

時間帯別人口を用いた購買確率推計に関する研究

島崎 康信・関本 義秀・柴崎 亮介

A Study on the estimation of the purchase probability using the time slot population based on person flow

Yasunobu SHIMAZAKI, Yoshihide SEKIMOTO and Ryosuke SHIBASAKI

Abstract: Recently, the need of person flow analysis arises in various aspects (situation) such as disaster prevention, marketing, traffic planning. In this paper, we investigated about the relationship between the time slot population and the purchase probability utilizing linear mixing model, for the application of area marketing based on person flow data.

Tokyo metropolitan person trip survey in 1998 was utilized as a time slot population data. Commercial statistics (1 km mesh) in 1999 was applied as data of actual purchase. Data was summarized utilizing the People Flow Analysis Platform of the University of Tokyo.

Keywords: 人の流れ (person flow), パーソントリップ調査 (person trip survey), 購買確率 (purchase probability), リニアミキシングモデル (linear mixing model)

1. はじめに

近年、防災・マーケティング・交通計画等の様々な局面で、「人の流れ」の把握の必要性が高まってきている。ある通信キャリアでは、携帯電話の通信で取得される基地局情報等を集計することにより、モバイル空間統計を作成し、まちづくりや防災計画分野での有用性を報告している¹⁾。

エリアマーケティングの分野においても、近い将来には、駅の改札利用時刻や、携帯電話ユーザーの位置情報等のデータが動的な統計情報として流通するようになり、店舗立地・販売促進等の活動も、大きく変化することが予想される。

人の流れを活用した動的情報の研究としては、パーソントリップ調査データと国勢調査、交通センサスとの関係を比較検証した事例（関本ほか、2010）や、滞在時間と商業統計との関係を比較した事例（島崎ほか、2009）等の萌芽的な研究がある。

店舗立地・販売促進等のエリアマーケティングの観点では、時間帯別人口や、平日・休日別人口の把握や、その購買確率の時間変化等が有効であると考え

えられる。しかし、現時点において、そのような分析に有効であり、質的・量的に整理された統計情報を入手することは困難である。

そこで本稿では、時間帯別人口と購買確率の関係を分析し、その関係性を検討することを目的とする。時間帯別の購買確率の推計には、リニアミキシングモデルを応用した分析モデルを提案する。

人の流れデータには、東京都市圏交通協議会の平成10年度東京都市圏パーソントリップ（以下、PT）調査のマスタデータを使用し、東京大学 空間情報科学研究センターの動線解析プラットフォーム²⁾を利用して、人の流れの断面データである時間帯別人口を集計した。この時間帯別人口に対して、本稿で提案するリニアミキシングモデルで時間帯別の購買確率を推計し、時間帯別人口と購買確率の関係性を分析する。購買結果データには、平成11年度商業統計調査のコンビニエンスストア年商品販売額の1kmメッシュ集計を用いる。

また、分析モデルの検証として、コンビニエンスストア利用状況のアンケート調査結果と比較することで、分析モデルの有効性も検討する。

2. 分析モデル

本稿で用いるリニアミキシングモデルの基本的な考え方は、画像のピクセル内の構成要素の割合を推計する方法と類似している。（Shimabukuroほか、1991）、（Holbenほか、1993）

任意の時間帯におけるメッシュ別の推計人口は、動線解析プラットフォームにより集計できるので、1人あたりの購買単価 t 、平日の日数 d 、購買確率 p を乗算し、全ての時間帯で加算すれば、購買結果である商品販売額になることをモデル化した。

各時間帯の推計人口 x_i について、購買確率 p は0から1の間で変化するので、購買確率 p を連続的に変化させ、その条件下における誤差金額 e を算出する。

また、平日の購買確率と休日の購買確率は大きく異なることが推測される。しかし、平成10年度東京都市圏PT調査では平日の1日間を調査対象としているので、動線解析プラットフォームで取得可能な時間帯別の推計人口も平日に限られる。このため、本モデルでは、商品販売額 r に平日の購買確率 b を乗じることとし、平日の購買確率 b も連続的に変化させるようにした。つまり、購買確率 p と平日の購買確率 b を同時に連続的に変化させ、その条件下における誤差金額 e の二乗が最小になる購買確率 p と平日の購買確率 b の値を本モデルの最適解として採用する。

$$b \cdot r = \sum_{i=1}^n (p_i \cdot t \cdot d \cdot x_i) + e \quad (1a)$$

$$F = e^2 = \left(b \cdot r - \sum_{i=1}^n (p_i \cdot t \cdot d \cdot x_i) \right)^2 \quad (1b)$$

