RF-008

整数計画法による変数選択を用いた店舗選択モデル

佐藤俊樹 † 高野祐一 ‡ 中原孝信 ♮

† 筑波大学大学院システム情報工学研究科 ‡ 専修大学ネットワーク情報学部 ♮ 専修大学商学部

1 はじめに

現代の消費者は，最寄り品の購買に限ってもスーパーマーケット，総合スーパー，百貨店，コンビニエンスストア，ドラッグストアなど，様々な小売業態において購買機会を有する．また，その中でも同じカテゴリーの商品を扱う業態が多数存在するため，消費者にとって店舗のスイッチング・コストは低くなっている．このような状況で，小売店が消費者の店舗選択をはじめとする購買行動を把握することは，他店との差別化などの戦略を検討する上で重要な指針となるであろう．

消費者の店舗選択行動に関する研究は数多く存在し，それらの観点は店舗属性，店舗イメージ，ストア・ロイヤルティの三つに大きく分類することができる．消費者に店舗間の相違が認識され，その相違が消費者の店舗選択において重要である場合に，店舗は特異な存在となる．その特異化の要因となるのが店舗属性であり，論文 [5] では，食品スーパーの店舗選択の要因として，品揃え，利便性，価格，チラシの影響を評価した．論文 [11] では，商品の価格や品揃え以外の要因として，駐車料金の低さや来店に要する時間の短さ，レジの待ち時間の短さなどを考慮し，店舗選択確率との関係を示した．また，論文 [16] では，消費者が買い物に出かける際の状況を日常使いとまとめ買いに分け，店舗選択に影響を及ぼす属性との関係を示した．消費者は店舗属性を認識し評価することによって店舗イメージを形成するが，論文 [4] はその店舗イメージがストア・ロイヤルティに与える影響を検証しており，また論文 [2] では，店舗イメージとストア・ロイヤルティの因果関係を示した．そして，論文 [15] では顧客の店舗選択の要因が商品レベル，店舗レベル，個人レベルの三つに分類されている．商品レベルは，品質・価格・品揃え・チラシなどの各商品が持つ要因で，店舗レベルは，店舗の規模・利便性・レジの待ち時間・サービス品質などの要因である．そして個人レベルは，買い物時間や買い物の目的，デモグラフィックから成る要因である．

消費者の店舗選択には，このような複数レベルの要因が相互に関係していると考えられるが，本研究では，消費者の店舗選択における商品レベルの要因を明らかにする．具体的には，消費者の複数店舗における購買状況を把握できるスキャンパネルデータを対象に，ロジスティック回帰モデルと変数選択を利用して店舗選択モデルを構築する．本研究で扱う店舗選択モデルは，特定の店舗と他店における購買行動の違いをモデル化しており，各店舗を選択する際に想起される商品群を明らかにできる点が特徴である．

ロジスティック回帰モデルに対する変数選択では，赤池情報量規準（AIC）[1] やベイズ情報量規準（BIC）[18] などの適合度指標が利用されることが多い．また，変数増減法 [6]，L1 正則化 [7, 10]，メタヒューリスティクス [14, 19, 20] などの様々な変数選択手法が提案されている．しかしながら，これらの手法では適合度指標の意味で最適なモデルを選択できる保証は無い．

一方で，近年では計算機とアルゴリズムの性能向上を背景として，整数計画法を用いた変数選択手法が注目を集めている [3, 8, 9, 12, 13]．本研究では Sato et al. [17] によって提案された変数選択手法を利用する．この手法では，非線形な目的関数に区分線形近似を施すことで，変数選択問題を混合整数線形計画問題として定式化する．また，情報量規準の最小値との差が一定範囲内に収まることを，理論的に保証できるという利点もある．しかし，先行研究 [17] では構築されたモデルの予測精度が検証されておらず，このことが課題として挙げられる．

以上を考慮して，本研究の目的は以下の 2 点とする：

- 整数計画法を用いた変数選択手法 [17] を利用して店舗選択モデルを構築し，その予測精度を検証する．
- 実際のスキャンパネルデータから店舗選択要因を

Variable Selection via Integer Programming for Developing Store Choice Model
Toshiki SATO†, Yuichi TAKANO‡ and Takanobu NAKAHARA♮
‡ Graduate School of Systems and Information Engineering, University of Tsukuba
‡ School of Network and Information, Senshu University
♮ School of Commerce, Senshu University

