

【研究ノート】

小規模小売店舗の商圈分析のための店舗の魅力度推定手法

李 廷恩

1. はじめに

(1) 背景

マーケティング地理学は、顧客の購買行動、店舗の経営、マーケティング活動間の相互関係を地域及び空間的観点から観察し、その変化や類似性を体系化する学問である (Davis, 1976)。特に、小売市場の商圈分析は、商圈の地理的空間の構造及び顧客の購買行動を把握するのが主な目的である。小売市場は細分化されつつ、店舗間の競争も激しくなるという環境の中、よりミクロな観点からの商圈分析アプローチが必要とされる。ミクロな観点での商圈分析のためには、顧客の購買行動の分析・理解がマーケティング戦略の樹立に重要な情報となる (Kotler, 1991)。

顧客の購買行動及び店舗の選択行動を理解するための既存手法は大きく2つに分類する。1つは記述及び決定的手法 (descriptive-determinist approach) で、もう1つは経験及び確率的手法 (explicative-stochastic approach) である。前者は、顧客に対するアンケートを用いたアプローチや中心地理論に基づいた商圈の地理的パターンを分析するのにたいして、後者は、顧客の店舗選択傾向が1つの店舗だけの選択より複数の店舗を選択する可能性から、選択の確率的概念を含んでいる。

上記に述べられた詳細な商圈分析のためには、店舗や顧客に対する地理的情報はもちろん、ある店舗とその競争店舗の売上情報や顧客の好み、各店舗の魅力度など、詳細なデータが多く集まるほど、より確実な分析結果が得られる。

(2) 研究目的

小規模の小売店舗の場合、前述のような競争店舗の売上情報や各店舗の魅力度を測るための詳細情報の取得は情報収集のための時間や費用の面で極めて難しく、現実的には対象となる店舗自体の売上データだけが入手可能なケースが一般的である。今までの研究では、十分なデータが揃った状況での消費者の購買行動分析や商圈分析などを通じたマーケティング戦略樹立を目的とするものが多く、小規模小売店舗のように商圈分析に必要とされる入手データに制限がある場合を扱う手法は少ない。

このような背景から、本研究では、情報の入手に制限がある小規模の小売店舗を対象に、その売上データと競争店舗の位置情報だけが取得可能なときの商圈分析を行うための手法を提案する。すなわち、売上分布に基づき、その形成に最も大きな影響を与える競争店舗を抽出する手法の提案が研究の目的である。

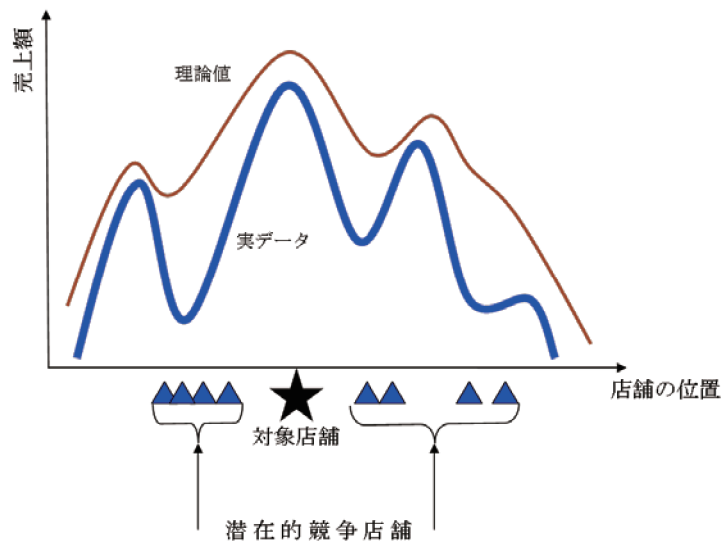


図1. 対象店舗の実際売上分布と理論値

図1の曲線は、店舗の売上分布を理論値と実データの場を分けて示している。上の曲線は理論値であり、対象店舗の中心とする領域の売上額は高く、中心から離れるほど売上額が減る分布になっている。しかし、実際の売上分布は理論値とは異なる形状を見せている。理論値と実際の分布値の差が特に大きい位置については、図1の三角記号で表されたように競争店舗の存在が重要な原因として考えられる。したがって、入手データに制限があっても、対象店舗の売上データと競争店舗の位置情報が揃えば、小規模小売市場の商圈分析は可能だといえる。

