

POSデータを用いた研究の周辺

平成18年7月20日

佐藤忠彦 筑波大学大学院ビジネス科学研究科

於：一橋大学

発表の概要

- マーケティングについて
- 小売業のデータについて
- 状態空間モデルを活用した効果的価格戦略実現のための研究の紹介

流通環境のマクロ的な変化

- 市場の成熟化・消費者嗜好の多様化
 - 新商品数の増加, 商品改廃の激化
- 小売店頭価格の下落
 - 消費者の参照価格が下落
- 小商圈化, 小売業の上位集中傾向
 - マイクロ・マーケティングの必要性が増大
 - FSPなどによる顧客囲い込み
- 価格政策・流通販促の効果の低下
 - 価格プロモーションの悪循環構造
 - 小売業からメーカーへの協力要請の増大

小売マーケティングを高度化し, 効果的に販売活動を行うことが必要

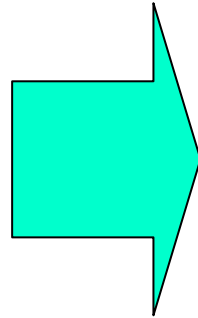
科学的アプローチが重要

小売業の役割の変化

商品が戦略の中心

業種

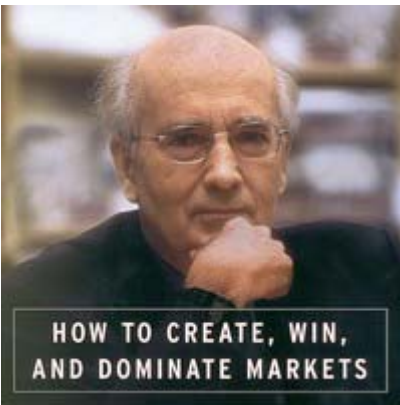
- もの中心
- 販売代理業
- 商品に愛情
- 売り切り追求



消費者が戦略の中心

業態

- ニーズ中心
- 購買代理業
- 消費者に愛情
- 売れ筋追求



マーケティングと市場

<マーケティング> : Kotler, *Marketing Management ME*

- マーケティングは社会活動のプロセスである。その中で個人やグループは、価値のある製品やサービスを作り出し、提供し、他者と自由に交換することによって、必要なものや欲するものを手に入れる

<市場> :

