.1 Predictive Model

In this section, predictive models are described. We predict the number of uses by multiple regression model.

$$\hat{y} = \hat{b}_0 + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \ldots + \hat{b}_p x_p . \tag{2}$$

In equation (2), the expressions are given below.

$\hat{y}$: Predictive value of response variable

$x_1, x_2, \ldots, x_p$	: Explanatory variable
$\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_p$	: Estimated value of partial regression coefficient
$\hat{b}_0$	: Estimated value of intercept
p	: The number of explanatory variable

The partial regression coefficients and the intercept is estimated by least-squares estimation. Response variable is the total number of uses of each restaurant and lunch-delivery per week. Options of explanatory variables are in Table 2. One of the three explanatory variables about the passage of time is selected by forced-entry method in order to conduct time-series prediction. Each model "(a)SW" and "(b)SW" is made by selecting each explanatory variables (a) and (b) and then selecting other factors by stepwise procedure. Each model "(a)SD", "(b)SD" and "(c)SD" is made by selecting each explanatory variables (a), (b) and (c) and then selecting other factors by stepwise procedure. We estimate the partial regression coefficients by using usage history data of week 1 to week 33, and predict the number of customers by using usage history data of week 34 to week 50.

We predict the number of customers by moving-average model (MA) and then compare the proposal models with this model. Equation (3) is MA with values of past four weeks.

$$\hat{y}_t = \frac{\sum_{i=1}^{4} y_{t-i}}{4}. \tag{3}$$

In equation (3), the expressions are given below.

$\hat{y}_t$	: Predicted value on week t
y_{t-i}	: Actual value on week $t-i$

4.2 Evaluative Method

In this section, evaluative method is described. The evaluative measure is MAPE (Mean Absolute Percentage Error).

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{|\hat{y}_t - y_t|}{y_t} \times 100\right). \tag{4}$$

In equation (4), the expressions are given below.

$\hat{y}_t$	: Predictive value on week t
y_t	: Actual value on week t
n	: Total number of weeks

Table 5: MAPEs of the total number of uses

	Predictive model		
	(a)SW	(a)SD	(b)SW
MAPE1	51.73%	34.23%	23.50%
MAPE2	52.71%	24.93%	20.73%
	Predictive model		
	(b)SD	(c)SD	MA
MAPE1	49.89%	16.14%	30.35%
MAPE2	46.29%	14.26%	28.69%

less error rate than that of MA   larger error rate than that of MA

We compare the proposal models with MA and evaluate them by utilizing MAPE of week 34 to week 50 and MAPE of week 38 to week 50.

4.3 Predictive Results and Evaluation

In this section, the predictive results and evaluation are shown. Table 4 shows MAPEs of the total number of uses of all restaurants and lunch-deliveries. In Row "MAPE1" of Table 4, MAPEs of week 34 to week 50 are shown. In Row "MAPE2" of Table 4, MAPEs of week 38 to week 50 are shown. In the Table 4, shaded cells have smaller MAPE than that of MA, and doted cells have larger MAPE than that of MA. It is found that MAPE of the model "(c)SD" is less than that of MA; and besides it is the least.

5. APPLICATION OF THE DIFFERENTIAL EQUATION MODEL

In this chapter, we apply the differential equation model of booms to usage history data collected from the online lunch-support and health-management system of company A. in section 5.1, the differential equation model of booms is introduced. In section 5.2, the differential equation model of booms is applied.

5.1 The Differential Equation Model of Booms

In this section, we introduce the differential equation model of booms. A boom is a social phenomenon in which some commodity, fashion or the like is suddenly prevailed among people and then is forgotten by most of them shortly. This model is based on two assumptions. The first is that regarding a boom of a certain commodity, each of the customers is considered to be in one of the following four states at a time.

State1(before a boom)	: The state in which a customer has not consumed the commodity yet.
State2(during a boom)	: The state in which he has begun to consume it after the start of the boom.