商圈分析の手法としてはロジットモデルを用いた店舗選択確率を計算する手法に基づく。ここに、顧客が店舗を選択する確率が店舗の売上額と比例すると仮定を立て、競争店舗の影響力を推定モデルの定義を行う。次に、店舗選択のモデル化及び仮想データを利用したシミュレーションを行う。その結果から、モデルの性能に影響を与える要素を把握し、それに基づき実際のデータを用いた分析を行う。実データを用いた結果から、ある店舗の売上データの分布に影響を与えると思われる競争店舗を抽出する。

2. 手法説明

(1) 店舗選択モデルの説明

本研究で用いられる手法は、選択行動が確率として説明可能な、ロジットモデル (Ben-Akiva and Lerman, 1985; Fortheringham, 1988) に基づき定義される。このモデルは複数の選択要素があるときに、その中から1つを選択する確率を求められるものであり、式1のように定義する。

$$p_i(k) = \frac{A_i \exp(-\alpha d_{ki})}{\sum_j A_j \exp(-\alpha d_{kj})} \quad (1)$$

$$0 \leq p_i(k) \leq 1, \sum_{i \in C_n} p_i(k) = 1, j \in C_n$$

式1で、 $p_i(k)$ は位置 k に存在する顧客が店舗 i を選択する確率を示す。 C_n が商圈内に位置する選択可能な店舗のリストで、 n は選択肢の数を、 j はその番号を示す。また、 A_i は店舗 i に対する魅力度、 α は距離に対する常数として定義される。式1の意味を説明すると、店舗選択の確率 $p_i(k)$ は店舗 i に対する魅力度である A_i には比例し、顧客 k と店舗 i 間の距離 d_{ki} には反比例する。

一般的に、実際のマーケティングで、店舗の魅力度は店舗のマーケティング戦略や知名度、顧客の好みなどといった様々な要素で構成されている。しかし、小規模の小売店舗には、このような詳細な情報の入手はなかなか難しいことである。そのため、本研究では、店舗の魅力度 A_i を店舗が持つと思われる魅力を相対的指標として扱うことにする。

(2) 提案手法

本研究で、実際利用可能なデータは、ある店舗の売上分布とその店舗の競争店舗に対する位置情報であることから、式1をそのままデータに適用するのは不可能であり、未知パラメータの推定をするためには、モデルの推定が必要である。

入手データの制限から未知パラメータ推定の必要性があり、式2のように式1に定義されている、店舗選択モデルを変更した。式2の左辺、 W_k は空間単位 k に対する店舗 i の実売上を示すものである。右辺には、モデルにより推定される未知パラメータが含まれている。ここで、空間単位 k は、商圈をメッシュ表現した際のメッシュを示す。式2の n_k は、空間単位 k に対する顧客数であり、その値は例えば、特定年齢代人口に比例すると言う仮定で、人口データから推定される。また、 M は顧客の平均購入額を示すもので、その値は対象店舗の売上額と人口データから得られる。

店舗 i を対象店舗として定義すると、式3の $P_i(k)$ 空間単位 k に対する対象店舗の実売上であり、その値は対象店舗の売上データから計算可能である。これは、式1の店舗選択モデルの確率とは同じ値ではないが、店舗の選択行動確率は、その店舗の売上額に比例すると定義することにより、店舗の魅力度のような未知パラメータの推定が可能になる。

$$W_k = Mn_k \frac{A_i \exp(-\alpha d_{ki})}{\sum_j A_j \exp(-\alpha d_{kj})} \quad (2)$$

$$s.t. \sum_{j=1}^N A_j = 1, \forall A_j \geq 0, i \in j$$

$$E = \min \sum_k \{P_i(k) - W_k\}^2 \quad (3)$$

結果的に、式2を用いることにより、式3の条件を満たす、未知パラメータの推定が可能になる。すなわち、モデルの推定は、 E の値を最小とする式3を解くことに

なる。これによって、未知パラメータである店舗 j の魅力度を表す $A_j(j=1, \dots, N)$ と距離常数 α の推定値が計算される。式3に定義されているように、これらの未知パラメータは $P_i(k)$ と W_k の最小二乗により値を得られる。

未知パラメータ、 A_j の推定により、対象店舗の競争店舗のうちどれが最も影響力の大きい店舗かがわかる。つまり、 A_j の推定値が大きいと競争店舗の魅力度も大きいことになる。

