

---

# マーケティング・データ解析の潮流

専修大学 生田目 崇

## 日本におけるマーケティング

- 19世紀 三井呉服店で反物の量り売りを実施(顧客のニーズに合わせた販売戦略のはじまり)
- 1905年 オハイオ州立大学にて講義科目「マーケティング」が登場
- 1955年 日本生産性本部がアメリカ視察 帰国後の会見でマーケティングの重要性が紹介された
- 1982年 セブン・イレブンジャパン POSシステム導入

## マーケティングの定義

### ■ AMA(2004年)

マーケティングとは、組織と利害関係者両者にとって有益となるよう、顧客に向けて価値を創造、伝達、提供したり、また顧客との関係性を構築するための、組織的な働きと一連のプロセス

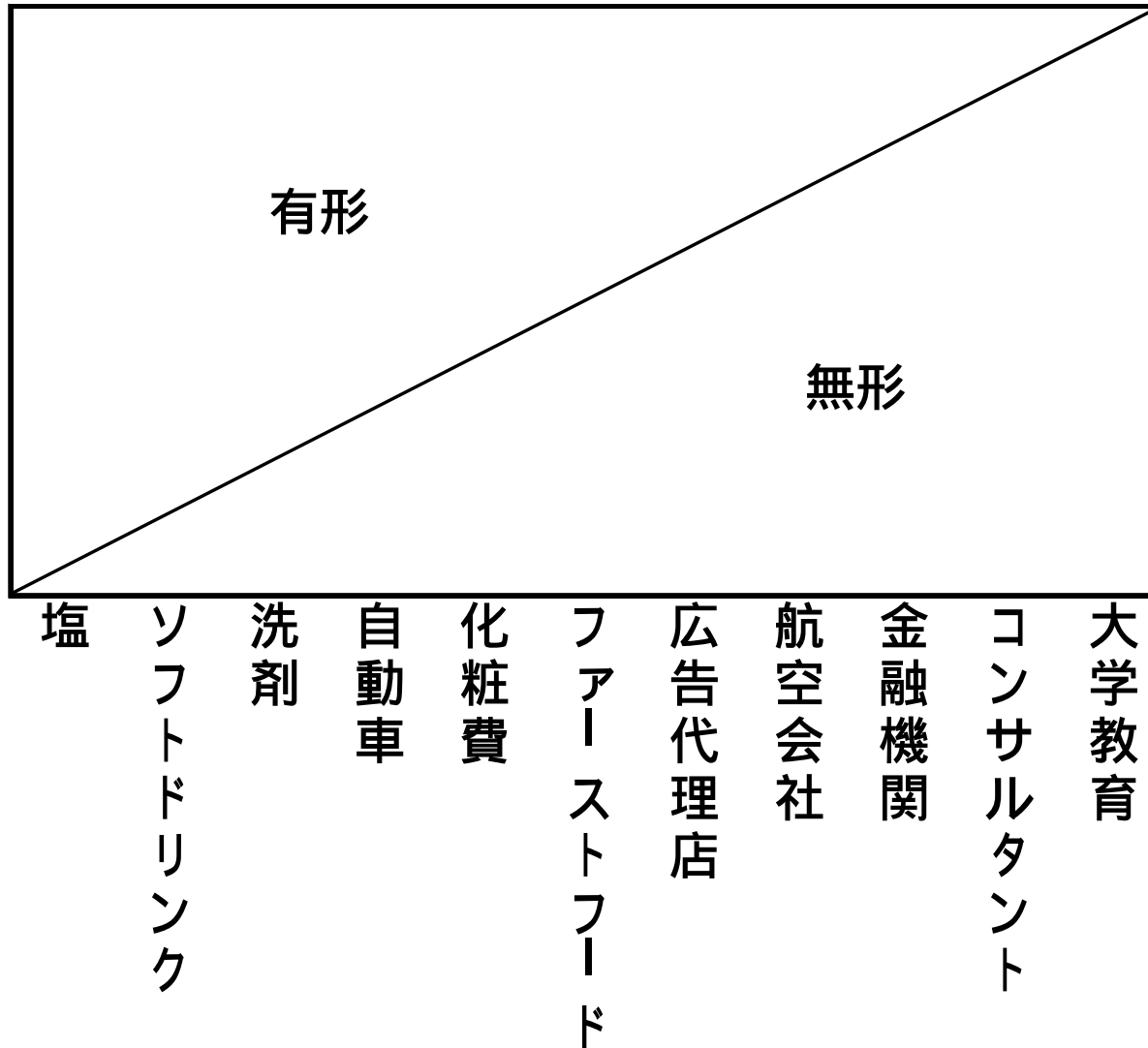
### ■ 日本マーケティング協会(1990年)