- 特定のニーズやウォンツを持つ潜在的顧客から成っており、彼らはニーズやウォンツを満足するために、取引をする意志と可能性を持つ

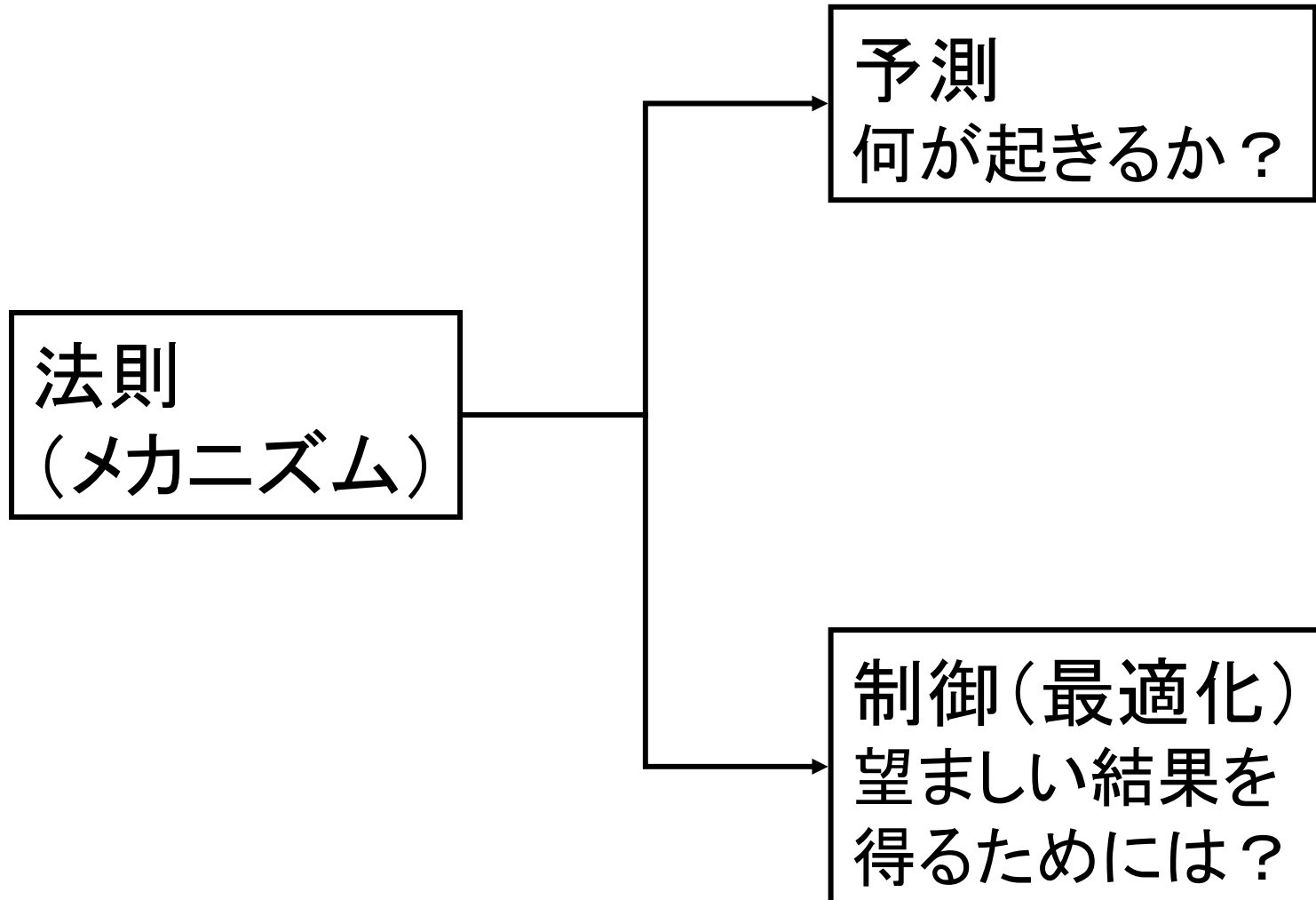
マーケティング・サイエンスとは

- キーワードは「データ」、「論理」、「市場」、「マネジリアル」の4つ
- マーケティングにおける諸問題に対応したマネジメント・サイエンスの一部門
- データと論理に基づいて市場を捉えるための基本的考え方および具体的方法を探求するものであり、つねに意思決定者にとって有用であるというマネジリアルな視点から発想されるもの

マーケティング・サイエンス・アプローチ

- マネジャーがいつも行っている問題発見・分析・解決のノウハウを論理的に整理して明示的なかたちに表現すること
- そのノウハウとマーケティング・サイエンスの知見を現実のデータをベースに融合していくこと
- 「仮説設定→データ収集・分析→実践→学習（仮説改訂）」というサイクルを何度となく繰り返すことにより、正確に市場をとらえるためのノウハウを明示的かつ有形な企業全体の共有資源として育てていくこと

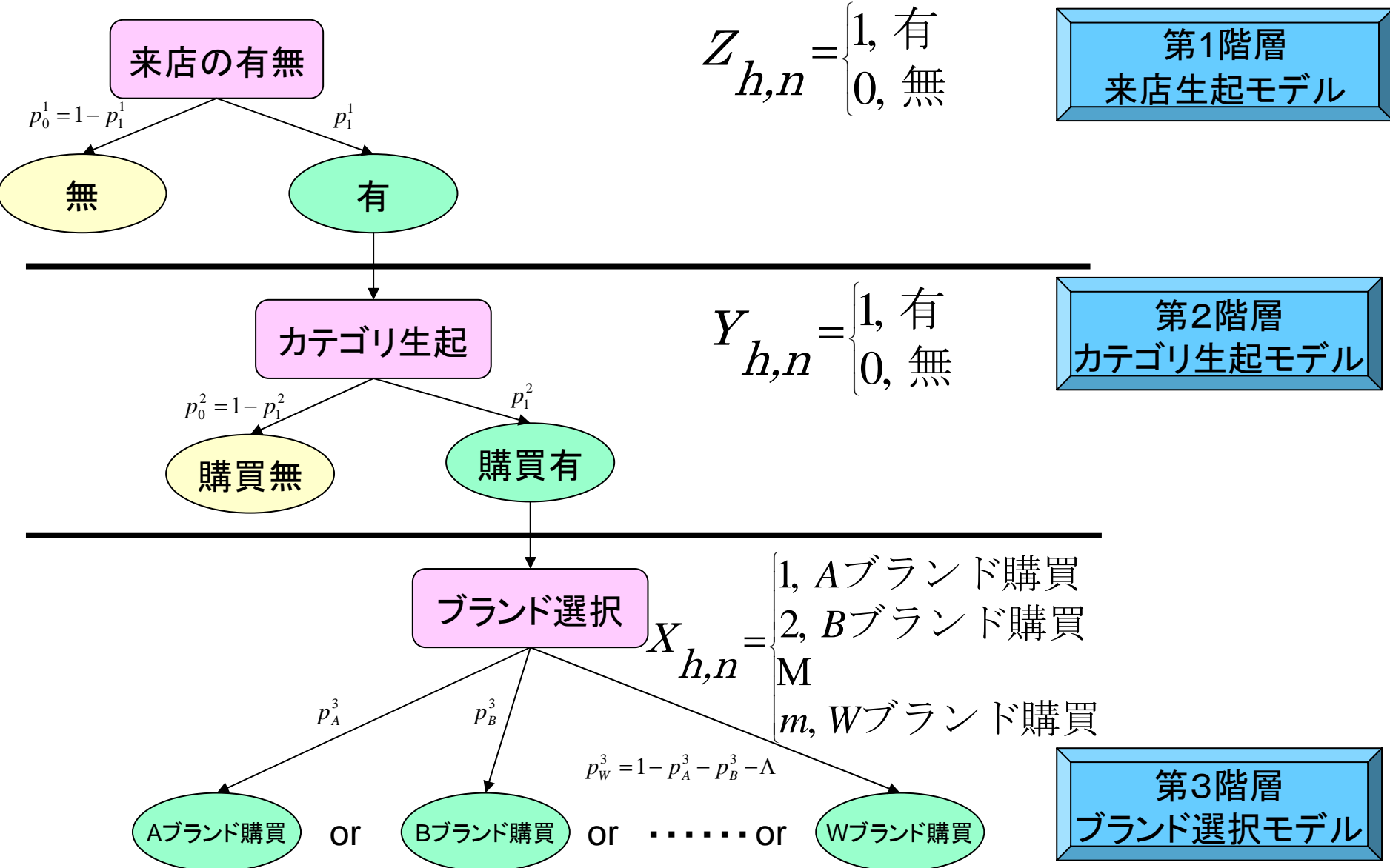
マーケティング・サイエンスにおける モデル化の目的



マーケティング・サイエンスの領域(1)