モデル定義の次は、仮想データを利用し、店舗選択行動シミュレーションを行う。その次に、実際データをモデルに適用する。仮想データの利用の際は、競争店舗数が少数（2店舗）のケースと多数（9店舗）の二段階に分け、小数店舗の時は空間的要素が店舗の魅力度にどのような影響を及ぼすのかを見る。多数店舗のケースでは、店舗間の魅力度の順位を推定し、仮想データによる状況の説明が正確にできるかを見る。実際データのモデルへの適用では、その結果がデータの表示による商圈状況を適切に説明しているかを確認する。

3. 手法の適用

提案された手法の検証のため、まずは、仮想データを用い、モデルの性能に影響があると思われる要素を分析するなど多様な観点からの実験を行う。次に、その結果に基づき、実データをモデルに適用する。モデルの性能評価は、計算結果が実際のデータから記述される商圈の状況をどれだけ説明できるかで判断される。

(1) 店舗選択行動シミュレーション

ここでは、仮想データを用いた店舗選択行動シミュレーションを行う。図2には、そのフローが記述されている。まずは、シミュレーションに必要な仮想データを作成する。

必要な仮想データは、対象店舗の売上分布と対象店舗及び競争店舗の位置情報である。特に、店舗の位置情報を作成するためには、対象店舗を中心に、店舗間の位置属性として「方向」と「距離」のオプションを設定する。方向の場合、対象店舗を商圈の中心に置き、2等分した時、競争店舗が同一の区域に位置するとS (Same)、反対の区域に位置するとO (Opposite) の記号でシミュレーションケースを区分する。これは、店舗間の配置

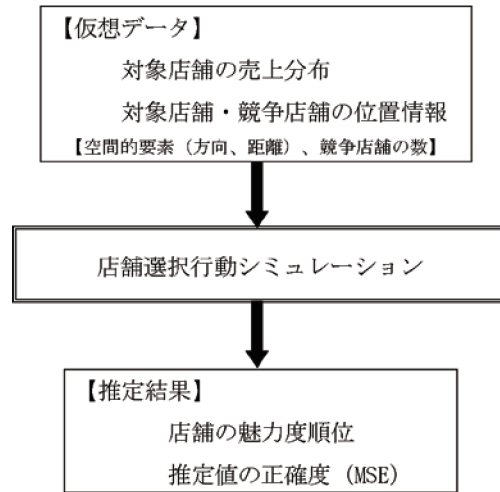


図2. 店舗選択行動シミュレーションのフロー

を偏らないようにするためである。距離については、対象店舗と競争店舗間の距離をN (Near)、M (Moderate)、F (Far) といった3つのレベルに分ける。これらは、相対的な基準であり、Nを基準とし、その2倍の距離をM、3倍の距離をFと定義する。また、競争店舗の数は「少数（2店舗）」と「多数（9店舗）」のケースに分ける。競争店舗数が2の場合、位置属性である「方向」と「距離」がモデル性能に与える影響の理解が目的である。それを元に、影響力が小さい属性は、競争店舗数が9の場合では考慮しないこととする。ここでの店舗数の指定は、入手したデータに基づき定めた値である。

これらの位置属性に関するオプションの詳細は図3に示してある。例えば、店舗1と2の方向オプションは、点線で区分された同一の領域に位置するので「S」になる。店舗1と3あるいは、2と3の場合は点線を基準に

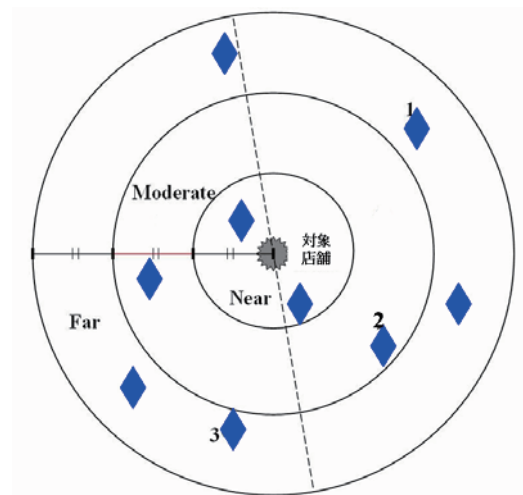


図3. 店舗選択行動シミュレーションの空間的要素

異なる領域に位置するので、方向オプションは「O」となる。

作成された仮想データを用い、店舗選択行動シミュレーションを行い、次に、その結果を持って、提案手法の評価を行う。その時の評価指標は、表1に示されているように、店舗の推定魅力度の順位正確度（全店舗対象、競争店舗対象）と店舗の推定魅力度の平均MSE（二乗誤差）の3つがある。これらの評価指標は、推定計算結果のうち、式2の、店舗の魅力度 A_j の値が0以上で、合計が1という条件を満たす「許容解」と呼ばれるものを対象とする。