- b : 平日の購買確率 ($0 \leq b \leq 1$)
- r : 商業統計で1ヶ月当りの商品販売額[万円]
- i : 時間帯の区分 ($i=1, 2, 3 \dots$)
- p_i : 時間帯 i における購買確率 ($0 \leq p \leq 1$)
- t : 1人当りの購買単価[万円]
- d : 平日の日数 (20日を設定)
- x_i : 時間帯 i における推計人口 (生産年齢人口)
- e : 誤差金額[万円]

時間帯の区分 i を細かくするか、推計人口 x_i を性別・年齢階級別の属性別にすることで、より精度の高い分析が可能と考えられるが、同時に本モデルでの計算量が級数的に増加してしまう。そこで本稿では、時間帯および推計人口の属性について、3.1.に記載の通りとする。

3. 分析データの処理

3.1. PT 調査データ

平成10年東京都市圏PT調査データ（東京都市圏交通計画協議会）は、1998年10～12月の土・日・祭日及び月・金を除く平日の1日間を調査時点としている。都市圏全体の5歳以上の人口32,896,705人に対するPT調査の有効サンプル数は、883,044人なので、標本率は約2.68%となっているが、拡大係数を用いて全人口を推計することにする。

動線解析プラットフォームでは1分間隔でのデータ取得が可能だが、本稿では、検証用のコンビニエンスストア利用状況のアンケート調査を勘案し、7:30、10:30、13:30、16:30、19:30、22:30の6つの時間帯とする。また、推計人口を属性別に分類することはせずに、15歳以上65歳未満の生産年齢を対象に人口を1kmメッシュで集計を行った。

3.2. 商業統計調査データ

本稿では、PT調査時点と最も近い平成11年商業統計調査データの年間商品販売額1kmメッシュを用いる。分析対象の項目は「コンビニエンスストア年販売額」とした。

3.3. コンビニエンスストア調査データ

1人当りの購買単価として、本稿では、(社)フランチャイズチェーン協会のコンビニエンスストア統計調査月報の平均客単価（既存店ベース）579円を採用する³⁾。

また、分析モデルで推計される時間帯別の購買確率は、(社)中央調査社の全国成人のコンビニエンスストア利用状況（2001年5月）と比較する⁴⁾。

4. 分析モデルの結果

本モデルで推計されたコンビニエンスストアの平日の購買確率の地域分布を示したのが、図-1である。1ヶ月（30日間）の内、平日は20日と仮定したので、平日と休日の購買確率の差異が全くないと仮定すると、平日の購買確率は20/30（67%）となる。そのため図-1の中で、平日の購買確率が67%より高い地域は、平日の購買確率が高く、休日の購買確率が低い地域と判読できる。逆に、平日の購買確率が67%より低い地域は、平日の購買確率が低く、休日の購買確率が高い地域と判読できる。休日の集客が多い新宿駅・渋谷駅等の主要駅周辺では、平日の購買確率が低く推計されていることが判読できる。平日のデータしか入手できない場合でも、平日と休日の購買確率の差異を推計できることは、本モデルの特徴である。

平日の時間帯（7:30、10:30、13:30、16:30、19:30、22:30）別の購買確率が推計された中で、19:30時点の購買確率を地図化したものが図-2である。19:30時点では、オフィス街・山手線内の地域を除くような郊外で購買確率が高く、平日の購買行動をある程度表現していると思われる。但し、推計された購買確率が高めにしているのは、各時間帯（例 6:00～9:00）の推計人口を一つの時間断面（例 7:30）で表現しているため、実際の時間帯に出現する人数を過少に推計しているためだと考えられる。