マーケティングとは、企業および他の組織がグローバルな視野に立ち、顧客との相互理解を得ながら、公正な競争を通じて行う市場創造のための総合的活動である

### ■ 国の環境・法律などの相違のために定義は多少異なる 時代とともにマーケティングの領域が変化

### ■ いずれにしても、小売業に限定することのない、すべての組織(もしくは個人)にかかわる活動として定義

# 製品vsサービス



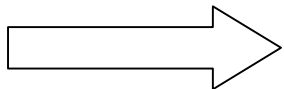
# データを用いてなにを知りたいか？

## ■ マーケティングの4P

- 製品 Product
- 価格 Price
- プロモーション Place
- 流通 Promotion

## ■ マーケティングの4C

- |             |                          |
|-------------|--------------------------|
| - Product   | 顧客価値(Customer Value)     |
| - Price     | 顧客コスト(Customer Cost)     |
| - Promotion | コミュニケーション(Communication) |
| - Place     | 利便性(Convenience)         |



さまざまなデータから以上

## データの点からの大きな転換

### ■ POSデータの登場 (POS: Point of Sales 販売時点情報)

- 1982年 セブン・イレブン・ジャパンによる小売業へのPOSシステムの導入以後, チェーン・ストアなどを中心にPOSシステムによる業務改善, 売上処理が一般的に

### ■ POSデータ

- 小売業の場合, レジを通った「すべて」の「単品」が「販売された時点」に記録
- 元々は在庫発注業務, レジ作業の効率化のための導入
- 売上の時系列把握による, 売れ筋商品, 死に筋商品の把握
- レシートデータからの同時購買分析
  
- ただし, 顧客に関するデータふくまれない

## FSP, スキャナ・パネル・データ

### ■ Frequent Shoppers Program (ポイントカード) データ

- ポイントカード別売上 = POSデータ + 顧客情報 (= ID付POSデータ)
- 顧客の購買履歴を把握できる
  - 時点をまたいだ同時購買の分析
  - 顧客別のアプローチ

### ■ スキャナ・パネル・データ

- FSPデータのデータを調査会社がデータ取得を目的とすることを宣言してモニタを募集
  - ストア・スキャン: 店舗コーザル, プロモーションデータ
  - ホーム・スキャン: 店舗をまたいだ世帯購買データ

## ウェブ・アクセス・ログ・データ

### ■ インターネットのページの閲覧記録

- POSデータが「結果」であるのに対して「プロセス」のデータが取得可能
- リクエストされたページだけでなく、どのページから来たか(リファラ)が記録されるため、訪問者の「動線」を把握することができる。Cookieを使えば、時点をまたいだ訪問履歴も取得可能
- コンバージョン(目的ページへの到達)に至らない、いわば立ち寄りデータも把握できる (= 実店舗でいうところの回遊行動の把握)