把握する．

本論文の構成は以下のようになる．2 節では，ロジスティック回帰モデルと情報量規準について説明する．3 節では，整数計画法による変数選択問題の定式化を示す．4 節で数値実験の結果を報告し，5 節で本論文のまとめを述べる．

2 ロジスティック回帰モデルと情報量規準

各サンプル $i = 1, 2, \dots, n$ に与えられた 2 値のラベル $y_i \in \{-1, 1\}$ を，p 個の説明変数から成るベクトル $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})^\top \in \mathbb{R}^p$ によって判別する問題を考える．ロジスティック回帰モデルでは，以下のロジスティック関数を使用してラベルの生起確率をモデル化する：

$$\Pr(y \mid \boldsymbol{x}) = \frac{1}{1 + \exp(-y(\boldsymbol{w}^\top \boldsymbol{x} + b))}.$$

切片 b と偏回帰係数 $\boldsymbol{w} = (w_1, w_2, \dots, w_p)^\top$ は以下の対数尤度関数が最大となるように求める：

$$\begin{aligned}\ell(b, \boldsymbol{w}) &= \log\left(\prod_{i=1}^{n} \Pr(y_i \mid \boldsymbol{x}_i)\right) \\ &= -\sum_{i=1}^{n} \log(1 + \exp(-y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b))) \\ &= -\sum_{i=1}^{n} f(y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b)).\end{aligned}$$

ただし，$f(v) = \log(1 + \exp(-v))$ とし，この関数はロジスティック損失関数と呼ばれる．この関数 f は 2 階微分が正であるため凸関数である．

多くの情報量規準は，選択した変数の集合を $S \subseteq \{1, 2, \dots, p\}$ とし，それ以外の変数の偏回帰係数の値を 0 にすることで，

$$-2 \max\{\ell(b, \boldsymbol{w}) \mid w_j = 0 \ (j \notin S)\} + F(|S| + 1) \quad (1)$$

と表される．F は選択した変数の数 $|S|$ に対するペナルティの大きさを表し，$F = 2$ の場合は AIC [1] に，$F = \log n$ の場合は BIC [18] に対応する．これらの情報量規準の値が小さい変数集合 S が望ましいとされる．

3 整数計画法による変数選択

0-1 決定変数 $\boldsymbol{z} = (z_1, z_2, \dots, z_p)^\top$ を導入し，$j \in S$ のとき $z_j = 1$，それ以外のとき $z_j = 0$ とする．これにより，情報量規準 (1) を最小化する変数集合 S を決定する問題は次のように書ける：

$$\min_{b, \boldsymbol{w}, \boldsymbol{z}} \ 2 \sum_{i=1}^{n} f(y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b)) + F\left(\sum_{j=1}^{p} z_j + 1\right) \quad (2)$$

$$\text{s.t.}\ z_j = 0 \Rightarrow w_j = 0 \quad (j = 1, 2, \dots, p), \quad (3)$$

$$z_j \in \{0, 1\} \quad (j = 1, 2, \dots, p). \quad (4)$$

制約 (3) は，タイプ 1 の特殊順序集合（SOS1）制約による表現が可能である．SOS1 制約は，集合の要素のうち非ゼロの値をとるのは高々1 つであることを意味する．よって，$j = 1, 2, \dots, p$ に対して集合 $\{1 - z_j, w_j\}$ に SOS1 制約を課せばよい．SOS1 制約は標準的な整数計画ソルバーに実装されている．

問題 (2)–(4) の目的関数は非線形であるため，このまま解くことは難しい．そこで，Sato et al. [17] では，目的関数に区分線形近似を施し，混合整数線形計画問題に変形する方法が提案されている．

点集合 $V = \{v_1, v_2, \dots, v_m\}$ を与え，点 v_k におけるロジスティック損失関数の接線を次で表す：

$$h(v; v_k) = f'(v_k)(v - v_k) + f(v_k) \quad (k = 1, 2, \dots, m).$$

このとき，ロジスティック損失関数は接線集合によって次のように下側から近似することができる：

$$\begin{aligned}f(v) &\approx \max\{h(v; v_k) \mid k = 1, 2, \dots, m\} \\ &= \min\{t \mid t \geq h(v; v_k) \ (k = 1, 2, \dots, m)\}.\end{aligned}$$