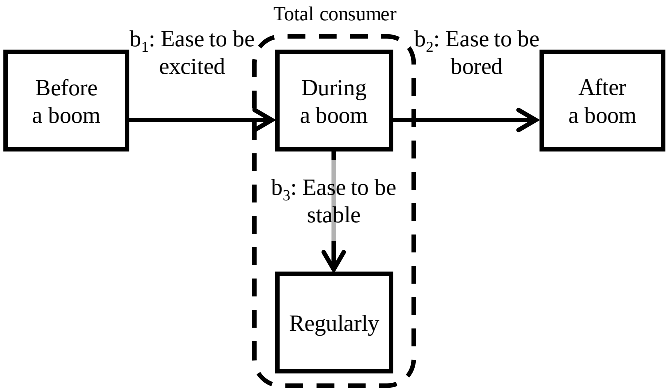

Figure 3: State transitions of customers

State3(after a boom) : The state in which he has stopped consuming it.
State4(regularly) : The state in which he consumes regularly.

Second is that the increasing rate of the number of the customers of each state is assumed to depend only on populations of the former states. This model is a model to describe the change in the numbers of the customers in these four states.

At a time t, $y_1(t)$, $y_2(t)$, $y_3(t)$ and $y_4(t)$ denote the number of the customers of State1, State2, State3 and State4 respectively. It is defined that the total number of customers keeps a constant value S. In other words, the following formula is established at any time.

$$y_1(t)+y_2(t)+y_3(t)+y_4(t)=S. \quad (5)$$

After the start of the boom at a time T, it is assume that the transition of the state of the customers is the following three ways: (i) <State1 to State2> a customer who has not consumed the commodity becomes to consume it by influence of the boom, (ii) <State2 to State3> a customer who has consumed the commodity by influence of the boom becomes to stop consuming it, (iii) <State2 to State4> a customer who has consumed the commodity by influence of the boom becomes to consume regularly. Figure 3 shows these state transitions of customers by arrows. Therefore, we formulate the following equation (Nakagiri and Kurita, 2004).

$$y_1'(t)=-b_1y_1(t), \quad (6)$$
$$y_2'(t)=b_1y_1(t)-(b_2+b_3)y_2(t), \quad (7)$$
$$y_3'(t)=b_2y_2(t), \quad (8)$$
$$y_4'(t)=b_3y_2(t). \quad (9)$$

In equations (6), (7), (8) and (9), it is defined the expression of parameter b_1, b_2, and b_3 are given below.

b_1 : Transition rate of $y_2(t)$ from $y_1(t)$. This means how easily customers are excited by the boom.
b_2 : Transition rate of $y_3(t)$ from $y_2(t)$. This means how easily customers are bored.
b_3 : Transition rate of $y_4(t)$ from $y_2(t)$. This means how easily customers are stable as to consumption after the boom.

Before a time T when the boom starts, it is assume that rate of k (≥ 0) of all customers have already consumed the commodity regularly, and the rest of all customers do not consume it. Therefore, the number of customers in each state is the following constant value.

$$y_1(t)=(1-k)S, \quad (10)$$
$$y_2(t)=0, \quad (11)$$
$$y_3(t)=0, \quad (12)$$
$$y_4(t)=kS. \quad (t\leq T) \quad (13)$$

In equations (6), (7), (8) and (9), these are initial conditions at a time T.

By solving equations (6), (7), (8) and (9) under the initial conditions (10), (11), (12) and (13), general solutions of equations (6), (7), (8) and (9) are obtained easily.

$$y_1(t)=(1-k)S\cdot e^{-b_1(t-T)}, \quad (14)$$

$$y_2(t)=\frac{b_2(1-k)S}{b_1-b_2-b_3}\left(-e^{-b_1(t-T)}+e^{-(b_2+b_3)(t-T)}\right), \quad (15)$$

$$y_3(t)=\frac{b_2(1-k)S}{b_2+b_3}+\frac{b_2(1-k)S}{(b_2+b_3)(b_1-b_2-b_3)}(b_2+b_3)e^{-b_1(t-T)}-\frac{b_2(1-k)S}{(b_2+b_3)(b_1-b_2-b_3)}b_1e^{-(b_2+b_3)(t-T)}, \quad (16)$$

$$y_4(t)=\frac{(b_2k+b_3)S}{b_2+b_3}+\frac{b_2(1-k)S}{(b_2+b_3)(b_1-b_2-b_3)}(b_2+b_3)e^{-b_1(t-T)}-\frac{b_2(1-k)S}{(b_2+b_3)(b_1-b_2-b_3)}b_1e^{-(b_2+b_3)(t-T)}. \quad (t\geq T) \quad (17)$$

5.2 Application

From preliminary analysis and choice factors analysis by multiple regression analysis, it is found that the total number of users increase when number of available restaurants