年	月	日	曜日	時	分	秒	ポート	ホスト名	ユーザエ	リクエスト	リファラUR	ユーザID
2006	1	1	Sun	0	0	3	-1	c2b00-254	Mozilla/4.0	http://www	http://pag	wYJAnr3p
2006	1	1	Sun	0	7	27	-1	eatcf-704p	Mozilla/4.0	http://www	NA	uMBZghYQ
2006	1	1	Sun	0	36	58	-1	fkcf-01p3	Mozilla/4.0	http://www	http://www	ErLs0Hr4w
2006	1	1	Sun	0	40	18	-1	p2094-ipbf	Mozilla/4.0	http://www	http://sea	QfcTdL4W
2006	1	1	Sun	1	3	13	-1	i220-108-	Mozilla/4.0	http://www	http://www	p7JhVed2Q

セッションID	UNIX時間	ユニークID	ディスプレイ	ディスプレイ	訪問回数	URL(パラ)	広告情報	検索エンジン	リスティン	キーワード
0B7884AA	1.14E+09	2.01E+08	1280	1024	1	http://www	Adwords2	Google	Ad	adwords
F98F16438	1.14E+09	2.01E+08	1024	768	1	http://www	Adwords2	NA		adwords
4E3295889	1.14E+09	2.01E+08	1024	768	1	http://www	sibulla.co	Google	Search	シビラ
C9CA7E7C	1.14E+09	2.01E+08	1024	768	1	http://www	sibulla.co	Yahoo	Search	シビラ
5605F2345	1.14E+09	2.01E+08	800	600	1	http://www	sibulla.co	http://www	kan-net.com/access	



## マーケティング分析

- マーケットに関するさまざまなデータから、マーケティング戦略上有効な情報を取り出す分析が、さまざまな角度から行われている
  - 古くから「マーケティング・リサーチ」という分野



例: 4P)  
製品戦略  
価格戦略  
チャネル戦略  
コミュニケーション戦略

営業実績  
アンケート  
顧客データ  
マクロデータ

統計分析  
多変量解析

## 多変量解析諸手法

- 回帰分析: 因果関係のモデル, 変数の影響の推測
- 判別分析: 事前に与えられているグループを判別するための基準を得る手法
- 主成分分析, 因子分析: 多変量を少数の成分に集約する手法
- コレスポンデンス分析: サンプルと変量の反応のパターンからサンプル間もしくは変量間の類似について分析
- クラスタ分析: 変量のパターンから同等と考えられるグループを作成
- 特に, コレスポンデンス分析や主成分分析などは, 「知覚マップ」の作成に有効で商品同士の相違を視覚的に把握するのにしばしば使われている.

## マーケティング独自のモデル

### ■ 耐久財の普及モデル

- Bassモデル 累積普及率 = 革新者 + 模倣者

### ■ 反復購買のモデル

- Assessor シェア = トライアル購買 × 反復購買

### ■ 商圈のモデル

- Huffモデル 店舗の魅力 = 店舗力 × 面積 / 移動距離

### ■ 営業員配置

- CALLPLAN

- 最大化 訪問効果 - 移動コスト  
制約条件 営業時間 + 移動時間

## 個人への対応

- 集計されたデータではなく, 非集計データの分析
- 1974年 Logit Model (D. McFadden)

$$\Pr\{U_i > U_j | i \neq j\} = \Pr\{V_i + \varepsilon_i > U_j + \varepsilon_j | i \neq j\} = \frac{\exp\{V_i\}}{\sum_j \exp\{V_j\}}$$

- ロジット・モデルの拡張
  - Nested logit model: 段階的選択行動
  - Mixed logit model: ロジットモデルの条件緩和

## データ・マイニング, テキスト・マイニング

- 大量のデータが取得できるようになり, 変量間の関係についてあらかじめ仮説が設定しづらくなってきた
- 仮説を検証するのではなく, むしろ発見することを目的に「マイニング」手法が近年注目されている
  - 統計分野だけでなく, 人工知能や情報工学の成果も総動員したデータ分析
- データ・マイニング = 定量データの分析
- テキスト・マイニング = 言葉の分析 (例: アンケートのフリーアンサー, お客様窓口での対応記録, ウェブコンテンツの分析)
- SAS Enterprise Miner, SPSS Clementine, 数理システム Visual Mining Studio など, マイニングシステムが分析プラットフォーム化