- 消費者行動
 - 消費者の購買行動のプロセス（問題認識，情報探索，評価，知覚，選好，購買決定），**ブランド選択モデル**，購買後行動
- マーケティング戦略の決定
 - **市場細分化**，ターゲティングと戦略ポジショニング
- 製品デザイン
 - **製品ポジショニング**，プロダクトマップ
- プライシング
 - **数量と価格の関係**，小売価格戦略，差別的価格戦略，価格設定の心理的側面

消費者の購買までのプロセス



ブランド選択モデル:ロジットモデル

家計 n のブランド i に対する確率的効用 U_{ni} は,

- マーケティング変数や家計属性などから成る非確率的効用 V_{ni}
- 測定誤差やその他の要因から成る確率的誤差項 ε_{ni}

の和で表される

$$U_{ni} = V_{ni} + \varepsilon_{ni}$$

代替案の集合を $i=1,2,\dots,I$ とする時, 家計 n がブランド i を選択する確率 p_{ni} を, 「ブランド i の確率的効用が全代替案の中で最大になる」確率として定義する

$$p_{ni} = \Pr(U_{ni} \geq U_{n1}, U_{n2}, \dots, U_{nI})$$

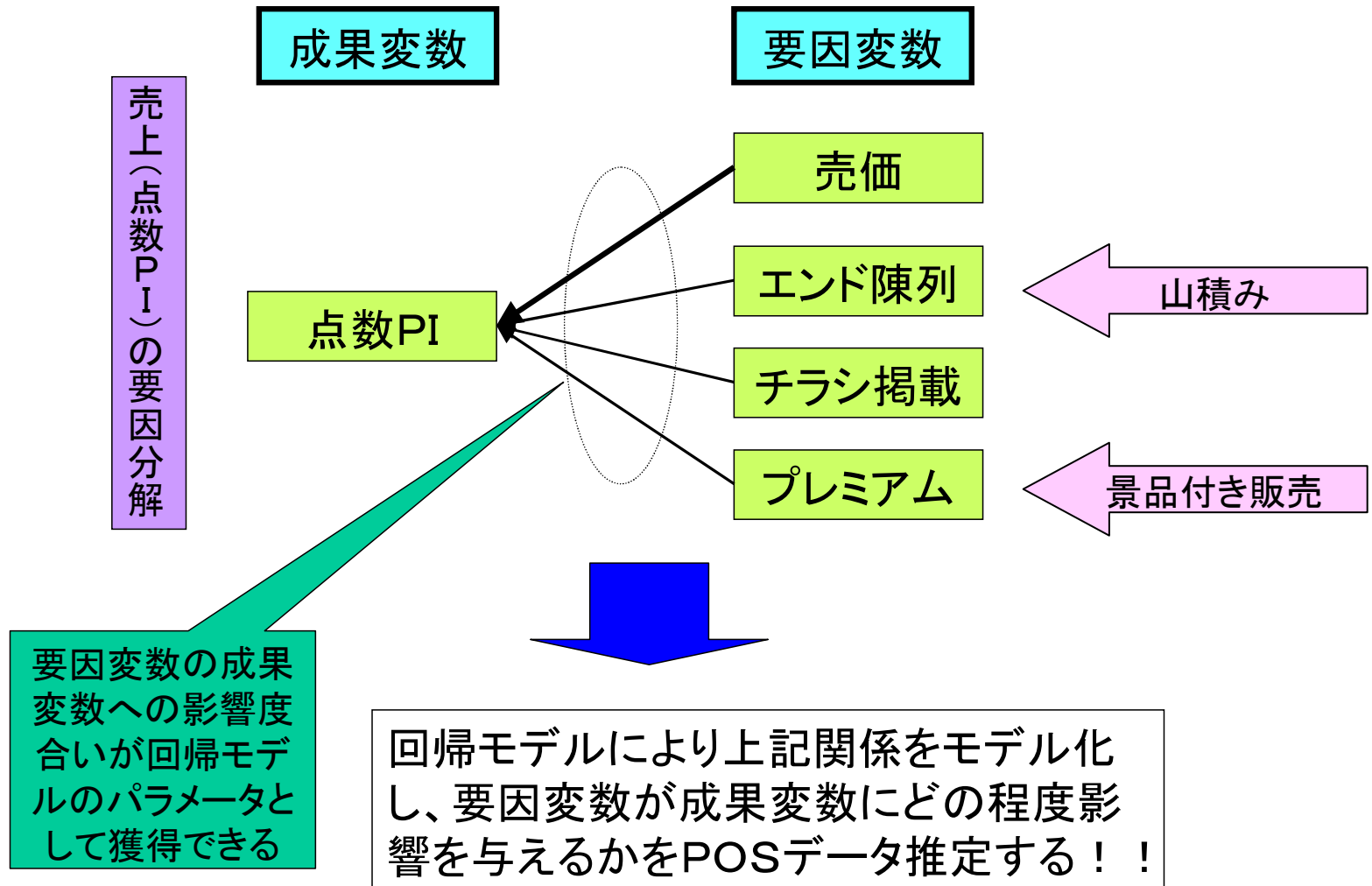
この時, 確率的誤差項 ε_{ni} が極値分布に従うと仮定すると

$$p_{ni} = \frac{e^{V_{ni}}}{\sum_{1 \leq j \leq I} e^{V_{nj}}} = \frac{e^{V_{ni}}}{e^{V_{n1}} + e^{V_{n2}} + L + e^{V_{nI}}}$$

マーケティング・サイエンスの領域(2)

- 消費者コミュニケーションと広告
 - コミュニケーション・プロセス, 広告のマネジメント・プロセス, **広告効果**, 広告予算の決定, メディア計画, 広告内容の決定
- プロモーション
 - プロモーション計画, **プロモーション効果の測定**, プロモーションのモデル
- 流通と営業
 - マーケティング・チャネルに関する意思決定分析, 流通に関する意思決定分析, 営業に関する意思決定分析