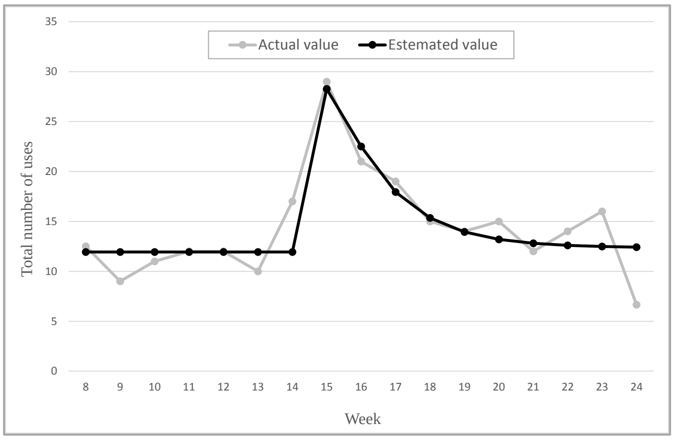

Figure 6: Estimated values when restaurant A4 becomes available

Table 5: Estimated parameter of each increasing of restaurant and lunch-delivery

Increased	A2 and A3	A4	B1	B2
b_1	0.4880	0.6301	2.4532	0.4214
b_2	4.6198	2.2531	3.9856	1.1605
b_3	0.0686	0.0097	0.0763	0.0857

increase. It seems to be a boom. In this section, we apply the differential equation model of booms to usage history data collected from the online lunch-support and health-management system of company A, so as to confirm that it is a boom actually. We assume that the time a restaurants or lunch-deliveries become available is start of a boom. Week 4 and week 5 are regarded as the same boom, because restaurant A3 become available soon after restaurant A2 becomes available.

Parameter b_1, b_2, and b_3 are estimated by least-squares estimation. Table 5 shows these parameters of each increasing of restaurant and lunch-delivery. Figure 4 shows one of the application that the total numbers of uses per weeks is estimated from week 8 to week 24. The phenomenon that the total number of users increase when number of available restaurants and lunch-deliveries increase seems to be a boom as shown in Figure 4.

6. SUMMARY

In this study, we considered the trend of people who have lunch in business district, by using the usage history data collected from the online lunch-support and health-management system of company A under the dynamic situation. We analyzed the choice factors when the number of available restaurants and lunch-deliveries increases by multiple regression analysis. As the result, we found that factors such as location, price and novelty have influence for uses of restaurants and lunch-deliveries. The prediction with these factors achieve a certain result. It seems that the phenomenon that the total number of users increase when number of available restaurants and lunch-deliveries increase is a boom, by application the differential equation model of booms.

In the future, we will apply the differential equation model of booms to prediction, by estimating parameter b_1, b_2, and b_3 with the choice factors. In addition, qualitative factors such as genre of food, taste and service should be found, because we only found quantitative factors in this study.

REFERENCES

Akihiko Fujino, Msaaki Ohba, Hisashi Yamamoto, (2008) Store Choice Orientations and Intertype Shopping Behavior toward Grocery Stores, *Japan Industrial Management Association*, 58, 459-464

Bucklin, R. E. and Latin, J. M., (1992) A Model of Product Category Competition Among Grocery Retailers, *Journal of Retailing*, Vol.68, No, 3, pp. 271-293

Kunihito Horinouchi, Naoki Takabayashi, Masaaki Ohba, Hisashi Yamamoto, (2014) A Study of Sales Increase and/or Decrease by Campaign Using a Differential Equation Model of the Growth Phenomenon, *Industrial Engineering & Management Systems*, Vol 13, No 3, pp.289-296.

Nakagiri and Kurita, (2004) ON A DIFFERENTIAL EQUATION MODEL OF BOOM. *Transactions of the Operating Research Society of Japan*, 47, 83-105.

Seiya Omonishi, Ryosuke Sawada, Junki Sekiya, Feng Lingling, Yukio Maruyama, Xiao Xiao, Hisashi Yamamoto, (2016) A Proposal of a Model for Analyzing Restaurant User Trend in a Business District(in Japanese), *Japan Industrial Management Association*

Urbany, J. E., Dickson, P. R. and Kalapurakal, R., (1996) Price Search in the Retail Grocery Market, *Journat of Marketing*, Vol.60, pp.91-I04