## マイニングのための諸手法

### <データマイニング> 大規模データ, 非線形

- 決定木(ディシジョン・ツリー)
- SVM: 非線形判別手法
- アソシエーション: if thenルールの抽出:例)ビールとおむつ
- ニューラルネットワーク: 非線形予測手法
- K-means法: 探索的クラスタ分析

### <テキストマイニング>

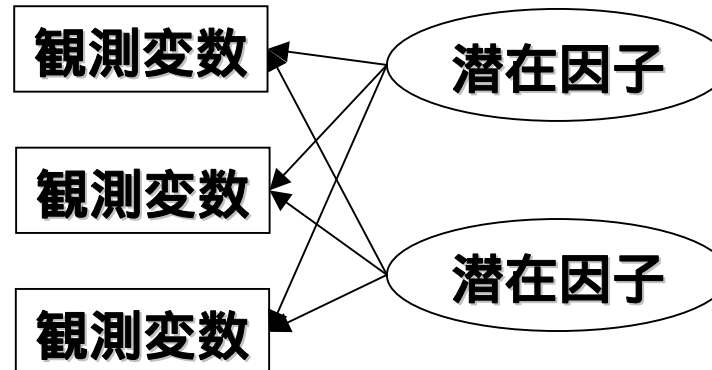
- キーワードの抽出
- 係り結び
- 評判分析

## 潜在構造

### ■ 購買データ = 顧客の心理的判断から出てきた外的反応

- マーケティングの究極の目的は、顧客の「心」を知り、それを満足させるような行動 = 顧客の心において外に出てこない本質的なものを把握

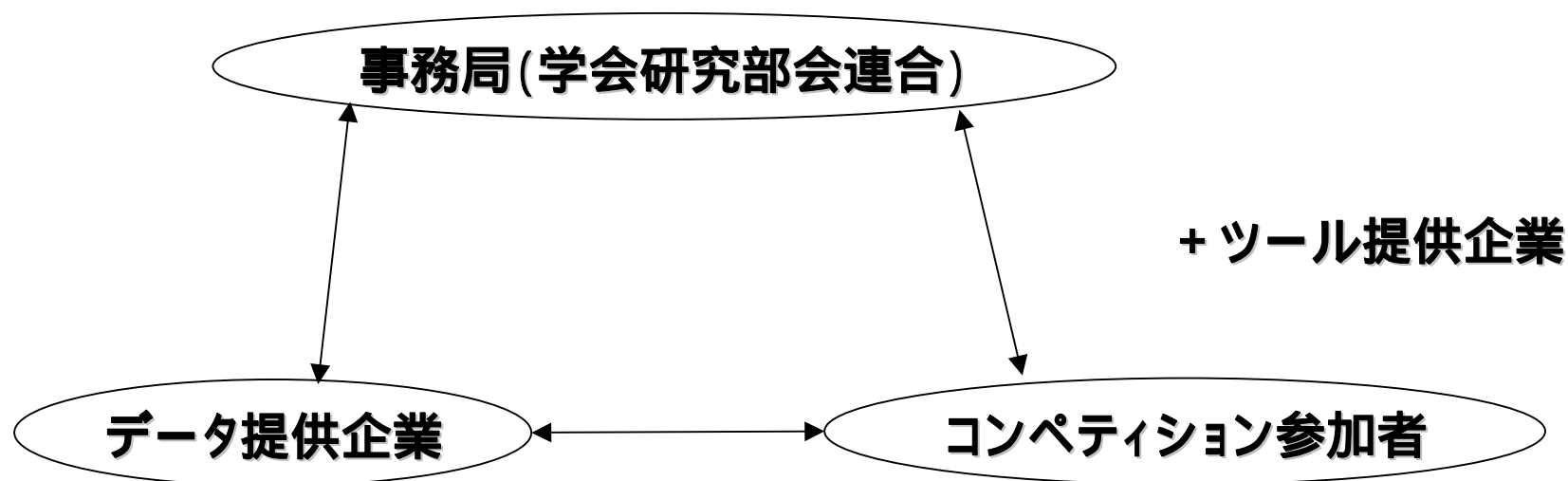
### ■ 古くは因子分析



### ■ これらを拡張した共分散構造分析, 潜在構造モデル

## マーケティング分析分野の産学連携

### ■ データ解析コンペティション(経営科学系研究部会連合協議会主催)



- さまざまな企業よりデータを提供いただき,産学の研究者がそのデータの分析手法,分析結果の有効性を競う.今年14年目の活動.45チーム200名が参加



## 運営体制

### ■ 経営科学系研究部会連合協議会

- 日本OR学会マーケティング・インテリジェンス研究部会
- 日本マーケティング・サイエンス学会ID付POSデータ活用研究部会
- 日本マーケティング・サイエンス学会市場予測のための消費者行動分析研究部会
- 日本データベース学会ビジネスインテリジェンス研究グループ
- 早稲田大学マーケティング・コミュニケーション研究所CRM研究会
- 株式会社NTTデータ技術開発本部
- 株式会社産業科学研究開発センター