図-3は、霞ヶ関の1kmメッシュの推計人口（青線）と購買確率（赤線）の時間帯別変化である。推計人口の時間変化と購買確率の時間変化は、同調して変化する訳ではなく、時間帯（例 19:30）によっては推計人口が少なくても購買確率が高くなっていることが判読できる。また、霞ヶ関を含む1kmメッシュの平日の購買確率は85%であったので、休日の人口が少なく購買確率も低くなっていると推測され、現実の状況を表現できていると思われる。

5. コンビニ利用状況アンケートとの比較検証

本稿で提案している分析モデルで推計された平日の時間帯別の購買確率の妥当性を検証するため、

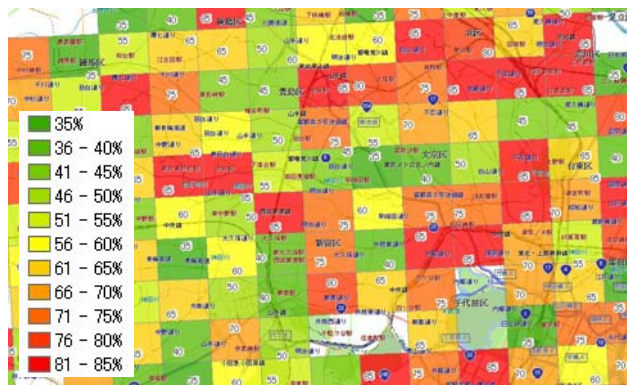


図-1：平日の購買確率[%]（コンビニエンスストア）

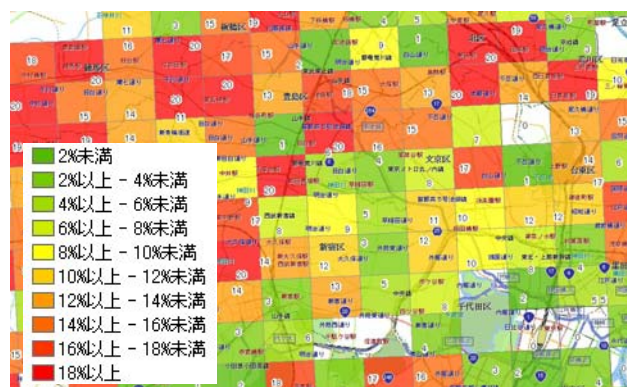


図-2：19:30 時点における購買確率[%]（コンビニエンスストア）

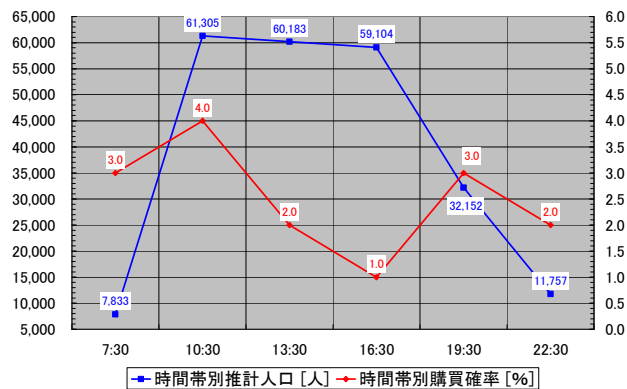


図-3：推計人口と購買確率の時間変化（霞ヶ関）

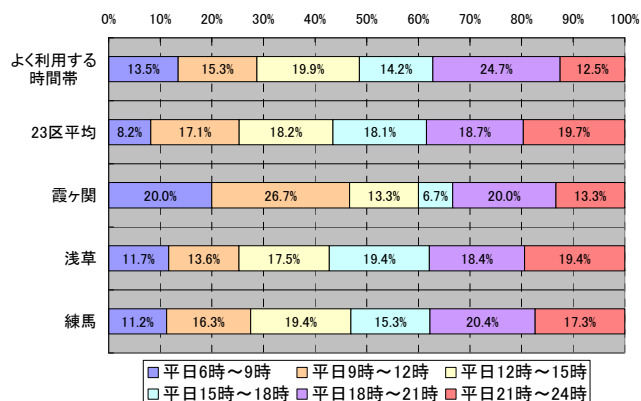


図-4：時間帯別の購買確率の構成比 [%]