1番目に、「全店舗対象の店舗の推定魅力度の順位正確度」は、すべての店舗に対する推定順位がシミュレーションの入力値による店舗の順位と一致するケースをカウントする。2番目の「競争店舗対象の店舗の推定魅力度の順位正確度」に対しては、仮想データの競争店舗順位の上下関係が一致するケースをカウントする。3番目の「店舗の推定魅力度の平均MSE (Mean Square Error)」は、実際の入力した店舗魅力度と推定結果値の間の誤差を平均化したものである。これは、誤差の値が小さいほどより正確な推定結果として判断する。

表1は、少数の店舗数（競争店舗数：2）に対する店舗選択行動シミュレーションの結果を示したものである。ケース_IDは、「-」記号を中心に左右に数値があり、左の数値は競争店舗の数を、右の数値は方向及び距離オプションに対するケース番号の意味を持つ。ケースごとに100回のシミュレーションを行い、その中、式2の店

舗の推定魅力度 A_j に対する条件を満たす許容解がいくつあったかを表1の「許容解の総数」に示している。許容解のうち、店舗の推定魅力度による順位が対象店舗を含む、すべての店舗に対して一致する場合を「店舗の推定魅力度の順位正確度」の「全店舗対象」に示しており、「競争店舗対象」の数値は、許容解のうち、全競争店舗の推定魅力度による順位の上下関係が一致する場合に該当する。

順位推定の正確度から見ると、2つの競争店舗が同一の商圈区域に存在し、距離オプションがNである場合は、他のケースより多い正確度（全店舗の推定順位）を見せた。それ以外には、商圈区域の区分に対するオプションより、対象店舗からの距離がMかFの場合、競争店舗対象の推定結果がより多い正確度を見せた。具体的に、ケース_ID、2-7、2-8、2-10から2-12が推定の正確度が高い結果である。これらの平均MSE値は0.1以下であり、他のケースより小さい誤差を示している。

これらの結果により、提案されたモデルでは、競争店舗の商圈内配置区域には影響されず、対象店舗との距離が両方ともMかFの場合に推定正確度が上がることがわかった。

表2の内容は、店舗数が多数の場合（競争店舗数：9）の店舗選択行動シミュレーション結果である。まず、ケース_IDの9M-1の場合を例として説明すると、9は競争店舗数を、Mは全競争店舗に対する配置の距離オプションを示す。また、「-」右辺の1は、各店舗の魅力度の入力値パターンを意味する。ここでは、少数店舗の

表1. 店舗選択行動シミュレーション結果（競争店舗数：2）

ケース_ID	距離 & 方向オプション	店舗の推定魅力度の順位正確度			平均 MSE
		全店舗対象	競争店舗対象	許容解の総数	
2-1	N/N/S	30	9	55	0.089
2-2	N/N/O	10	16	52	0.246
2-3	N/M/S	13	9	35	0.246
2-4	N/M/O	21	14	53	0.129
2-5	N/F/S	12	12	32	0.289
2-6	N/F/O	23	17	67	0.131
2-7	M/M/S	23	28	97	0.070
2-8	M/M/O	22	30	100	0.075
2-9	M/F/S	18	17	54	0.129
2-10	M/F/O	20	28	89	0.083
2-11	F/F/S	22	32	99	0.070
2-12	F/F/O	20	29	99	0.069

* 距離のオプション：N (Near)、M (Moderate)、F (Far)
方向のオプション：S (Same)、O (Opposite)

表 2. 店舗選択行動シミュレーション結果 (競争店舗数: 9)

ケース_ID	店舗の推定魅力度の順位正確度			平均MSE
	(1)	(2)	許容解の総数	
9M-1	4	1	19	0.106
9M-2	4	2	19	0.098
9F-1	16	9	52	0.026
9F-2	11	7	42	0.033

* 距離のオプション: M (Moderate)、F (Far)
 「店舗の推定魅力度の順位正確度」(1): 上位3店舗の順位が昇順の時
 (2): 上位3店舗の順位が1、2、3のいずれである時

