

POS データ・人流計データ・滞留計データによる ISM 分析手法の開発

野 田 哲 夫*

Development of Interpretive Systems Mode Analysis Technique
by using POS, Circulation and Delayed Data of Human

Tetsuo NODA

キーワード：ISM 分析, FSP データ, POS データ, 人流計・滞留計データ, マーケット・バスケット分析

はじめに

当研究室（情報経済論）においては、2003 年度より島根県出雲市の経営コンサルタント会社（有）アシスト¹（業務内容はマーケティング・プロモーション、経営コンサルティング、調査分析など）と以下のような共同研究を進めてきた。

- ・2003 年度：「高精度小規模店舗売上予測・立地判定システムの開発と業務の効率化」（経済産業省 IT 活用型経営革新モデル採択事業）
- ・2004 年度：「市場シミュレーションシステム及び飼育生産履歴・工程管理・経営効率化システムの構築」（経済産業省 IT 活用型経営革新モデル採択事業）
- ・2005 年度：「POS データ・人流計データ・滞留計データによる分析手法の開発とその実施」

特に「高精度小規模店舗売上予測・立地判定システムの開発と業務の効率化」においては売上予測・立地判定システムの構築のために、出店予定店舗地域における実店舗の実地調査データ（敷地、敷形、交通量、看板、視界などの立地条件）および GIS データ（人口、人口密度、年齢別人口、世帯別人口、住宅などの地理情報データ）を調べ、売上高との相関を求め、統計的手法により相関が高いデータを説明変数、売上高を被説明変数として SPSS を利用した「対話型変数選択法（変数減少法）」による回帰分析を行い、実地調査データと GIS データによる売り上げ予測の方程式を導出した。この研究成果は拙稿「小規模店舗売上予測システムの統計分析に見られる小規模店舗の出店と存続の課題」²として出稿した他、アシストにおいて「小規模店舗の売上予測システム」³として製品化され、同社のコンサルティング業務における戦略的商品

¹ <http://www.izumo.ne.jp/~a0707/>

² 野田 [2005] 参照。

³ <http://www.izumo.ne.jp/~a0707/work/syokibo.htm> を参照。

*島根大学法文学部教授

として位置づけられている。

また、「POS データ・人流計データ・滞留計データによる分析手法の開発とその実施」においては、(株)バブ日立工業(広島県呉市)⁴が開発した「来店者カウント」、「滞留者密度カウント」、「男女判別カウント」などのシステムと従来からある「POS データ」との組み合わせから発展し ISM (In-Store Merchandising Analysis)＝顧客の商品選定プロセス分析を行った。

「POS データ・人流計データ・滞留計データによる分析手法の開発とその実施」においては POS データ・人流計データ・滞留計データによる ISM の手法設計、出雲市内スーパーマーケットへの人流・滞留観測システム設置の指導、POS データ・人流計データ・滞留計データの収集と SPSS Clementine によるマーケット・バスケット分析⁵などを行い、顧客の姿や商品の販売特性を踏まえ、現在購入に至っていない顧客に今後購入してもらう為にどのような手段を講じればよいか、これらの点に付いて定量データの分析結果から客観的に結論を導き出すことを目的として出雲市内の店舗において実証実験とその実験結果に基づく検証を行った。これは ISM 分析に「新たな指針となる指標」を提供するものである。本稿ではこの研究の経過と成果に基づいて開発された手法について論説する。

1. ISM 分析の基本と発展

(1) ISM 分析手法について

消費者の店舗選択を考えた時、消費者は商

圏内の各店舗の特性を良く知っており、自らの目的に応じて業種・業態を使い分けているという事実がある。この理由として、消費者の業種・業態選択行動は、小売の提供する価格よりも、消費者の利便性やライフスタイルなど、もっぱら消費者側の要因で決まることが考えられる。その結果、主婦を対象とした調査でも、平均として年間の総購入金額の約 60% をスーパーマーケットでの買物支出に使っていると言われている。

