

おむすび屋における統計データ解析を用いた炊飯量決定について

宮内由梨花^{†1} 小野廣隆^{‡2}

概要：あるおむすび屋では1日に必要な炊飯量は当日の商品の売れ行きに応じて決定される。当日使うことのできるごはんは、来客のピークであるお昼時に向けての準備が必要であるため、朝のうちに決定しておく必要がある。そのため、当日の朝に準備しておいた炊飯量に過不足があると廃棄や欠品などの問題が発生してしまう。現状の店舗では、その日に必要なお飯の量を前年や前週の実績に基づき推測しているが、ご飯の過不足が発生する日が存在するため問題になっている。これに対し、本研究では過去データに対する重回帰分析を行うことにより、現状の感覚予測に依らず1日の必要炊飯量を決定するために必要な要素を模索する。さらに、実際にある年のデータから求めた必要炊飯量を次年度データにあてはめることで、高精度の炊飯量を予測が可能であることを示す。

キーワード：重回帰分析, 需要予測

Statistical Data Analyses for Deciding Amounts of Cooking Rice

YURIKA MIYAUCHI^{†1} HIROTAKA ONO^{‡2}

Abstract: In a rice ball shop, amount of rice to be cooked in one day is determined according to the sales of the day, but a problem is that it is necessary to decide the amount in the morning for preparation. If the amount of rice cooked in the morning is more or less than the necessary and sufficient amount, it causes disposal or shortage of rice; this could be much loss. In the current shop, the amount of cooking rice is determined based on the ideal amount of the same day of the previous week or the previous year, the performance is not always good or even bad. In this paper, we consider how we estimate the amount of cooking rice not by an intuitive sense but by more analytical way. By choosing appropriate factors of previous year's data, our multiple regression analyses can predict the amount of cooking rice with high accuracy.

Keywords: Multiple regression analysis, Demand forecast

1. はじめに

1.1 研究内容

本研究では、実在するおむすび屋を対象として、その店舗において1日に売り上げられる商品を作るために必要なお米の量を決定する。この問題の特徴としては、過不足をなくするために1日1日すべてに正確な予測を求めなければならないという点がある。また、複合施設の中に存在しているため、多数の必要炊飯量に関わる情報を考慮しなければならないという点も挙げられる。そのため、回帰分析を繰り返し、必要な要素を導入して回帰式の精度の向上を示す。

本論文の構成は以下のとおりである。引き続き1章で研究の対象となる店舗や炊飯についての説明を行い、2章でいくつかの説明変数を考察し解析を行う。そして、導き出した回帰式を用いて3章で導入のシミュレーションを行う。

1.2 店舗の特徴

今回対象とした店舗は、児童向けの有料複合施設の中にあるお店の1つである。施設内の他店舗には、雑貨屋など

に並び、当該店舗の主な商品であるお米を除いたメニューをモチーフとした飲食店も並んでいる。当該店舗では、毎朝その日必要な分のおむすびを作り、売れ残りは当日中に廃棄する。そのため毎朝、1日に売れる商品数に応じたお米を用意することが必要とされる。

1.3 炊飯量とは

本研究で求める炊飯量とは、当該店舗で商品を作るために準備するお米の量である。その量が多すぎるとお米の廃棄量や従業員の仕事量が増えてしまうという問題がある。少なすぎると商品の供給が間に合わず、お客様に提供する商品がなくなってしまう。また、お米が炊き上がるまでの間に従業員が手を余すことなどが問題視される。さらに、この施設におけるお米の炊飯には以下の条件を有する。

1. 毎朝、高々1升のお米が用意されていること
2. 所有する2つの炊飯器で用意することのできるお米の量は1合刻みで0~1升と5合刻みで2~4升であること。
3. お米を炊くには炊飯量に応じた時間が必要とされること。