ケースに適用された商圈区域のオプションや対象店舗との距離オプションのばらつきは考慮されていない。その理由としては、第一に、表1の結果から位置属性のうち、方向オプションは計算結果にあまり影響がないことである。第二に、店舗数の増加によりケースの数は幾何級数的に増えてしまい、それを全部計算するのは無意味である。したがって、ここでは、距離オプションが全ての競争店舗に対し、MかFか、また、店舗魅力度の入力値パターンを2種類にすることにより、全部4つのシミュレーションケースを作成した。

結果を見ると、許容解の総数や上位3店舗の順位正確度、平均MSEなど対象店舗からの距離がFのケースに提案モデルがより正確な結果を出したことがわかる。平均MSEの値から見ると、距離オプションがMの場合は平均MSE値が約0.1であるのに対し、Fの場合は、約0.03である。これは、対象店舗に対する競争店舗間の距離がMかFの場合、提案モデルを利用し、影響力の高い競争店舗の抽出が可能であることを意味する。

(2) 実際データの適用

ここでは、モデルを実際のデータに適用し、対象店舗の有力な競争店舗を抽出するための過程及びその結果について説明する。図4は実際のデータを入手した対象店舗の位置とその商圈を示している。商圈は東京都内のある地域であり、対象店舗は出前そばやである。白線の内側が研究対象地域(対象店舗の商圈)である。その中に丸で表されているのが対象店舗である。

入手したデータには、顧客の名前と位置(住所)や注文時間帯、注文金額などを含んでいる。データの収集対象となる期間は1月7~

31日で、顧客の数は234である。各顧客に対し複数回の注文履歴が存在し、1日に3つの営業時間帯(11:00~14:00、14:00~18:00、18:00~21:00)に分けて注文内容が記録されている。

提案手法を用い、対象店舗に最も影響力のある競争店舗を抽出するためには、上記に記述したようなテキスト形式の元データを空間データに変換する必要がある。例えば、顧客の位置情報は住所から座標に変換し、その位置情報を50mのメッシュ上にマッチングさせ、その位置に該当する売上金額を各メッシュの属性として割当てる。各メッシュに対する人口1人当たりの売上金額は、メッシュの売上総額を商圈内の人口分布が丁丁目単位で提示されているのを利用し、それが丁丁目内では均等に分布されているという仮定で計算した人口で割った数値を表示している。また、競争店舗の位置情報は、その住所から座標情報に変換し、点データとして対象店舗の商

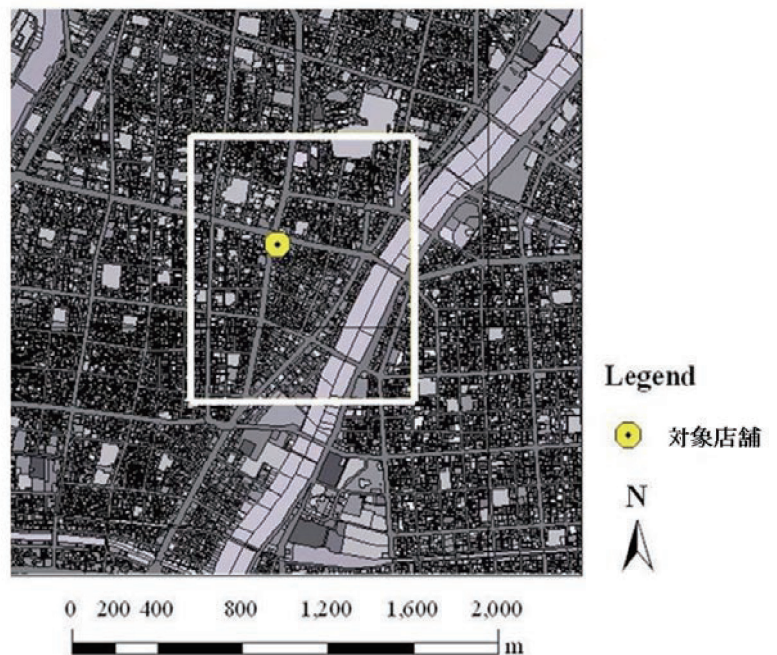
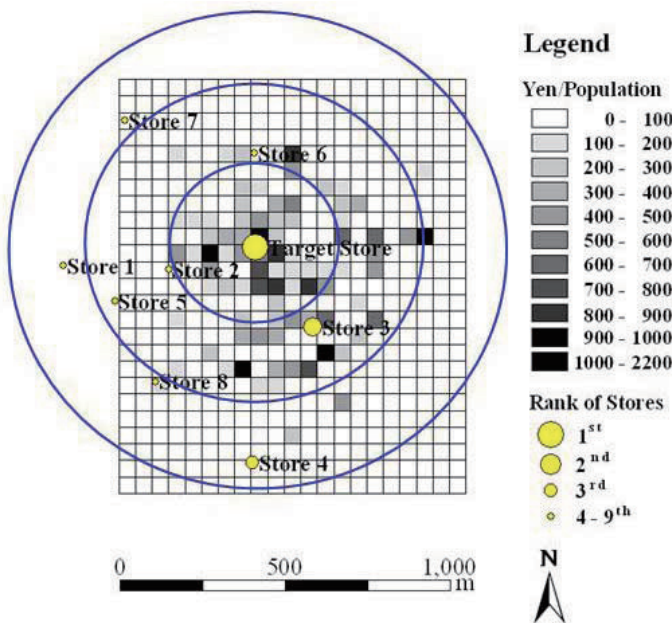