一方、価格の面から考えると、価格は顧客のトライアルには効くが、各顧客がロイヤリティまで至るには、更に消費者の利便性や店員との個人的つながりが重視されると言われている。

これらの事実を実際の個店に応用し、顧客の実態を把握して将来に向けて検討を行おうとするならば、現在の顧客の姿や商品の販売特性を踏まえ、現在購入に至っていない顧客に今後購入してもらう為に、どのような手段を講じればよいか、これらの点に付いて定量データの分析結果から客観的に結論を導き出すことが必要となる。目的とする分析は ISM 分析 (In-Store Merchandising Analysis) と呼ばれる。

従来から FSP データ (Frequent Shoppers Program Data)⁶ や POS データ (Point Of Sales Data) を用いたデータマイニングの可能性についてその使用法は、「クラシフィケーション (優良顧客発見)」、「クラスタリング (階層特性)」、「マーケット・バスケット分析 (同時購買商品の把握)」、「時系列パターン分析 (需要変動パターン)」の 4 つがあることが指摘さ

れている⁷。人流・滞留分析は、上記に加えて新しい手法として検討され「来店者カウント」、「滞留者密度カウント」、「男女判別カウント」(下記 2 参照) が新たにデータとして加わり、以前からある「POS データ」との組み合わせから発展し ISM 分析の「新たな指針となる指標」を提供することを目的とする。

(2) ISM 分析システムの概要

今回、バブ日立工業株式会社が開発し ISM 分析において使用する一連の計測システム及びそのシステムから出力される計測値は、以下の通りである。

計測システムの名称	アウトプット	備考
来店者カウントシステム	人数カウント	
	平均滞留時間	
	平均滞留者数	
男女判別システム	男女別カウント	
滞留者密度カウントシステム	混雑度カウント・滞留度カウント	
(動線追跡システム) ^{*1}	(通行量、歩行速度、追跡軌跡)	(動線データ)

^{*1} 動線追跡システムは、現段階では機器未完成の為、今回設置されなかった。

次に、それぞれのシステム設置の目的と判別方法であるが、

- ①来店者カウントシステムの目的は、
- ・測定エリアへの入退場者数カウントと分析への応用

- ・測定エリアの通過者数カウント及びその方向の計測と分析への応用
- ・1 次処理として算出される滞留者数及び平均滞留時間の分析への応用

- ②男女判別システムの目的は、
- ・測定エリアへの入退場者数カウントの男女数算出と分析への応用

- ③滞留者密度カウントシステムの目的は、
- ・滞留の多少を判別する為の指標を算出 (滞留度) . . . 主として目視にて判別
 - ・混雑の多少を判別する為の指標を算出 (混雑度) . . . 主として目視にて判別
 - ・顧客の店舗内での平均的な行動結果 (引き返し) の確認 . . . 主として目視にて判別
 - ・人の流れを狙った商品配置・機器配置 . . . 主として目視にて判別
 - ・効率的商品配置 . . . 分析
 - ・混雑度を使った商品の販売特性 . . . 分析

- ④動線追跡システムの目的 (現状未設置) は、
- ・顧客を特定 (ターゲット) した買い回り行動の把握
 - ・平均的な動線パターンの把握と分析への応用

⁶ 顧客全体にサービスするのではなく、特に利用してくれる顧客 (フリークエント・ショッパー) に対して特別のサービスをする作戦。例えば、小売流通業では、顧客に会員カードを持たせ、購買額を累積的に確認して各種のボーナスを提供している。同時に、FSP で得られた詳細な顧客データを介して顧客の購買行動を把握し、CRM に活かすことが可能。例えば、FSP データから優良顧客を選別し、得られた優良顧客の購買履歴から併売分析を行い、各商品を意味付けしてサブ・カテゴリーに再分類し、これに基づき実際の商品の棚割を行うことなどが実践されている。