当該店舗では現状として、前年や前週の実績に基づいて当日必要なお米を人間の感覚で予測している。しかし、前

^{†1} 九州大学 経済学部 経済工学科
Department of Economic Engineering, School of Economics,
Kyushu University

^{‡2} 九州大学大学院 経済学研究院 経済工学部門
Department of Economic Engineering, Faculty of Economics,
Kyushu University

年からの予測を大きく上回る来客により、炊飯が間に合わず商品を切らしてしまう場合や、多数の来客に備え炊飯を繰り返した結果、炊飯量が多すぎて廃棄となるお米が多く出る場合が少なからずある。

本研究では、そのような感覚予測では誤差が大きくなる場合の要因を推測し、当該店舗の過去データを用いた重回帰分析により、実際に導入し皆が使えるような回帰式を模索し、炊飯量を決定する。

2. 重回帰分析

以下では、実際に過去のデータを用いて回帰分析を行う。本研究は、予測で求めた炊飯量を理想の炊飯量に近づけることにより、実際の炊飯量より過不足を減少させることを目的とする。ここで今回の研究において用いる理想炊飯量について定義する。本研究では過去データにおける実際の炊飯量を基準として、その日廃棄されたお米の量や商品の数から炊飯の余剰を推定し、当日の品切れの状況から炊飯の不足を考え理想炊飯量を決定した。具体的な方法としては、実際の炊飯量からその日の廃棄量をもとに理想炊飯量を定める。一方、お米が不足した日は商品が途切れた時間をもとに理想炊飯量を定めた。ただし、ここでは販売する際に実際の販売量より多くの商品を備えておく必要も考慮して、実際の炊飯量に対して当日の廃棄量が5合にあたる約1.7kgを下回った場合の廃棄量はゼロとみなす。

回帰分析には定休日である2015年1月1日を除いた2014年9月1日から2015年8月31日までの1年分のデータを用いる。ここで、基準となる実際の炊飯量と理想の炊飯量の誤差2乗和は**136.93**であった。

2.1 確定要素を用いた回帰分析

まず始めに、曜日や天気のように当日の朝に考慮できる確実な要素と理想炊飯量との関係性を調べる。回帰分析はフリーソフト：**R i386 3.2.2[1]**を用いて行った。

ここでは以下の記号を用いて説明する。

y : 予測炊飯量

x_i : 曜日に関する 0-1 関数

(1,2,3,4,5,6,7,8=月,火,水,木,金,土,日,祝)

w_i : 天気に関する 0-1 関数

(1,2,3,4=晴れ,くもり,雨,その他)

その他の天気には、台風・大雪を含む。

h : 休日に関する 0-1 関数

o : 営業時間に関する 0-1 関数

(1) 曜日のみを変数とした場合

現状の感覚予測の際には、前年の同月同曜日を参考にし

て炊飯量を決定することが多いため、初めに年間を通した曜日と理想炊飯量の相関を調べて回帰式を得た。ここでは、月曜日を基準に日曜日までの七曜と祝日までを含めて考える。その結果、以下の回帰式が得られた。

$$y = 3.96 + \begin{cases} 0x_1 \\ -0.96x_2 \\ -0.90x_3 \\ -0.95x_4 \\ -0.62x_5 \\ 1.53x_6 \\ 2.85x_7 \\ 4.70x_8 \end{cases}$$

この結果の誤差2乗和は**892.7525**となった。

この分析では現実の炊飯量の誤差に比べても非常に大きく誤差が開いた結果となったため、新たな変数を検討する。

(2) 天気を変数に加えた場合

当該施設は全天候型の施設であるが、家族連れを主なターゲットとする大衆向けの施設であるため、以下では曜日に加え天気との相関についても考える。晴れを基準として、くもり、雨、また子供の外出に大きく影響を与える大雪や台風といった悪天候をその他として考えると以下の結果が得られた。

$$y = 3.95 + \begin{cases} 0x_1 \\ -0.93x_2 \\ -0.92x_3 \\ -0.98x_4 \\ -0.58x_5 \\ 1.51x_6 \\ 2.83x_7 \\ 4.74x_8 \end{cases} + \begin{cases} 0w_1 \\ -0.14w_2 \\ 0.35w_3 \\ -0.86w_4 \end{cases}$$