セールスプロモーションの効果測定 ～データからの効果測定の考え方～



マーケティング・サイエンスの領域(3)

- テストとコントロール
 - テストマーケティング, プロダクトライフサイクル, 売上予測, マーケティング・ミックス・モデル
- その他
 - 顧客プロフィール分析, 顧客資産, ブランド資産, ブランド力測定, 無形資産のモデル化, インターネット・マーケティング

モデルのイメージ



$$y = ax + \varepsilon$$

売上

価格

誤差

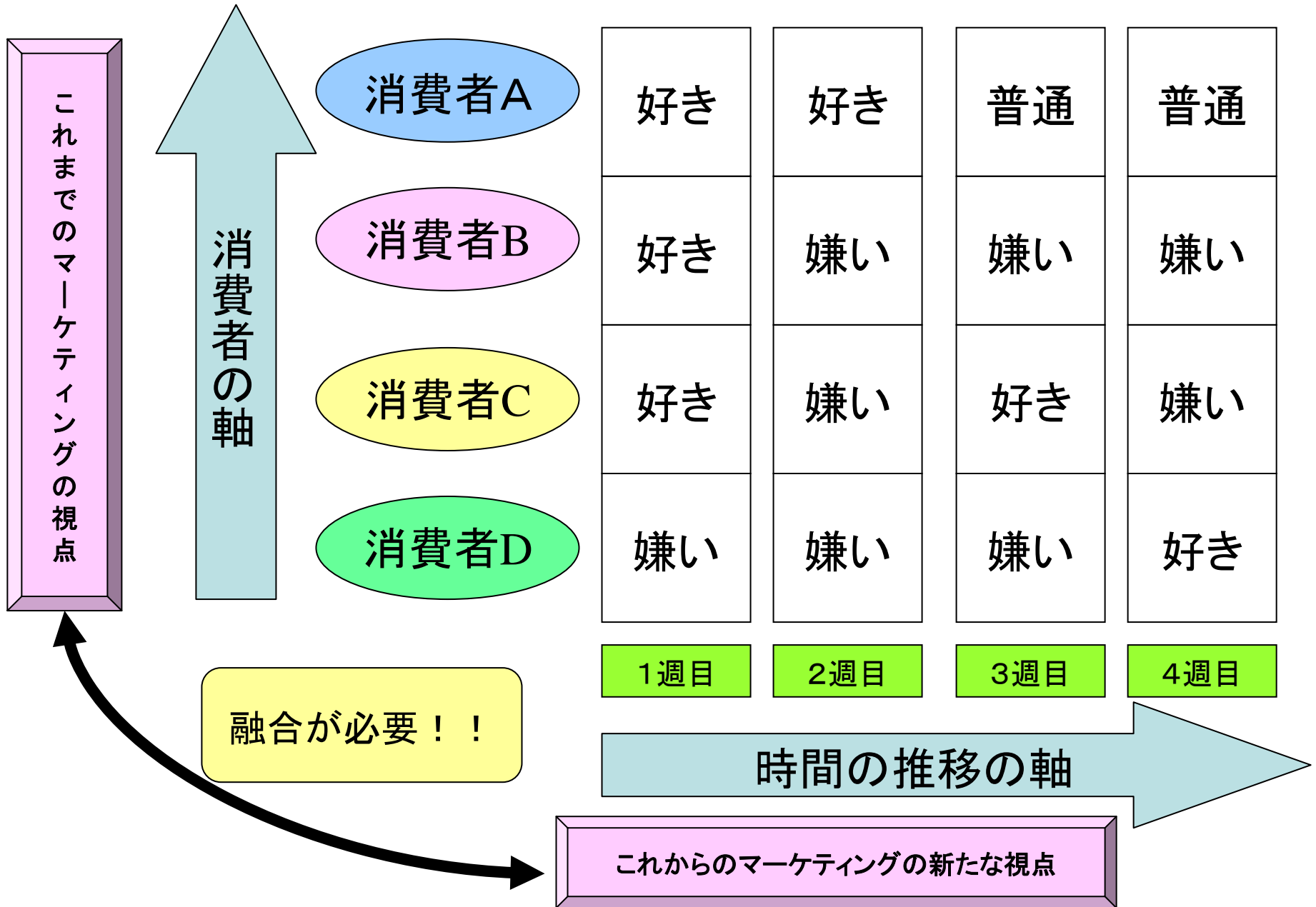
分析で用いられる手法

先端的な手法

- 記述統計
 - グラフ, 相関, 平均, 分散, 各種検定法 実験計画 等
- データマイニング的手法
 - 決定木, ニューラルネットワーク, C5.0 等
- 伝統的な多変量解析手法
 - 因子分析, 主成分分析, クラスタ分析, MDS(多次元尺度構成法), 数量化, 回帰分析, 線形判別分析 等
- 統計的なモデル
 - 非集計モデル(ロジット, プロビット), 時系列モデル, 階層ベイズモデル, 共分散構造分析, サポートベクターマシン, ベイジアンネットワーク 等

実務的な手法

個人の異質性／時間的異質性



分析やモデル化に用いるデータ

- 1次データ

- アンケート調査
- 実験計画

- 課題に対応するために消費者調査などを新たに実施することによって得られるデータを1次データという
- 別の目的によって収集され、社内外に既に存在しているデータを2次データと呼ぶ

- 2次データ(特に重要)

- POSデータ
- 顧客データ

- 購買履歴データによる顧客データベース(フリークエント・ショッパーズ・プログラム)
- ユーザー登録による顧客データベース
- キャンペーン応募データによる顧客データベース
- 会員組織化による顧客データベース