⁷ 古賀 [1998], 63 頁～85 頁参照。

⁴ <http://www.bhic.bhk.co.jp/>

⁵ データマイニングの利用法の 1 つで、POS データや EC サイトのトランザクション (取引) データを分析して、“一緒に買われる商品”の組み合わせを発見する探索的データ分析のこと。分析結果は、店舗内レイアウトや棚割り、陳列の計画、特売品・併売キャンペーン商品の選定、商品仕入れなどを検討するときの参考にする。

・動線の無駄の排除（評価測定する方法に付いては今後検討）
 ・通行量の表示とデッドポイントの把握である。

これら出力されるデータの中から、今回は特に「人の動き（顧客動線）」と「購買傾向」の把握を行うことになる。

(例1) 一番出入りの多い時とPOSデータとの関係

(入退場者数、滞留者数 来店者カウント - POS購買者数、購買金額、購買特性)

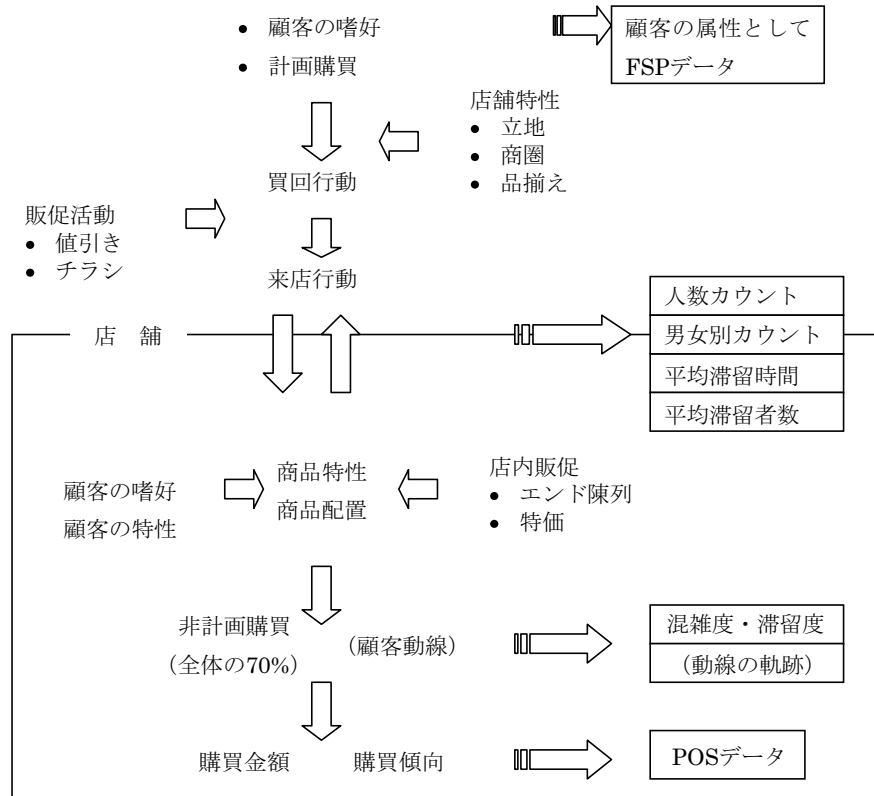
(例2) 人が集中している場所とその時に売れている商品特性の関係（滞留者密度カウントシステム）