図 4. 対象地域及び対象店舗の位置

圏の地図上に明示できる。その結果、加工されたデータは、図5のように、店舗の位置は点として、対象店舗の売上分布は売上額別に色付けされたメッシュに示されている。

提案手法を用い、対象店舗及び競争店舗の魅力度を推定した結果は、図5に大きさが異なる点で表示されている。魅力度の順位が一番高いのは、対象店舗であり、その次に順位が高いのは店舗3、4である。シミュレーションの際も上位3つの店舗だけを考慮したので、実際のデータでも同じく上位3位の店舗だけを見る。推定結果と対象店舗の売上分布の形状を比較してみると、売上分布値が部分的に急減する位置に影響力が大きい競争店舗(図5のStore 3、4)が立地していることがわかる。

モデルの性能評価の要素で対象店舗と競争店舗との距離オプションから、モデルの正確度(平均MSE)が高かったケースは競争店舗がMあるいはFの距離に立地したときであった。実際、商圈を対象店舗を中心とする3つの距離区域(N、M、F)で分けたとき、有力の競争店舗がM、Fの距離区域に入っていることがわかった。これにより、対象店舗の売上データと競争店舗の位置情報だけを用いた各店舗の魅力度の推定である、実データの提案モデルへの適用結果から、影響力の高い競争店舗を抽出した。これは、小規模小売市場の対象店舗に対する商圈の状況説明を可能にした。



* 店舗の順位は標識が大きいほど上位順位である。

図5. 対象店舗の売上分布と推定魅力度による店舗の順位

4. 結論

本研究では、マーケティング地理学で重要とされる、商圈分析について、小規模の小売店舗に焦点を当てた。現実的に小規模小売店舗が入手可能なデータは商圈内の競争店舗の位置情報と対象店舗の売上データが一般的であり、これらのデータだけで商圈分析を行うために、ロジットモデルを用い、未知パラメータの推定をする手法を提案した。用いた手法では、データの入手が困難な小規模小売店舗の商圈分析において、店舗の選択確率が対象店舗の売上に比例するという仮定を立てロジットモデルを用い、未知パラメータである店舗の魅力度を推定した。

提案された手法の性能評価のために、まず、仮想データで店舗選択行動シミュレーションを行った結果、モデルの推定正確度が高まる要素は、対象店舗から競争店舗間の距離(オプションがMあるいはF)であることがわかった。次に、実データを用いモデルに適用し、レコード数は約230個で少なめではあるが、推定した競争店舗の魅力度は対象店舗の売上分布の状況を有効に説明することが可能であった。

今後の研究拡張の可能性としては、入手データに関しては、より長期間のデータを収集し分析を行う方法も考えられる。また、店舗選択行動のシミュレーションの際に、モデル性能に影響があるような競争店舗配置に関する空間的要素の提案も1つの課題である。

References

Ben-Akiva, M. and Lerman, S. R. (1985) Discrete Choice Analysis, The MIT Press.

Davies, R. L. (1976) Marketing Geography : with special reference to retailing, Methuen & Co. Ltd., 223-229.

Fortheringham, A. S. (1988) Consumer Store Choice and Choice Set Definition, Marketing Science Journal, 7 (3), 299-310.

Kotler, P. (1991) Marketing Management : Analysis, Planning, Implementation and Control, 7th Edition, Prentice Hall.

[いー じょんうん]
 [(財) 土地総合研究所 調査部 研究員]