誤差2乗和は**883.1532**であった。

予想に反し、曜日のみの予測のときと比べ大きく精度を高めることはできなかった。天気に関する変数の係数から、やはり全天候型の施設であるため、家族連れが外出を避けるほどの強風や大雪では無い限り天気はあまり関係しないように考察される。そのため、以下では新たな定義による変数の導入を検討する。

(3) 休日の導入

曜日と天気の導入による回帰式の結果、とくに誤差が大きい日を検証すると予測炊飯量が理想炊飯量を大きく下回る平日が目立ち、特定の平日に予想を超える来客があるこ

とが推定できた。短期の需要予測としては電力に関する研究が数多くあり、そのなかでも GW や年末年始のような平日と異なった動きが見える時期は特異期間として分析されている[2][3]。そこで、ここでは当該施設が児童向けの施設であることを考慮し、子供たちの休日を定義し変数として導入する。

休日 h の定義は以下のようにする。

1. 土日祝
 2. 福岡市の幼稚園・小学校の運動会や参観日による代休
 3. 幼稚園の終業式から始業式の間の長期休み
 4. 幼稚園の卒業式から入学式の間の長期休み
- 対象の複合施設が存在する福岡市において、小学校の長期休みは全て幼稚園の長期休みと重複していたため、条件の対象を幼稚園とした。

$$y = 3.34 + \begin{cases} 0x_1 \\ -0.75x_2 \\ -0.78x_3 \\ -0.83x_4 \\ -0.40x_5 \\ 0.37x_6 \\ 1.69x_7 \\ 3.63x_8 \end{cases} + \begin{cases} 0w_1 \\ -0.11w_2 \\ 0.36w_3 \\ -1.42w_4 \end{cases} + 1.75h$$

誤差 2 乗和は **734.0014** となった。

係数の減少により、曜日の相関が小さくなり休日を表す変数が理想炊飯量に強い相関をもつことが考えられるので、引き続き変数として使用する。しかし、これまでの分析に比べると誤差は小さくなったものの現状の感覚予測に比べ、誤差が大きいので、引き続き新たな変数の導入を検討する。

(4) 営業時間の考慮

感覚予測をする際に、最も炊飯量が必要であると判断される日に連休の中日がある。これまでの分析では、休日をひとまとまりとして考えていたため、連休初日、中日、最終日がとくに区別されることなく分析されていた。ここで、連休の中日については複合施設の営業開始時間が通常の営業開始時間より早くなることを変数として導入する。営業開始が早い日に対して o という 0-1 変数を考える。

$$y = 3.37 + \begin{cases} 0x_1 \\ -0.76x_2 \\ -0.88x_3 \\ -0.91x_4 \\ -0.55x_5 \\ 0.31x_6 \\ 1.38x_7 \\ 2.51x_8 \end{cases} + \begin{cases} 0w_1 \\ 0.03w_2 \\ 0.24w_3 \\ -2.04w_4 \end{cases} + 1.59h + 4.22o \cdots \textcircled{1}$$

この結果、上記の回帰式が得られ、誤差 2 乗和は **506.0888** となった。これにより、誤差を大きく縮めることが出来たため、開店が早い日が必要な炊飯量の決定に大きく関わっていることがわかる。

2.2 入館者数を用いた回帰分析

ここまでは、確定要素のみを用いた分析を行ってきたが、ここからは入館者数 X を用いた分析を行う。ここでは、店舗への客数データが入手困難であったため、施設への入館者数で代替する。施設への入館者数と店舗への客数は比例し、結果として売り上げに影響すると推測したためである。理想炊飯量 y を目的変数として回帰分析を行った結果、以下の式が得られた。

$$y = 0.00299203X + 0.8141859$$

誤差 2 乗和は **169.1152** となり、これまでの確定変数での回帰式よりも精度が高いことがうかがえる。また、回帰分析式は図 1 に示す通り。ここからも、入館者数と理想炊飯量の相関が非常に高いことが予想できる。