マーケティング情報の現状

- 属人的意思決定からデータに基づく意思決定への変化（“感”→“サイエンス”）
 - 属人的意思決定は組織として脆弱
 - コンピュータの進歩
 - 活用可能なデータの増大
 - マーケティング・サイエンスの進歩



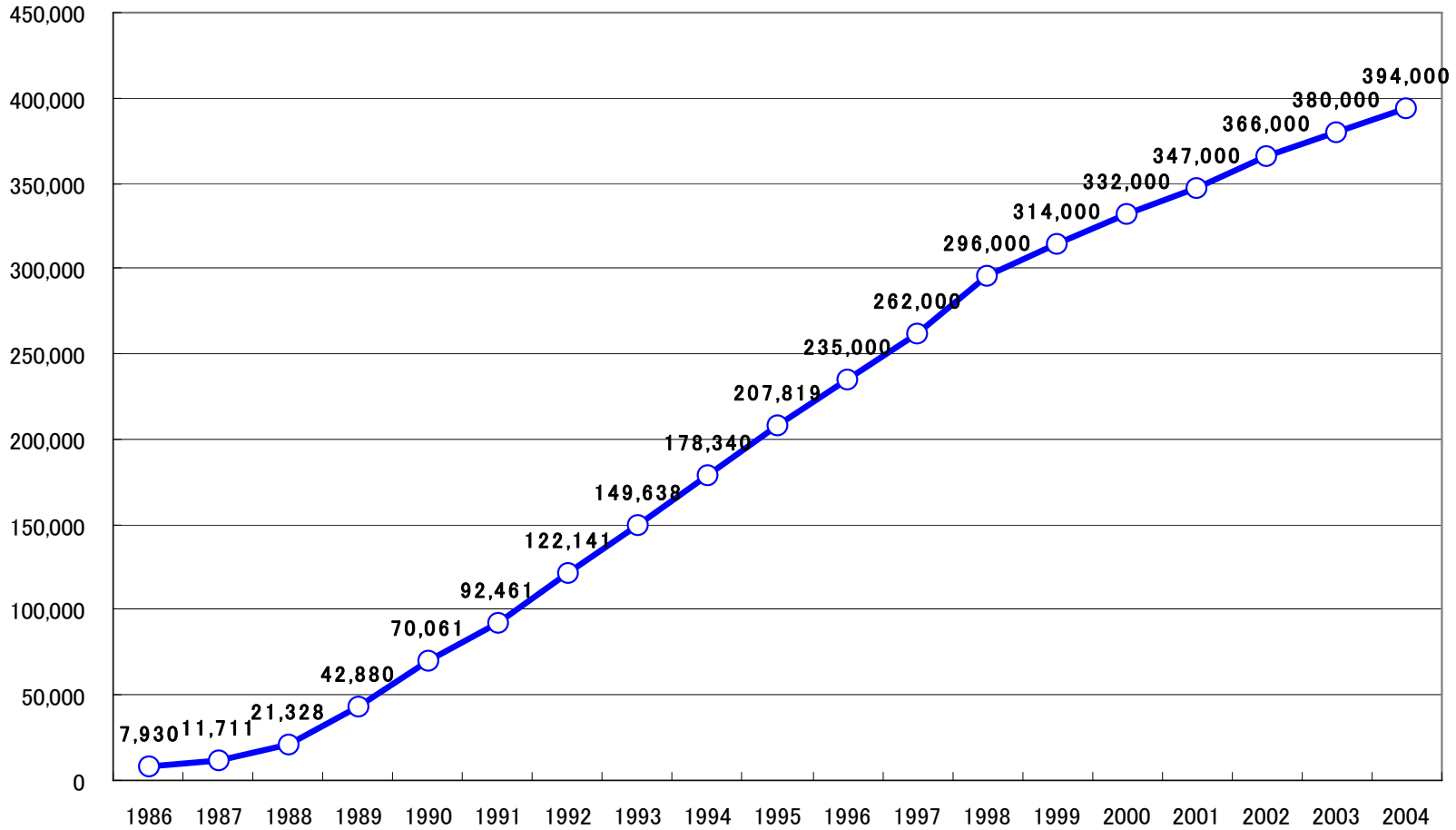
- マーケティングにおいて有用な情報を日常的かつ継続的に収集し、整理することが今日重要になってきている

小売業のデータに関する整理

- POSデータはGMS, SM, CVS及びDRUGなどの小売業においては100%獲得できる
 - いつ, 何が, 何個, いくらで売れたかについては分かる
- また, フリークエント・ショッパーズ・プログラム実施小売業の増大に伴い, POSデータの情報に誰が買ったかという情報を付加したデータが活用可能
- コーザル系のデータの整備はこれから・・・