(例3) 混雑度と商品に対する興味を確認する指標（滞留者密度カウントシステム）などが考えられる。

(3) 計測データの確認

評価検証する項目と計測データを図示すると下図のようになる。

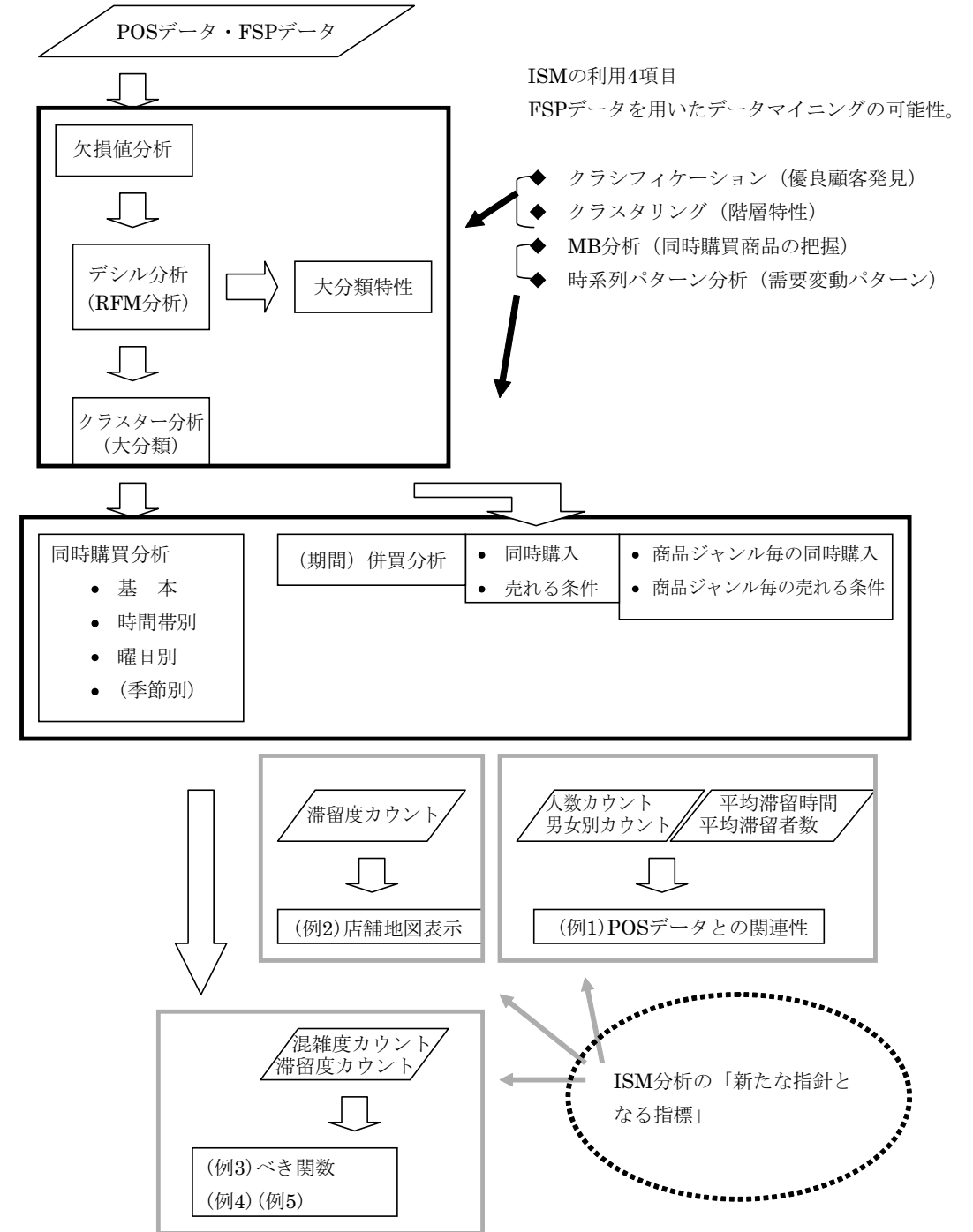
⇒ で示しているのが今回の実証実験で出力されるデータである。



(4) 分析手法とISM分の新たな手法

上記データを下にした分析プロセスが下図のとおりであるが、特に右下の部分今回の

ISM分析において「新たな指針となる指標」となる部分である。



ISMの利用4項目
 FSPデータを用いたデータマイニングの可能性。

- ◆ クラシフィケーション（優良顧客発見）
- ◆ クラスタリング（階層特性）
- ◆ MB分析（同時購買商品の把握）
- ◆ 時系列パターン分析（需要変動パターン）