たとえば $v_1 = -\infty, v_2 = -2, v_3 = 0, v_4 = 2, v_5 = \infty$ としたときの接線集合は図 1 のようになる．ただし，v_1, v_5 について以下が成立することに注意されたい：

$$h(v; v_1) = -v, \ h(v; v_5) = 0.$$

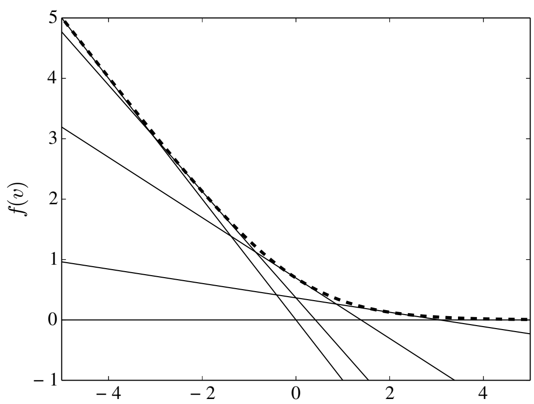

図 1: ロジスティック損失関数の区分線形近似

決定変数 $\boldsymbol{t} = (t_1, t_2, \ldots, t_n)^\top$ を導入し，ロジスティック損失関数に区分線形近似を施すと，問題 (2)–(4) は次のように変形できる：

$$\min_{\boldsymbol{b},\boldsymbol{t},\boldsymbol{w},\boldsymbol{z}} \ 2\sum_{i=1}^{n} t_i + F\left(\sum_{j=1}^{p} z_j + 1\right) \tag{5}$$

$$\text{s. t.} \ \ t_i \geq f'(v_k)(y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) - v_k) + f(v_k)$$

$$(i = 1, 2, \ldots, n; k = 1, 2, \ldots, m), \tag{6}$$

$$z_j = 0 \Rightarrow w_j = 0 \quad (j = 1, 2, \ldots, p), \tag{7}$$

$$z_j \in \{0, 1\} \quad (j = 1, 2, \ldots, p). \tag{8}$$

これは混合整数線形計画問題であり，標準的な整数計画ソルバーで扱うことができる．

4 数値実験

本研究では，株式会社マクロミルより提供されたホームスキャン方式のスキャンパネルデータを利用する．このデータは約 6,500 人のモニタによる 2012 年の 1 年間の購買履歴である．ホームスキャン方式であるため，各モニタの複数店舗における購買行動が記録されている．

モニタには事前に 5 件法によるアンケート調査が行なわれており，次の二つの質問項目の平均点が 3.5 以上のモニタを分析対象とした：

- 事前にチラシやネットでお買い得商品を確認する．
- 多少遠くても，品揃えが豊富なお店やお買い得商品の多いお店を選ぶ．

分析対象とした顧客は価格や品揃えを重視しており，店舗のスイッチング・コストが低いと考えられる．

スキャンパネルデータから分析対象とする 3 店舗を選択し，表 1 の 3 種類のデータセットを作成した．ここで，n はサンプル数，p は候補となる説明変数の数を表す．各データセットでは，対象とする 2 店舗の両方で購買履歴を持つ顧客に対し，1 回の来店を 1 サンプルとした．候補変数は，表 2 に示す JICFS 分類コードの大分類と中分類ごとの購買点数とし，すべてのサンプルに対して購買点数が 0 の変数はデータセットから除いた．例えばデータセット 7vsAE では，セブン&アイとイオンの両店舗で購買履歴を持つ顧客に対して，商品分類ごとの購買点数から，どちらの店舗で購入したかを判別するモデルを作成する．

4.1 予測精度の検証

5 分割交差検証によって，以下の 4 つの手法の予測精度を検証した：

表 1: データセット

略称	店舗ラベル		n	p
	正例	負例		
7vsAE	セブン&アイ	イオン	5,613	26
7vsSY	セブン&アイ	西友	4,557	23
AEvsSY	イオン	西友	2,987	24

FM 候補変数をすべて用いたモデル（フルモデル）．

SW$_{\text{const}}$ 説明変数無しの（切片のみの）モデルを初期解とした変数増減法．

SW$_{\text{all}}$ フルモデルを初期解とした変数増減法．

MIP 整数計画法による変数選択手法（問題 (5)–(8)）．接線を定義する点集合は，先行研究 [17] により有効とされている以下の点集合を用いた：

$$V = \{0, \pm 0.44, \pm 0.89, \pm 1.37, \pm 1.90, \pm 2.63, \pm 3.55, \pm 5.16, \pm\infty\}.$$