POS導入店舗数

POS導入店舗数



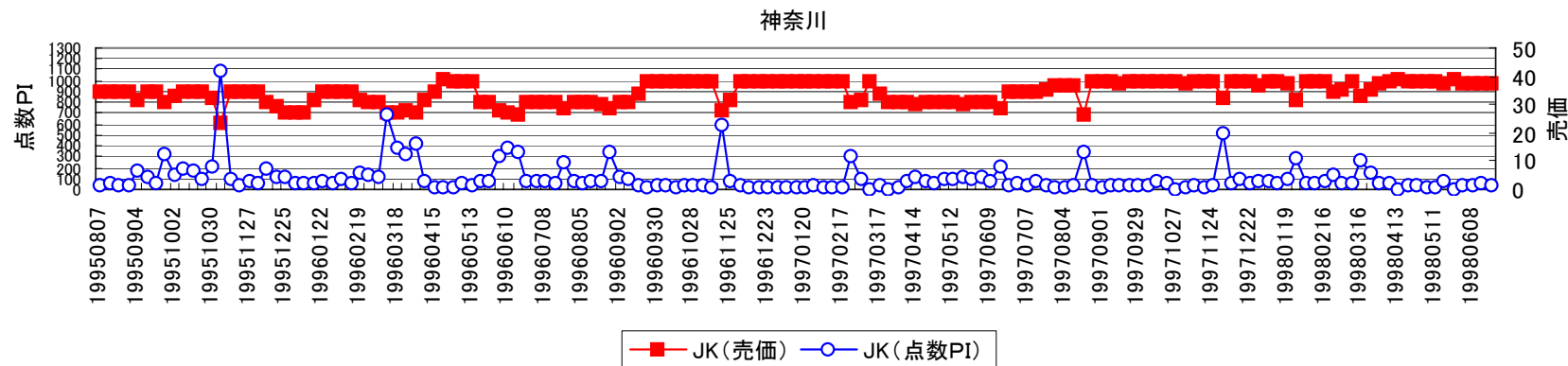
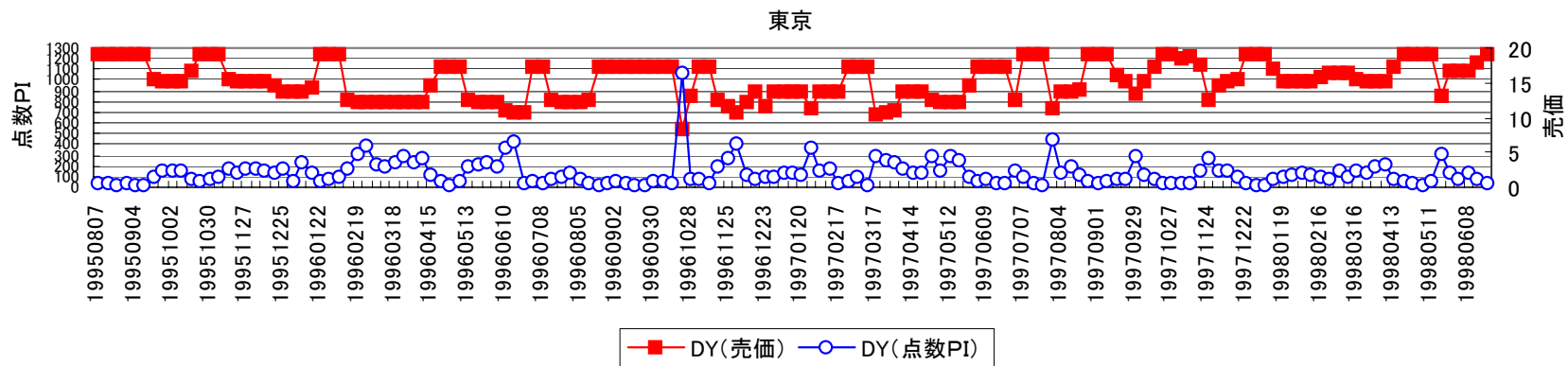
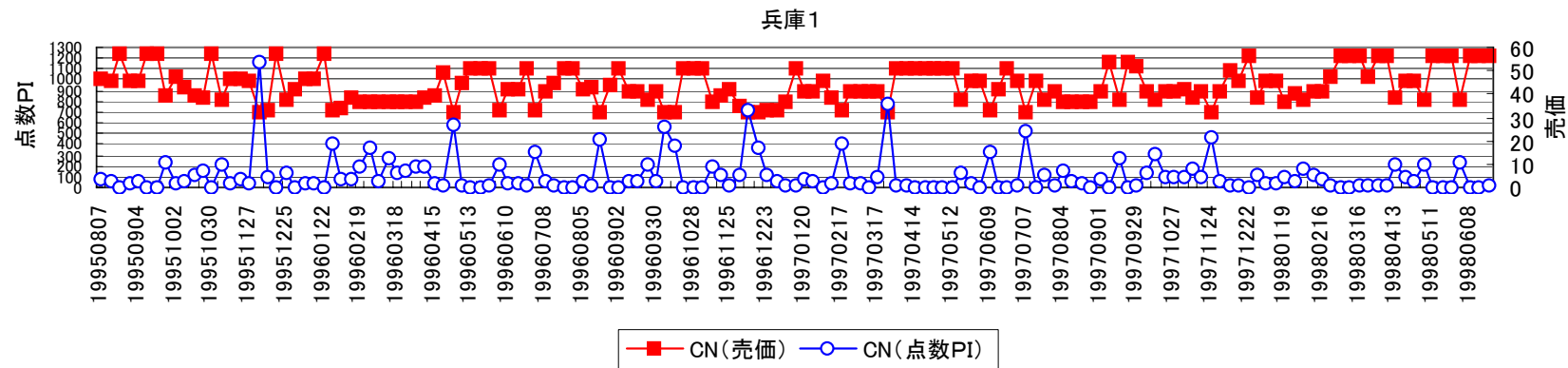
出典：流通システム開発センター

販売データ等提供機関(有料)