2. 人流計・滞留計データと ISM 分析 (In-Store Merchandising Analysis) の開発

(1) データ間の連関と分析

現在までの収集、あるいは収集予定のデータは滞留データに関しては棚（滞留区画）ごとの滞留時間、および店舗が集約した POS 端末データである。特に POS データに関しては下記統計的分析 (C~E) によって店舗売上高との連関を探ることが可能である。

両データの問題点としては、滞留データは、POS データの個人情報と関連されていない（滞留人物と購買人物の一致が確認できない）点である。*2

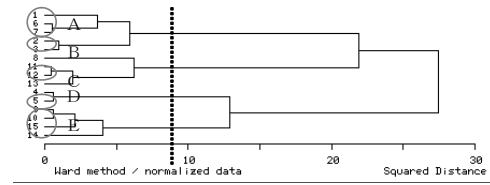
そこで、現状では、滞留データと POS データ（売上高）を関連させるためには、滞留データを時間ごとに特徴づけるグループに分け (A)，これを時間単位で POS データと連関、分析する手法が考えられる。

*2 動線追跡システムが完備されていないため。

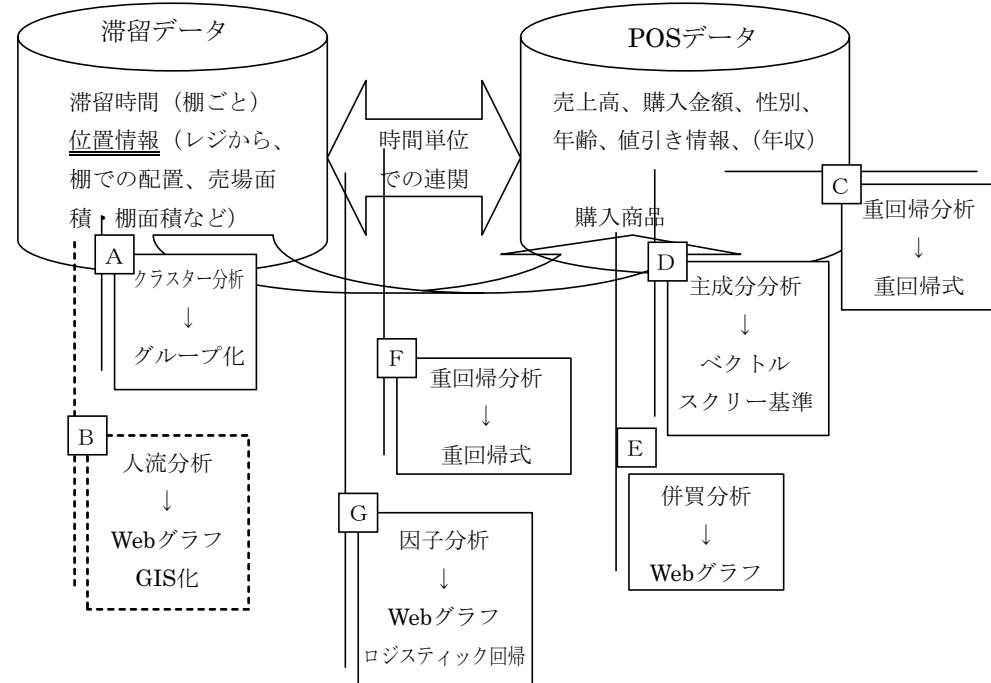
(2) アウトプット (例)

以上の分析から得られるアウトプットの形は以下のとおりである。データはダミーを使ってシミュレーションしてある。

A クラスタ分析：デンドログラムによるグループ化



特に滞留データを類型化するために、カメラ区画と時間ごとに分かれた各サンプルの滞留時間の他に、カメラ区画の位置情報（レジからの距離、面積）や、さらに小区画ごとの位置情報と小区画滞留時間のデータから、クラスタ分析を行い、各サンプルのクラスター距離の近いものをまとめる。その結果、例えば時間帯ごとの各区画のグループ化が可能となる。