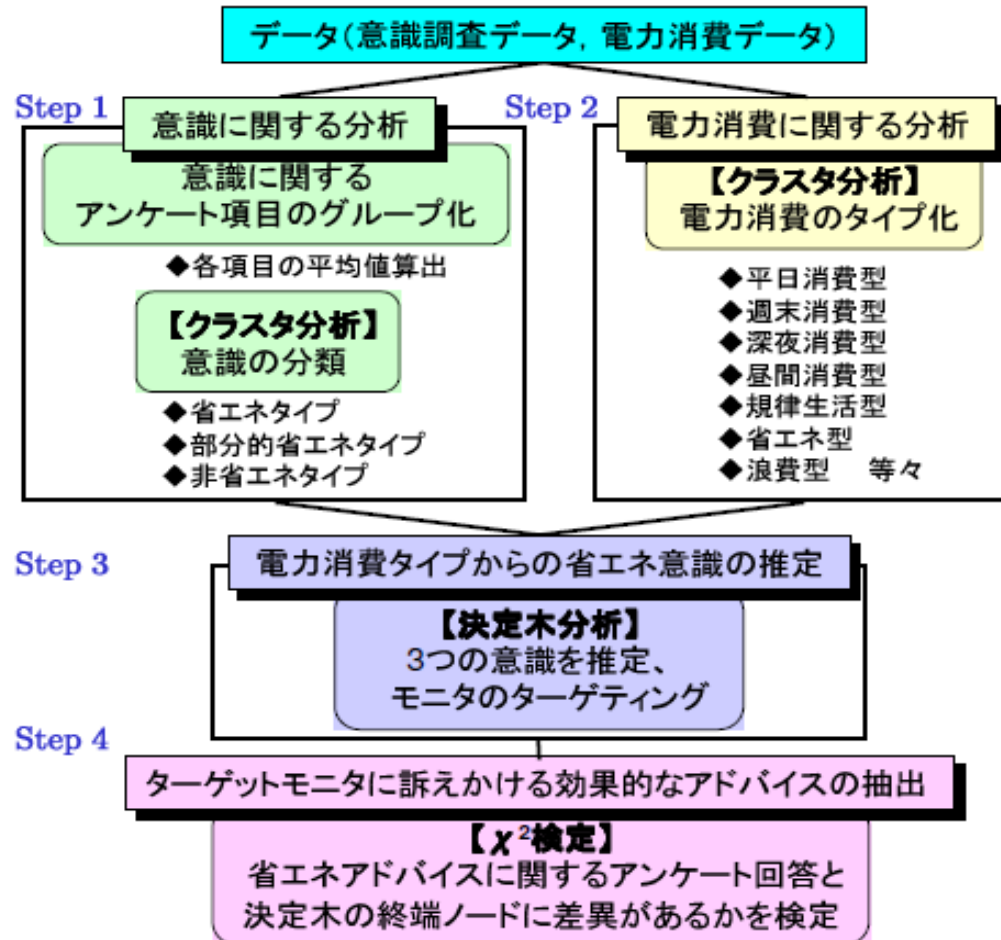
### ■ 学問分野, 人的交流, 次世代の研究者の育成

## データ解析コンペティションの歴史

H6年度	<u>食品購買行動</u> <ul style="list-style-type: none"> <li>● ストア・スキャン・データ</li> </ul>	H13年度	<u>流通CRM</u> <ul style="list-style-type: none"> <li>● ポイントカード・データ</li> </ul>
H7年度	<u>食品・日雑購買行動</u> <ul style="list-style-type: none"> <li>● ストア・スキャン・データ</li> </ul>	H14年度	<u>流通CRM</u> <ul style="list-style-type: none"> <li>● ポイントカード・データ</li> </ul>
H8年度	<u>日用雑貨品購買行動</u> <ul style="list-style-type: none"> <li>● ホーム・スキャン・データ</li> </ul>	H15年度	<u>10周年記念</u> <ul style="list-style-type: none"> <li>● 電力消費, ハウスカード, スーパー・ドラッグストアPOS</li> </ul>
H9年度	<u>観光行動</u> <ul style="list-style-type: none"> <li>● 旅行履歴, 意識アンケート</li> </ul>	H16年度	<ul style="list-style-type: none"> <li>● クレジットカード利用履歴, 加工食品POS</li> </ul>
H10年度	<u>食卓マーケティング</u> <ul style="list-style-type: none"> <li>● メニュー・データ</li> </ul>	H17年度	<u>アミューズメントPOS</u> <ul style="list-style-type: none"> <li>● CD販売店ID付POS</li> </ul>