本稿では、(社)中央調査社の全国成人のコンビニエンスストア利用状況(2001年5月)のアンケート調査結果と比較することにする。アンケート調査地域は全国ではあるが、調査時期が2001年3月で、有効回答数が1,406人なので、平均的な利用状況を把握する上では最適だと思われる。アンケートの中で、コンビニエンスストアをよく利用する時間帯を聞いている(複数回答可)が、この結果を構成比に換算して図化したものが、図-4の最上段である。

分析モデルで推計された時間帯別の購買確率についても同様に、構成比に換算した値を図-4の2段目以降に示した。23区平均は、分析対象とした23区の112メッシュ⁵⁾の平均値である。アンケート調査結果と比較すると、分析モデルの結果の方が、平日6時~9時と18時~21時の構成比が低くなっているが、概ねアンケート調査と近い結果となった。浅草、練馬の1kmメッシュについては、分析モデルで推計された時間帯別の購買確率(構成比に換算)とアンケート結果は、かなり似た構成比となり、全国平均と近い傾向があることが判読できる。但し、霞ヶ関の1kmメッシュは、アンケート結果(構成比に換算)と違いもあるので、地域性による差異か分析モデルの問題であるのかをより詳細に検討する必要がある。

6. まとめと今後の課題

本稿では、動線解析プラットフォームから集計される時間帯別の推計人口と、商業統計のコンビニエンスストア年商品販売額に対して、リニアミキシングモデルを適用することで、平日の購買確率、時間帯別の購買確率を推計するモデルを提案した。

本モデルによって推計された平日の購買確率、時間帯別の購買確率は、消費者の常識的な感覚と合う部分も多くあり、コンビニエンスストアの利用状況のアンケート調査と比較しても、提案モデルの有効性を感じることができた。

しかしながら、地域的な差異を細かく精度検証するには、より詳細な実測データとの比較が必要なため、POSデータ等を用いた検証を検討していきたい。

謝辞

本稿の研究は、東京大学空間情報科学研究センター寄附研究部門「空間情報社会研究イニシアティブ」の一環として行われた。ここに記して謝意を申し上げます。

注

1) NTT ドコモ モバイル空間統計の報道発表

(http://www.nttdocomo.co.jp/info/news_release/2011/05/24_01.html)

2) 東京大学 空間情報科学研究センター 人の流れ

プロジェクト (<http://pflow.csis.u-tokyo.ac.jp>)

3) (社)日本フランチャイズチェーン協会 コンビニエンスストア統計調査月報の2005年10月が、PT調査月が同じで、最も年次に近いデータであった。

(<http://www.jfa-fc.or.jp/>)

4) (社)中央調査社 全国成人のコンビニエンスストア利用状況、2001年5月 (<http://www.crs.or.jp/>)

5) 分析モデルの計算対象の1kmメッシュは23区内としたが、商業統計の秘匿メッシュでなく、推計された時間帯別の購買確率が0%でないメッシュを対象とした。

参考文献

関本義秀・薄井智貴・島崎康信・南佳孝・柴崎亮介(2010): パーソントリップデータを用いた時空間内挿処理と様々な統計データとの比較検証, 土木計画学研究・講演集, vol. 41 (CD-ROM)

島崎康信, 関本義秀, 柴崎亮介, 秋山祐樹(2009): 人の流れから算出される滞在時間と商業統計の関係性について, 第18回地理情報システム学会講演論文集, Vol. 18, 239-242.

Shimabukuro Y.E. and Smith J.A. (1991): The least-squares mixing models to generate fraction images derived from remote sensing multispectral data, IEEE Transactions on Geoscience and Remote Sensing, 29, 16-20
Holben B.N. and Shimabukuro Y.E. (1993): Linear mixing model applied to coarse spatial resolution data from multispectral satellite sensors, International Journal of Remote Sensing, 14, 2231-2240.