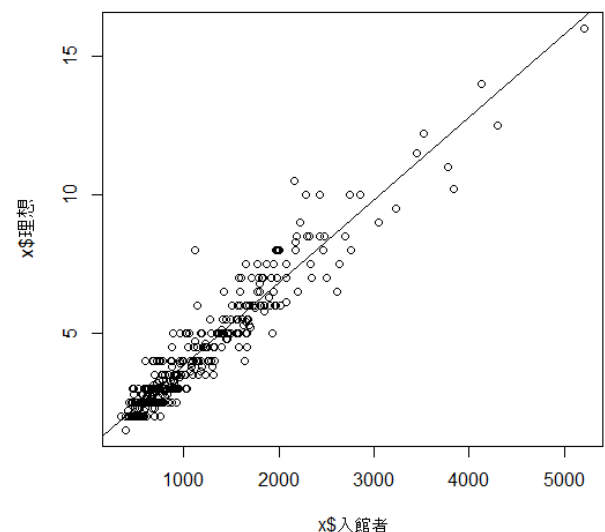


図 1 理想炊飯量と入館者の回帰曲線式

Figure 1 regression line

2.3 曜日・天気に入館者数を加えた回帰分析

最後に, 確定要素に入館者数を加えて分析してみると, 以下の回帰式が得られた.

$$y = 0.96 + \begin{cases} 0x_1 \\ -0.12x_2 \\ -0.13x_3 \\ -0.11x_4 \\ 0.01x_5 \\ -0.37x_6 \\ 0.05x_7 \\ 0.99x_8 \end{cases} + \begin{cases} 0w_1 \\ -0.04w_2 \\ -0.09w_3 \\ -0.50w_4 \end{cases} + 0.11h - 0.33o + 0.003X \cdots \textcircled{2}$$

誤差 2 乗和は **166.2949** となり, ここまでの回帰式の中で最も精度の高い結果が得られた. また, これまでの全ての要素を含めても回帰式の項数は最大で 6 つに留まるため, この回帰式の精度が高ければ, 実際に当日の朝用いることを考慮しても実用的である.

2.4 対象範囲の省略

ここで, 先に述べた理想炊飯量の決定の問題点について考える. この章のはじめに述べたように, 今回使用する理想炊飯量についてその値を決定するにあたって, 全ての日において実際の炊飯量をもとに決定した. そのため, 理想の炊飯量について厳密な必要炊飯量とは言えないということと, 実際の炊飯量と理想の炊飯量との差がない日が多く存在するという問題点が浮かび上がる.

このことを考慮して, ここでは過去のデータの対象とした 1 年間の中から実際の炊飯量と理想炊飯量の差が存在しない日を対象の範囲から外し, 実際と理想の差が生まれた日のみのデータから改めて回帰分析を行った. 対象となったのは 1 年間のうち定休日を除いた 364 日のうち 78 日間であり, その範囲における実際の炊飯量と理想炊飯量の誤差 2 乗和は変わらず **136.93** である.

(1) 確定要素のみによる回帰分析

先に出た変数から入館者数のみを除いたものを説明変数とした回帰式は以下の通りになった.

$$y = 3.05 + \begin{cases} 0x_1 \\ -0.37x_2 \\ -0.82x_3 \\ -0.92x_4 \\ -0.48x_5 \\ -0.77x_6 \\ 0.76x_7 \\ 1.51x_8 \end{cases} + \begin{cases} 0w_1 \\ 0.04w_2 \\ 0.45w_3 \\ -2.79w_4 \end{cases} + 3.11h + 3.12o \cdots \textcircled{1}'$$

ここで得られた誤差 2 乗和は **157.2181** となり, 実際の炊飯より精度としてわずかに下回る結果となった.