係数^a

モデル		非標準化係数		標準化係数	t	有意確率	B の 95% 信頼区間		共線性の統計量	
		B	標準誤差	ベータ			下限	上限	許容度	VIF
1	(定数)	1691.971	432.738		3.910	.000	841.984	2541.958		
	POS 購買者	12.227	13.428	.211	.911	.363	-14.149	38.603	.032	30.977
	混雑度	-3.401	12.920	-.061	-.263	.792	-28.779	21.977	.032	31.035
	A 区滞留	.306	6.140	.002	.050	.960	-11.753	12.366	.988	1.012
	B 区滞留	-3.293	6.175	-.024	-.533	.594	-15.422	8.837	.828	1.208
	E 区滞留	1.679	5.120	.014	.328	.743	-8.378	11.737	.936	1.068
	D 区滞留	-.347	5.312	-.003	-.065	.948	-10.781	10.087	.888	1.126
	C 区滞留	-1.735	5.326	-.014	-.326	.745	-12.196	8.726	.924	1.082
	G 区滞留	10.277	6.576	.071	1.563	.119	-2.640	23.193	.849	1.177
	F 区滞留	6.893	6.375	.050	1.081	.280	-5.630	19.415	.819	1.222

a. 従属変数：合計金額

例) 10月1日の午後7時~午後8時の No 1の棚 (乳製品)・・・Aグループ
11月5日の午後2時~午後3時の No 2の棚 (缶詰)・・・Dグループ

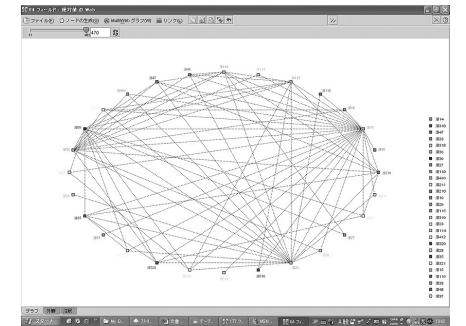
データと滞留データを関連付ける主成分を分析し、説明力の強い主成分が含まれているところまでとりあげる。
この主成分に絞って重回帰分析を行ってもよい。

C, F 重回帰式

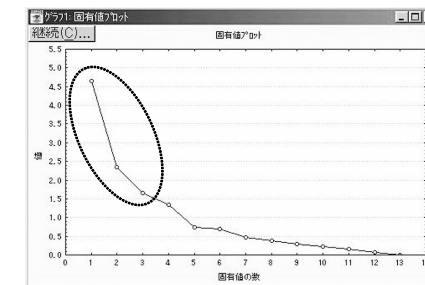
時間単位の POS データ、あるいは POS データと滞留データは連続した数量データであるので、相関分析を行い、また相関が高いものを絞って重回帰分析を行えば、データ間の関連を示す重回帰式を算出できる。

これはあくまで統計数量的な分析であるので、データ間の関連の信頼性には検討が必要である。

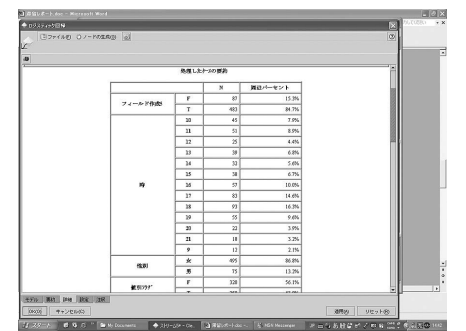
B, E, G Web グラフ, ロジスティック回帰



D スクリーン基準 (主成分の決定)



時間単位の POS データ, あるいは POS



時間単位の POS データ, あるいは POS

データと滞留データを連続数値データから離散データに変換し、Web グラフ化することによって、関連が強いものを抽出し、その関連を規定する因子を分析することが可能である。