変数増減法はデータ解析ソフト R[1] 3.1.2 の step 関数を利用し，問題 (5)–(8) の求解には数理計画ソルバー Gurobi Optimizer[2] 6.0.0 を使用した．予測精度指標として，ROC 曲線下の面積である AUC を採用した．AUC は 0 以上 1 以下の値をとり，値が大きいほど予測精度が良いとされる．

AIC・BIC を基準とした変数選択の 5 分割交差検証の結果を表 3，4 に示す．ここで，AIC・BIC は 5 分割交差検証の訓練データにおける平均値とし，訓練データへの適合度を表す．AUC は 5 分割交差検証の検証データにおける平均値とし，予測精度を表す．各データセットにおいて，最良の AIC，BIC，AUC の値を太字で表した．

まず，表 3 について考察する．すべてのデータセットにおいて MIP による AIC 値が最も小さく，整数計画法によって AIC 値の良い変数集合を選択できていることがわかる．3 種類のデータセットの AUC 値の平均を見ると，MIP の AUC 値の平均が最も高く，整数計画法によって予測精度の高いモデルを構築することができたと言える．

次に表 4 について考察する．3 種類のデータセットの BIC 値の平均は MIP が最も良いが，データセット 7vsAE において MIP は SW$_{\text{const}}$ よりも BIC 値が悪かった．この原因としては，ロジスティック損失関

[1] http://www.R-project.org
[2] http://www.gurobi.com

表 2: JICFS 分類コードの大分類と中分類

大分類（5 種類）	中分類（29 種類）
食品	加工食品，生鮮食品，菓子，飲料・酒類，その他食品
日用品	日用雑貨，医薬品，化粧品，家庭用品，DIY 用品，ペット用品，その他日用品
文化用品	文具・事務用品・情報文具，玩具，書籍，楽器・音響ソフト，情報機器，その他文化用品
耐久消費財	家具，車両用品，時計・メガネ，光学・写真関連品，家電，その他耐久消費財
衣料・身の回り・スポーツ用品	衣料・衣服，寝具・装飾品，身の回り品，靴・履物，スポーツ用品

表 3: 5 分割交差検証の結果（AIC 基準）

	手法	AIC	AUC
7vsAE	FM	6021.18	0.6006
	SW_{const}	6008.53	0.6014
	SW_{all}	6010.50	0.6002
	MIP	**6008.14**	**0.6019**
7vsSY	FM	4750.96	0.6595
	SW_{const}	4720.75	**0.6694**
	SW_{all}	4724.78	0.6688
	MIP	**4719.76**	0.6682
AEvsSY	FM	3028.47	**0.6534**
	SW_{const}	3017.47	0.6515
	SW_{all}	3031.40	0.6523
	MIP	**3017.46**	0.6529
平均	FM	4600.21	0.6378
	SW_{const}	4582.25	0.6408
	SW_{all}	4588.89	0.6404
	MIP	**4581.79**	**0.6410**

表 4: 5 分割交差検証の結果（BIC 基準）

	手法	BIC	AUC
7vsAE	FM	6159.63	**0.6006**
	SW_{const}	**6072.92**	0.5992
	SW_{all}	6083.63	0.5992
	MIP	6077.32	0.5884
7vsSY	FM	4876.23	0.6595
	SW_{const}	4777.77	0.6576
	SW_{all}	4785.36	**0.6694**
	MIP	**4769.60**	0.6679
AEvsSY	FM	3144.95	**0.6534**
	SW_{const}	**3054.00**	0.6510
	SW_{all}	3082.56	0.6434
	MIP	**3054.00**	0.6510
平均	FM	4726.94	**0.6378**
	SW_{const}	4634.90	0.6359
	SW_{all}	4650.52	0.6373
	MIP	**4633.64**	0.6358

数を近似する接線の数が不足していたことが考えられる．AUC 値の平均を見ると，BIC 値の平均が良いほど AUC 値の平均が悪くなっており，BIC 最小化は予測精度向上の意味では逆効果だったと言える．この原因として，今回のデータセットはフルモデルの予測精度が比較的高く，変数を減らす傾向の強い BIC には相性が悪かったことが考えられる．