(2) 入館者数を利用した場合

先ほどの分析のように, 今回の範囲においても確定要素に入館者数を加えた回帰式を求めた結果は以下の通りになった.

$$y = 1.48 + \begin{cases} 0x_1 \\ -0.40x_2 \\ -0.58x_3 \\ -0.80x_4 \\ -0.71x_5 \\ -0.59x_6 \\ -0.42x_7 \\ 0.51x_8 \end{cases} + \begin{cases} 0w_1 \\ -0.23w_2 \\ -0.24w_3 \\ -0.39w_4 \end{cases} + 0.43h - 0.45o + 0.003X \cdots \textcircled{2}'$$

誤差 2 乗和は **64.71485** となり, 現実の炊飯量より精度の高い予測を求めることが出来た.

1	日付	全部予測	予測差 ²	予測差
68	3月27日(金)	3.7	1.444804	1.2
69	7月19日(日)	12.4	1.838736	1.4
70	4月6日(月)	4.4	1.926544	1.4
71	8月23日(日)	8.5	2.362369	1.5
72	12月31日(水)	6.1	2.515396	1.6
73	9月15日(月)	8.9	2.540836	-1.6
74	11月24日(月)	7.6	2.553604	1.6
75	7月5日(日)	7.0	4.141225	2.0
76	3月16日(月)	4.7	4.695889	2.2
77	11月22日(土)	6.2	5.0176	2.2
78	5月4日(月)	14.9	5.602689	2.4
79	9月29日(月)	5.2	7.568001	-2.8
80			64.71485	

図 2 予測差

Figure 2 prediction interval

図 2 の右列はこの回帰式から求められる予測炊飯量とその炊飯量の理想炊飯量との差である。予測結果と理想炊飯量の差が少ない順に並んでおり、太枠内は特にその差を縮めることが出来なかった値である。誤差の 2 乗和は現実より精度が高まったものの、各日を見てみると最も誤差が多い日には 2.8 升の不足が発生するため、更なる精度の向上が期待される。

3. シミュレーションによる評価

本章では、これまでの分析をもとに導き出した回帰式を実際に店舗に導入した場合の結果を実証する。回帰分析では 2014 年から 2015 年におけるデータで解析を行ったため、本章では 2016 年の営業開始日である 1 月 2 日から 2 月 2 日までのデータと照らし合わせてここまで推定した式の精度を確認する。

① 確定要素のみのシミュレーション

はじめに、実際に店舗に導入した際の実用性を確かめるため、当日の朝に知ることが出来る確定要素のみを説明変数に用いて予測を行う。前年までのデータから求めた①式に 2016 年のデータを代入して予測式を割り出し、実際の炊飯量と比較した。

1	日付	曜日1 予測	予測差	予測差 ²
29	1月5日(火)	4.2	1.73	2.9929
30	1月7日(木)	4.1	2.08	4.3264
31	1月8日(金)	4.4	2.14	4.5796
32	1月24日(日)	4.3	2.3	5.29
33	1月2日(土)	9.5	3.49	12.1801
34				45.2655

図 3 予測の導入①

Figure 3 Introduction of prediction I

このときに得られた誤差 2 乗和は **45.2655** で、最も誤差がある日で 3.5 升の余剰となった。しかし次点では 2.3 升の余剰で収まった。また、この誤差が大きく生じた日をそれぞれ解析すると、最も誤差のある日は三が日に含まれる営業初日であり、次点は近年では珍しい大雪に見舞われた日であったため、過去における特定の日のデータと照らし合わせれば回避できる誤差の範囲内であると考えられる。また、2.4 節で抽出した①'式にも代入して推定すると、その誤差 2 乗和は **72.4308** となり、過去データと照らし合わせた場合と異なり、1 年分のデータから求められた回帰式より精度を欠く結果となった。