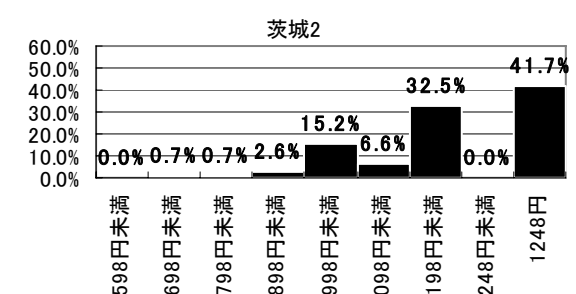
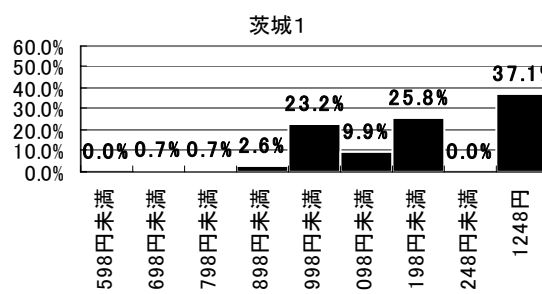
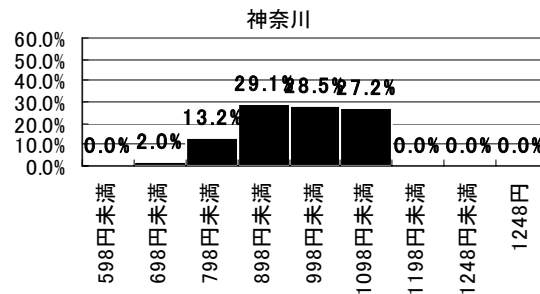
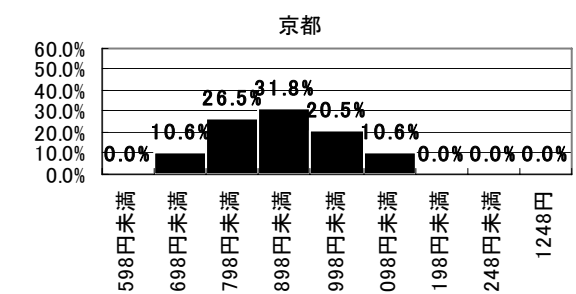
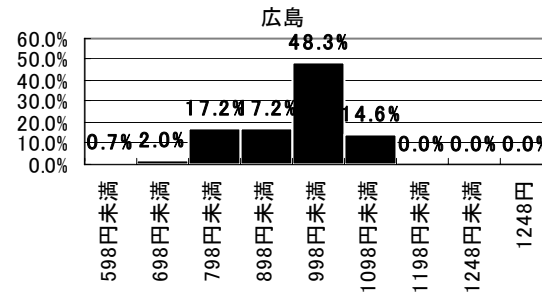
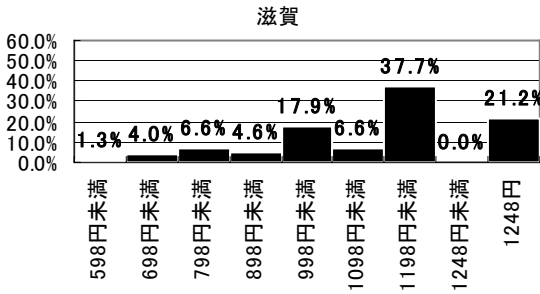
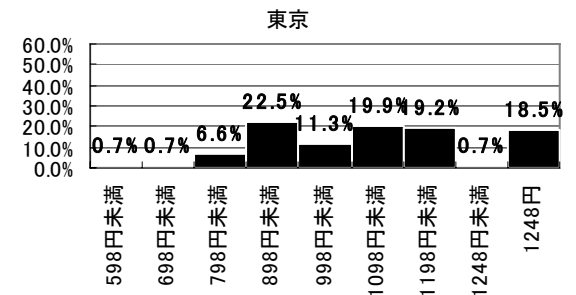
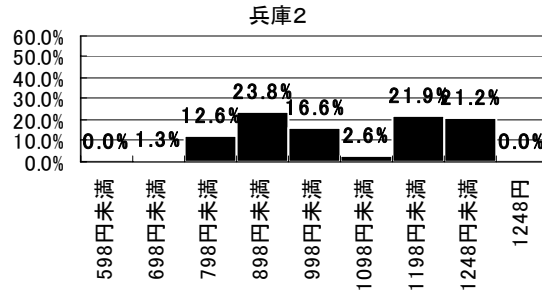
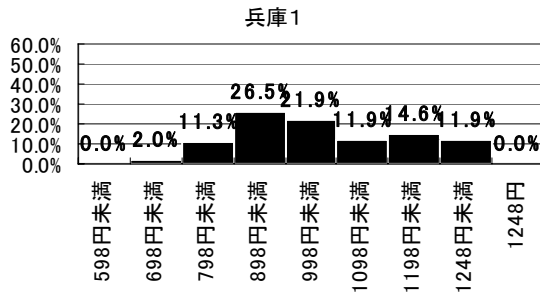
提供機関	データ種別	業態	その他
(財)流通経済研究所	POS、スキャンパネルデータ(ストアスキャン)、店頭コーザルデータ	GMS・SM、ドラッグ、コンビニ	自主収集も含めて流通システム開発センターのPOSも使用
インテージ	POS、スキャンパネルデータ(ホームスキャン)、店頭コーザルデータ	GMS・SM、ドラッグ、コンビニ、ホームセンター	日本最大のデータサービス会社
(株)ウイスダ	POS、FSPデータ	GMS・SM	IRIの日本法人だったIRJが母体となってできた組織。ASPサービス主体
(株)CSKネットワークシステムズ	POS、スキャンパネルデータ(ストアスキャン)	SM、ドラッグ、コンビニ	
凸版印刷(株)	POS	GMS・SM、コンビニ	流通システム開発センターのPOS
日本ユニシス情報システム(株)	POS	GMS・SM、コンビニ	流通システム開発センターのPOS
(財)流通システム開発センター	POS	GMS・SM、コンビニ	商品コードJANを統括する組織
日経メディアマーケティング(株)	POS	GMS・SM、コンビニ	自主収集も含めて流通システム開発センターのPOSも使用
エーシーニールセン・コーポレーション(株)	POS、スキャンパネルデータ(ホームスキャン:東急のQPR)、店頭コーザルデータ、テレビCM出稿データ	GMS・SM	世界的なデータサービス企業。日本では苦戦。東急エージェンシと共同で業務を行っている。
カスタマー・コミュニケーションズ(株)	FSPデータ、地理情報	GMS・SM	
(株)ビデオリサーチ	マス広告出稿データ、移動者購買スキャンデータ	—	業態によらず購買した商品全てのデータ
(株)NTTデータ	食マップ	—	世帯のおかずについてデータ
(株)Tカード&マーケティング	基本的にはTSUTAYAのFSPデータ(雑誌)	TSUTAYA	雑誌に購買に関するデータ