4.2 店舗選択要因の把握

表 5 は AIC・BIC を適合度指標として，MIP による変数選択を行なった結果である．ここでは，選択された説明変数の数と，その中で 5%有意な変数とそうではない変数の数を示している．AIC 基準の変数選択と比較して，BIC 基準の変数選択では選択される変数の数が少なくなっている．また，AIC 基準の変数選択では，5%有意ではない変数も選択されているが，BIC 基準の変数選択では 5%有意な変数のみが選択されている．

次に，具体的な店舗選択要因について分析する．AIC 基準の MIP によって選択された変数を表 6 に示す．ここでは，大分類・中分類に属する説明変数をそれぞれ（大）・（中）で表している．

まず，セブン&アイに着目する．データセット 7vsAE，7vsSY において，偏回帰係数が正値の変数はセブン&アイを選択する要因であり，偏回帰係数が負値の変数は他店を選択する要因であると考えられる．セブン&ア

表 5: 整数計画法により選択された説明変数の数

	AIC 基準			BIC 基準		
	7vsAE	7vsSY	AEvsSY	7vsAE	7vsSY	AEvsSY
候補変数	26	23	24	26	23	24
選択された変数	13	9	8	8	6	4
5%有意	8	7	5	8	6	4
5%有意ではない	5	2	3	0	0	0

イは低価格戦略をとらないコンビニ展開を中心としている．データセット 7vsAE，7vsSY では「(中）飲料・酒類」の偏回帰係数が負値となっており，今回分析対象としたような価格に敏感な顧客は，飲料や酒類をセブン&アイでは購入しにくいと考えられる．データセット 7vsAE の結果から，たばこ・健康食品などの「(中) その他食品」や，弁当・調理パンなどの「(中）加工食品」はセブン&アイで購入され,「（中）生鮮食品」はイオンで購入される傾向があることがわかり，セブン&アイのコンビニとしての特徴が見られる．

次にイオンに着目する．イオンはショッピングセンターやスーパー，ドラッグストアを抱えている企業である.「(中）生鮮食品」の偏回帰係数はデータセット 7vsAE では負値に，データセット AEvsSY では正値になっていることから，生鮮食品はイオンの選択要因であり，イオンのスーパーとしての特徴が現れている．また,「(大）耐久消費財」の偏回帰係数も同様であり，家具・カー用品などの耐久消費財全般はイオンの選択要因であり，イオンのショッピングセンターとしての特徴も見られる．また，データセット 7vsAE，AEvsSY の両方で「(中）飲料・酒類」の偏回帰係数が負値となっていることから，イオンの飲料や酒類はセブン&アイよりは強いが，西友には劣っていると言える．

最後に西友に着目する．データセット 7vsSY, AEvsSY では「（大）食品」「（中）家庭用品」の偏回帰係数が正値となっており，西友は食品全般と家庭用品が他の 2 店舗に比べて劣ると考えられる．一方で，データセット 7vsSY，AEvsSY では「(中）飲料・酒類」の偏回帰係数が負値となっている．飲料や酒類はどの店舗で買っても品質は同じであり，今回分析対象としたような価格に敏感な顧客は，毎日低価格戦略をとる西友で飲料や酒類を購入していると考えられる．

表 6: 整数計画法により選択された変数（AIC 基準）

説明変数	偏回帰係数	
7vsAE（正例：セブン&アイ，負例：イオン）		
（中）衣料・衣服	13.26	
（中）書籍	0.63	
（中）家庭用品	0.52	***
（中）その他食品	0.35	***
（中）加工食品	0.03	***
（中）飲料・酒類	−0.07	***
（中）生鮮食品	−0.20	***
（中）化粧品	−0.21	
（中）文具・事務用品・情報文具	−0.35	**
（大）耐久消費財	−1.13	**
（中）医薬品	−1.15	**
（中）DIY 用品	−2.06	
（大）衣料・身の回り・スポーツ用品	−13.26	
7vsSY（正例：セブン&アイ，負例：西友）		
（中）家庭用品	0.93	***
（中）書籍	0.69	
（大）日用品	0.15	**
（中）菓子類	0.10	***
（大）食品	0.10	***
（中）飲料・酒類	−0.20	***
（中）ペット用品	−0.63	***
（中）化粧品	−0.87	***
（中）DIY 用品	−13.52	
AEvsSY（正例：イオン，負例：西友）		
（大）衣料・身の回り・スポーツ用品	13.36	
（大）耐久消費財	1.11	*
（中）家庭用品	0.53	**
（中）生鮮食品	0.28	***
（大）日用品	0.07	
（大）食品	0.11	***
（中）飲料・酒類	−0.10	***
（中）化粧品	−0.36	