②入館者予測を用いたシミュレーション

入館者予測の精度

次に、入館者数を説明変数に入れた②式で炊飯量を予測する。1 日の実際に入館者数は、当日営業後でなければ値を得ることができないが、年間を通しての施設への入館者数は各日の予測が各店舗へと配布される。この予測入館者数のデータを用いることで実際の運用を可能にする。

まず実際にこれからデータの導入シミュレーションで用いる 2016 年の 1 月の 1 ケ月間の予測についてその精度の高さを確かめる。

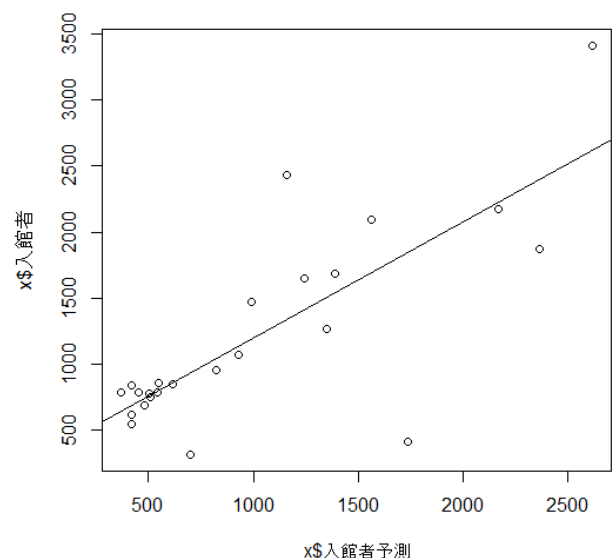
図 4 入館者の相関
相関係数 $R=0.7878433$

図 4 とその相関係数より、入館者と入館者予測の値は相関が低くはないが著しく値が異なる日が存在することが分かる。図 5 に予測値と現実の入館者数がとくに大きく離れた日をいくつか示す。

日付	入館者	予測入館者	予測差	予測差 ²
1月30日(土)	1102	1554	-452	204304
1月16日(土)	993	1476	-483	233289
1月3日(日)	2366	1874	492	242064
1月11日(月)	1565	2093	-528	278784
1月10日(日)	2619	3409	-790	624100
1月31日(日)	1247	2330	-1083	1172889
1月17日(日)	1160	2429	-1269	1610361
1月24日(日)	1740	416	1324	1752976

図 5 予測差の大きさ

図 5 は実際の入館者とその予測が最も大きく離れている日のデータであり、太枠内はその差が 500 人を上回っている日を表している。ここから、著しく入館者とその予測の値が離れている日の多くが土日祝などの休日であり、その日の予測入館者が 1000 人を超えていることがわかる。この結果より、入館者の予測の精度に関しては、ある程度の予測人数を範囲で区切れば、その精度を向上させることが出来ると予想される。

導入シミュレーション

(1) 全範囲を対象とする

実際に、先に求めた全ての要素を説明変数とした②式における未知の入館者数 X に、施設が公表している入館者予測を代入して予測を行う。

1	日付	全部1 予測	予測差	予測差 ²
29	1月25日(月)	2.6	1.063	1.129969
30	1月5日(火)	3.7	1.203	1.447209
31	1月2日(土)	7.2	1.21	1.4641
32	1月3日(日)	8.1	2.128	4.528384
33	1月24日(日)	5.8	3.84	14.7456
34				41.96579