H11年度	<u>金融マーケティング</u> <ul style="list-style-type: none"> <li>● 行動, 意識アンケート</li> </ul>	H18年度	<u>ウェブ・アクセス・ログ</u> <ul style="list-style-type: none"> <li>● あるサイトのウェブ・アクセス・ログ・データ</li> </ul>
H12年度	<u>金融マーケティング</u> <ul style="list-style-type: none"> <li>● 銀行取引サマリ</li> <li>● 行動, 意識アンケート</li> </ul>	H19年度	<u>オークション・データ</u> <ul style="list-style-type: none"> <li>● BtoB ECのオークション・データ</li> </ul>

# 分析例(1)

## ■ データ:電力消費データ(消費データ,意識,保有機器)



## 分析例(2):クレジットカード利用データからのデフォルト予兆発見

### ■ 顧客属性データ

#### – 顧客属性

- 加入日, 生年, 年齢, 性別, 住所区分, カード区分, 支払コース, 配偶者, 子人数, 居住区分, 勤務区分, 年収, 職種

#### – 限度額

- 総限度額, キャッシング限度額, ショッピング限度額

### ■ トランザクション・データ

#### – 月別トランザクション・データ

- 月別利用金額・回数, 月別総利用限度額, 月別利用残高, 支払結果

#### – 日別トランザクション・データ

- 利用日, 利用金額, 支払方法区分, 利用区分, 利用店舗名, 端末番号

# 分析フロー: デフォルト予兆顧客の抽出とその特徴把握

データ(個人属性データ, トランザクション・データ)