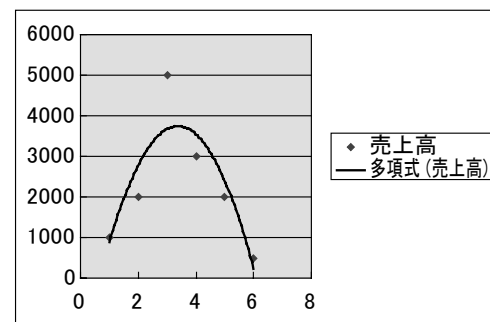
G-1

また、滞留時間のみで Web グラフ化し、店舗の位置情報と関連づければ、人流をある程度グラフィック化することも可能であろう。

G-2

(3) 両データの分析結果からの総合

① **A** でグループ化した各グループと、売上高などの POS データとの相関、回帰式の導出を行う。グループごとの結びつきの度合い(滞留の散布度)を得点化あるいは数量化し、POS データの数値データ(特に滞留の位置情報と関連する商品の売上げデータ)と相関、多項近似曲線およびその関数式を導出する。



② **B** の人流分析を時間ごとにパターン化、数量化し、①と同じように POS データとの相関、回帰式を導出する。

③ **A** のグループ化と **D** の主成分分析、あるいは **A** と **E** の併売分析との関連→論理的に説明する。

④ **G** の因子分析によって抽出された因子を論理的に再検討した上で、**C** や **F** の重回帰分析に反映させて、重回帰式の論理的な裏づけと精度を高める。この場合アウトプットは

回帰式として算出される。

商品の売上げ向上に結びつき、またビジュアル的にもわかりやすいのは①と②の方法による分析であると考えられる。ただしその場合、それぞれのグループ、散布度の数量化(次項参照)がカギになってくる。この際、マーケティングにおける数量化理論が参考にできると思われる。

数量化理論はマーケティング・データの分析において使われる理論で、アンケートなどで採集された顧客反応を質的(カテゴリー)なデータとしてつかまえ、有用な情報を汲み出すものである。

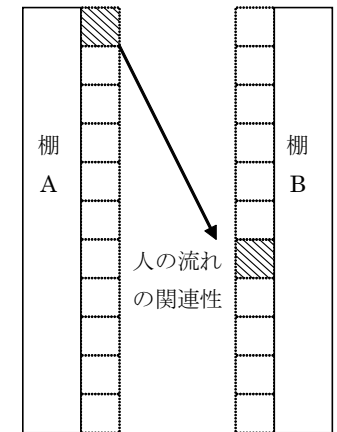
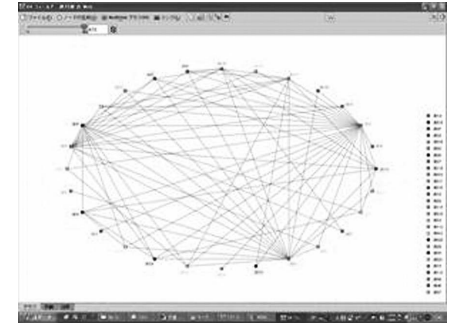
例えば、顧客が商品の購入に際して、その決め手となるものは、店舗のイメージ、商品の品揃え、鮮度、品質、価格、サービス等である。しかし、これらのデータの中には数値として取り扱えない質的なマーケティング現象が多く含まれており、そのままでは数量化できない。

そこで数量化理論では、その分析目的に応じて、on/off のアイテム・カテゴリー因子(層別因子)で、多数の人々の回答を反応パターンとして集め、得点を与え評価するものである。

今回は滞留を商品配置に対する顧客の反応と見て、滞留時間をいったん on/off のアイテム・カテゴリー因子(層別因子)に変換し、多数の顧客の反応パターンとして捉えて、**A** でグループ化したカメラ区画ごと時間ごとの各グループに得点を与えてみる。その得点化されたデータと売上高などの POS データとの相関、回帰式の分析によって①のような関数式の導出が可能となれば、商品配置による売上高の予測が可能となる。

3. 人流計・滞留計データによる ISM 分析の応用と課題

前節で開発した人流計・滞留計データによる ISM 分析手法に基づき、今回の実証実験(出雲市内スーパーマーケットへの人流・滞留観測システム設置の指導、POS データ・人流計データ・滞留計データの収集と SPSS Clementine によるマーケット・バスケット分析)を行った。第1節において課題とした「人の動き(顧客動線)」と「購買傾向」の把握に関して、今回は動線追跡システムが完備されていないため、人流計データの計測値、特に入場者数と退場者数とのデータから