図 6 予測の導入②

Figure 6 Introduction of prediction II

得られた必要炊飯量の予測の誤差 2 乗和は **41.96579** となり、予測された入館者を説明変数に含まない確定要素のみからなる予測より僅かに精度の高い結果となった。1 日ごとの予測差を検証すると、最も予測差の大きい日で 3.8 升の余剰があり、次点で 2.9 升の不足が生まれるといった結果となった。ここで最も予測差が大きくなった日も、先に述べた大雪に見舞われた日であり、前日より外出を避けるように呼びかけられていた日であるため、当日にある程度の予測を立てられると考えられる。しかし、次点は際立ったイベントなども当たらない普通の日曜日であるため、さらなる予測の改善が求められる。ここで考えられる原因の 1 つに、前年 12 月より導入された年間パスポートの影響が考えられる。このパスポートは、1 度目の来館時に来客が購入することができ、それ以降の 1 年間はパスポートを所有していれば入館料を払わずに施設へ入館することが出

来るというキャンペーンである。これは、回帰分析にデータを利用した年にはまだ実施されていなかったキャンペーンであり、施設への再訪者を大きく増加させようとする事象であるため、このことを含めて検討しなおした分析も必要とされる。また、2.4 節で抽出した②' 式にも代入をして予測を行った。この予測から導かれた誤差 2 乗和は **51.64293** となり、ここでも年間を通したデータから得た回帰式よりも精度が低い結果となった。そのため、実際の導入においては、データを省略して用いた回帰式より 1 年分のデータを全て用いた回帰式のほうが実用的であるといえる。

(2) 範囲を限定する

また、先ほどの予測の精度の検証に合わせ範囲を“入館者の予測が 1000 名を下回っている日”のみに限定した予測も行う。ここで、範囲の対象となったのは、先の 32 日のうち 21 日となった。

1	日付	全部1 予測	予測差	予測差 ²
17	1月26日(火)	2.3	0.331	0.109561
18	1月6日(水)	3.4	0.369	0.136161
19	2月2日(火)	2.1	0.644	0.414736
20	1月7日(木)	2.8	0.759	0.576081
21	1月25日(月)	2.6	1.063	1.129969
22	1月5日(火)	3.7	1.203	1.447209
23				5.711208

図 7 予測の導入③

Figure 7 Introduction of prediction③

この範囲に限定し②式を用いて予測を行った。その結果、誤差 2 乗和は **5.711208** となり、その予測差は 1.2 升以内に収まった。このことから、入館者の予測が 1000 人未満の日には先に求めた回帰式が有効であると言える。

4. おわりに

本研究では、重回帰分析に基づきおむすび屋における炊飯量の予測を行った。より精度の高い予測を行うため、様々な説明変数を導入し、回帰式を検討した。結果として、曜日・天気・休日・営業時間といった不変的な確定要素に施設の提供する入館者の予測数を加えた 5 つの変数を持つ回帰式が最も高い精度での予測を得られることがわかった。特に、予測される入館者の人数が 1000 人を下回る日においては、高い精度で予測が出来た。また、今回求めた式は少ない説明変数から得られる予測なので、誰でも簡単に素早く計算できるため、実際に当日の朝の予測に用いるために適したものとなった。しかし、最も誤差の出やすい休日における回帰式については精度が低い時もあり、より精度の高い回帰式を導く必要がある。

さらに、本研究では入館者予測値を施設提供のものに依っているが、より正確な需要予測の為には精度の高い入館者予測が必要であるため、これについても考えなければならない。このような研究としては例えば林田・渡辺の博覧

会における日別入場者数予測[4]などがあり, 本研究の発展の参考になるかと考えている.

参考文献

1“ The Comprehensive R Archive Network”

(<https://cran.r-project.org/>)

2 灰田武史, 武藤昭一. 重回帰手法に基づいた最大需要予測システムの開発. オペレーションズ・リサーチ, 平成 8 年 8 月号, 第 41 巻, 第 8 号.

3 田中英一, 長谷川淳, 伊藤正義. 重回帰分析と階層型ニューラルネットによる翌日需要予測. オペレーションズ・リサーチ, 平成 8 年 8 月号, 第 41 巻, 第 8 号.

4 林田和人, 渡辺仁史. 博覧会における日別入場者数変動に関する研究. 日本建築学会計画系論文集, 第 467 号, 81-88(1995)