*** : 0.1%有意，** : 1%有意，* : 5%有意

5 おわりに

本研究ではスキャンパネルデータを対象に，ロジスティック回帰モデルと変数選択を利用して店舗選択モデルを構築した．整数計画法による変数選択手法は，多

くの場合に変数増減法よりも AIC・BIC 値の良い変数集合を選択し，AIC 基準の場合には変数増減法よりも予測精度の高いモデルを構築した．一方で，BIC 基準の変数選択は予測精度の向上にはつながらなかったが，有意な説明変数を正確に選択していた．また，変数選択の結果を分析することで，各店舗における差別化要因を把握し，それぞれの店舗の特徴を明らかにすることができた．

謝辞

本研究の一部は，専修大学情報科学研究所の共同研究助成を受けたものである．

参考文献

[1] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

[2] Baker, J., Parasuraman, A., Grewal, D., & Voss, G. B. (2002). The influence of multiple store environment cues on perceived merchandise value and patronage intentions. *Journal of Marketing, 66,* 120–141.

[3] Bertsimas, D., King, A., & Mazumder, R. (2014). Best subset selection via a modern optimization lens. INFORMS Philip Morse Lecture for 2014–2015.

[4] Bloemer, J., & de Ruyter, K. (1998). On the relationship between store image, store satisfaction and store loyalty. *European Journal of Marketing, 32,* 499–513.

[5] Briesch, R. A., Chintagunta, P. K., & Fox, E. J. (2009). How does assortment affect grocery store choice?. *Journal of Marketing Research, 46,* 176–189.

[6] Efroymson, M. A. (1960). Multiple regression analysis. In Ralston A., & Wilf H. S. (Eds.), *Mathematical Methods for Digital Computers* (pp. 191–203). New York, NY: Wiley.

[7] Koh, K., Kim, S., & Boyd, S. (2007). An interior-point method for large-scale ℓ_1-regularized logistic regression. *Journal of Machine Learning Research, 8,* 1519–1555.

[8] Konno, H., & Takaya, Y. (2010). Multi-step methods for choosing the best set of variables in regression analysis. *Computational Optimization and Applications, 46,* 417–426.

[9] Konno, H., & Yamamoto, R. (2009). Choosing the best set of variables in regression analysis using integer programming. *Journal of Global Optimization, 44,* 273–282.

[10] Lee, S., Lee, H., Abbeel, P., & Ng, A. Y. (2006). Efficient L_1 regularized logistic regression. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence* (pp. 401–408). Menlo Park, CA: AAAI Press.

[11] Leszczyc, P. T. P., & Timmermans, H. (2001). Experimental choice analysis of shopping strategies. *Journal of Retailing, 77,* 493–509.

[12] Miyashiro, R., & Takano, Y. (2013). Mixed integer second-order cone programming formulations for variable selection. Technical Report, No. 2013-7, Department of Industrial Engineering and Management, Tokyo Institute of Technology.

[13] Miyashiro, R., & Takano, Y. (2015). Subset selection by Mallows' C_p: A mixed integer programming approach. *Expert Systems with Applications, 42,* 325–331.

[14] Pacheco, J., Casado, S., & Núñez, L. (2009). A variable selection method based on tabu search for logistic regression models. *European Journal of Operational Research, 199,* 506–511.

[15] Pan, Y., & Zinkhan, G. M. (2006). Determinants of retail patronage: A meta-analytical perspective. *Journal of Retailing, 82,* 229–243.

[16] Reutterer, T., & Teller, C. (2009). Store format choice and shopping trip types. *International Journal of Retail & Distribution Management, 37,* 695–710.

[17] Sato, T., Takano, Y., Miyashiro, R., & Yoshise, A. (2015). Feature subset selection for logistic regression via mixed integer optimization. Discussion Paper Series, No. 1324, Department of Policy and Planning Sciences, University of Tsukuba.

[18] Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461–464.

[19] Unler, A., & Murat, A. (2010). A discrete particle swarm optimization method for feature selection in binary classification problems. *European Journal of Operational Research, 206,* 528–539.

[20] Yusta, S. C. (2009). Different metaheuristic strategies to solve the feature selection problem. *Pattern Recognition Letters, 30,* 525–534.