(例1) 一番出入りの多い時と POS データとの関係について

人流計データの計測値と POS データ(商品のカテゴリーカルデータ)との相関結果

人流計データ計測値の評価と、POS 合計の比較統計量との相関係数(ピアソンの積率相関係数)

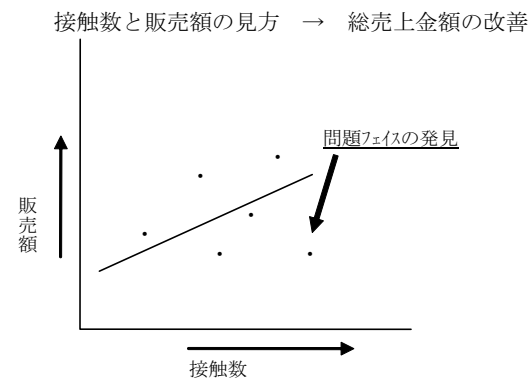
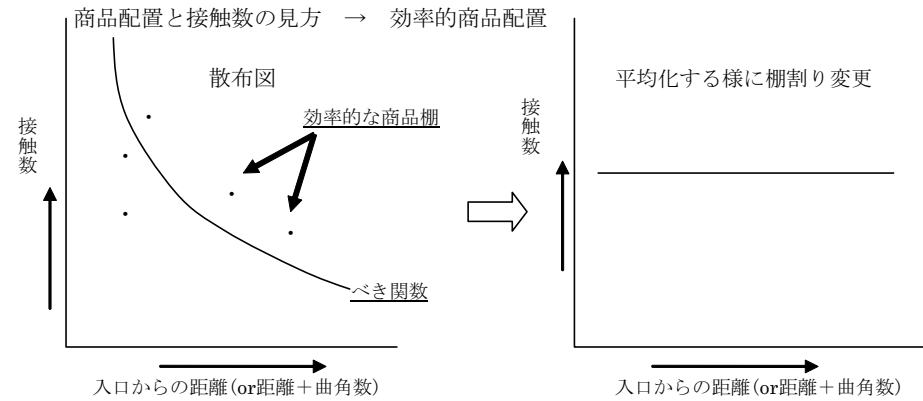
来場者数 0.220 / 退場者数 0.236
結果としては高い相関は得られなかった。今後動線追跡システムを完備することによる精度の改善が求められる。

(例2) 人が集中している場所とその時に売れている商品特性の関係人の流れの関連性分析

関連性の強いエリアとその流れの特性分析(Web グラフ、因子分析デシジョンツリー C 5.0)によって人の流れ・動線表示(行動パターン)をビジュアル表示し関連性の確認を行い、人が集中している場所とその時に売れている商品の関係が表れた。

(例3) 混雑度と商品に対する興味を確認する指標について

これは入口からの距離(商品配置)ー接触数(混雑度、滞留度)ー販売額、の関連であり距離(配置)・接触数(混雑度、滞留度)・販売額の3つの指標による3次元散布図として図示される。



【参考文献】

D. Hand, H. Mannila, P. Smyth [2001], Principles of Data Mining. MIT Press, Cambridge, MA.

W. Frawley and G. Piatetsky-Shapiro and C. Matheus [1992], Knowledge Discovery in Databases: An Overview. AI Magazine, pp. 213-228.

古賀広志 [1998], 「データマイニングの組織

的意義」『商大論集』第 50 巻, 所収。

福田剛志, 森本康彦, 徳山豪 [2001], 『データマイニング』, 共立出版。

野田哲夫 [2005], 「小規模店舗売上予測システムの統計分析に見られる小規模店舗の出店と存続の課題」, 『島根大学法文学部紀要 社会システム学科編 社会システム論集』第 10 号, 所収。