Step 1

デフォルトの定義

「未払い顧客」および「デフォルト顧客」の定義

Step 2

デフォルト・リスク算出モデル構築

二項ロジット・モデルの適用し,  
顧客ごとにデフォルト・リスクを算出

Step 3

モニタリングによる検証

予兆発見のタイミングを捉え,  
デフォルト濃縮率を算出

Step 4

デフォルト顧客の特徴把握

クラスター分析により, デフォルト顧客の  
カードの利用タイプを分類

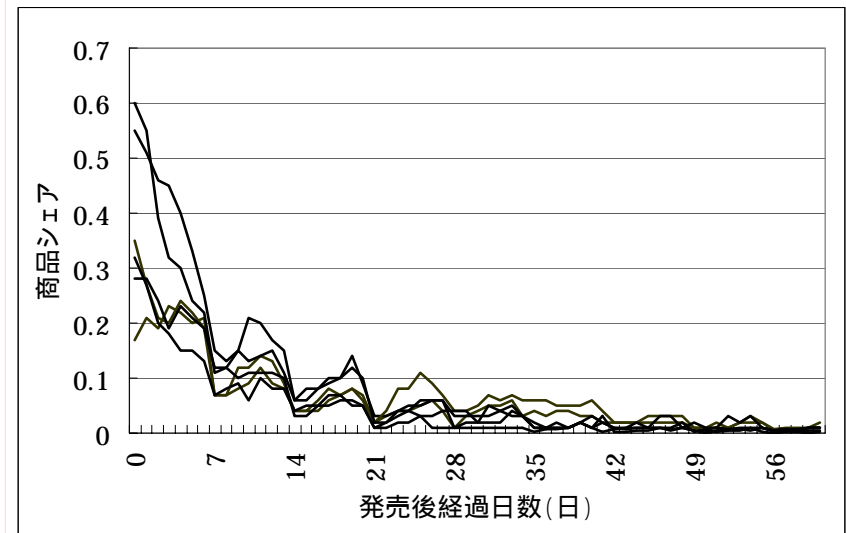
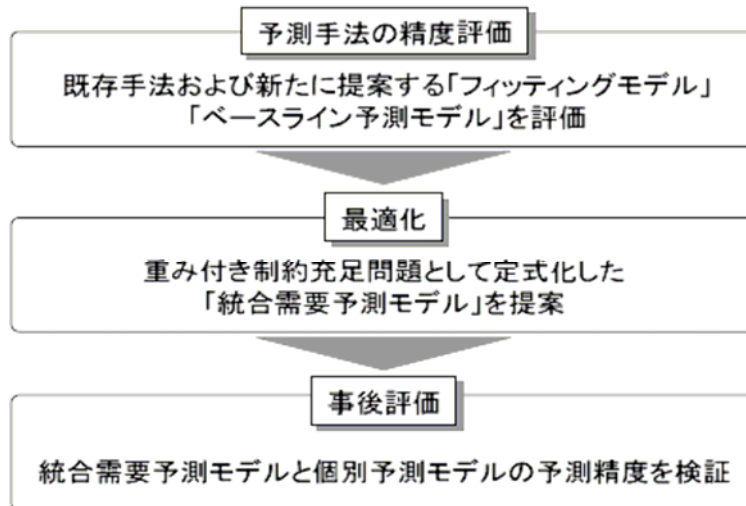
男性Cリボ型  
小額多用型

女性S選好型  
支払遅延型

## 分析例(3): CD販売データ予測

### ■ データ: CD販売のID-POSデータ

- 通常の商品と異なり, 販売直後が売上のピーク



- 予測期間別に適した予測手法の判定を組み合わせ最適化モデルとして定式化し, その結果を考察

## おわりに

- **新しい分析手法, 計算効率の追求は必要**
  - シミュレーションなども新しい方向性
- **分析結果の納得感, 驚き**
  - あまりにブラックボックスな分析では実務で利用不能
- **アクションにつながるプロセス**
  - 具体的なアクションにつながる事が大事