その他小売業が無償で提供するケースもある。ただし、メーカー及び卸売業に限る

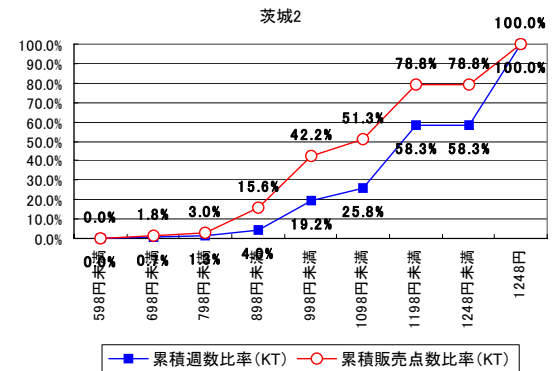
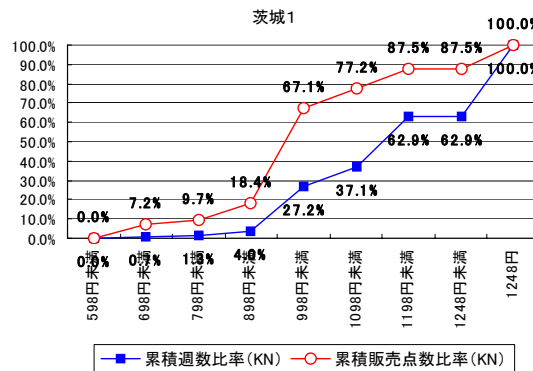
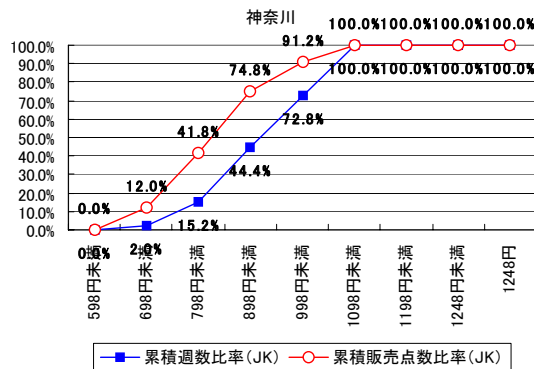
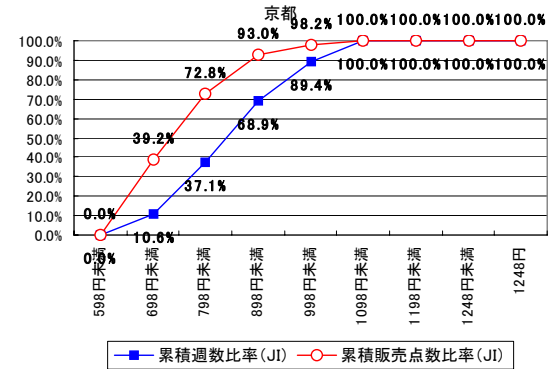
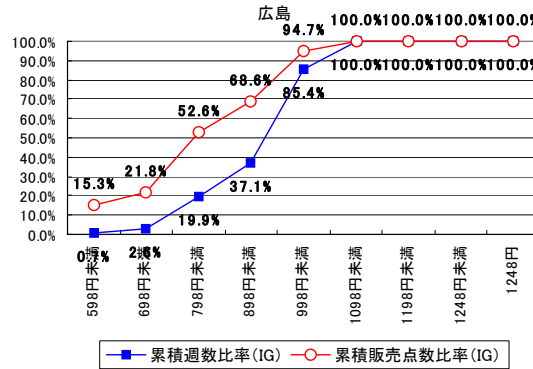
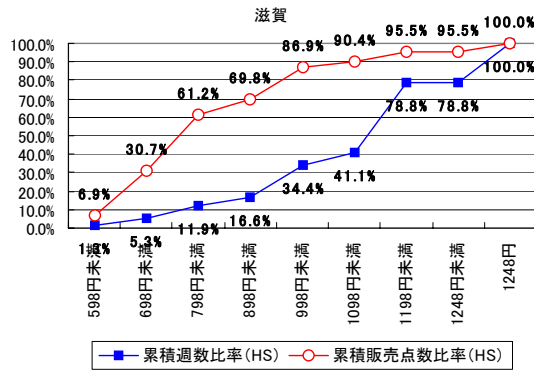
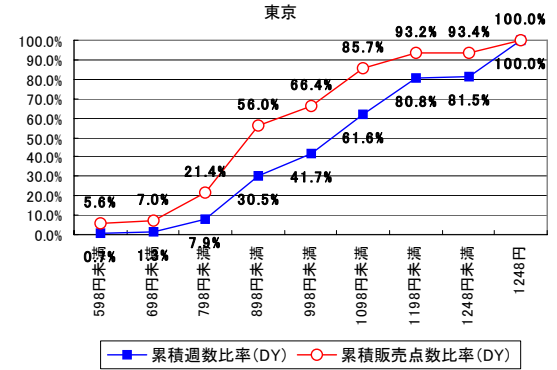
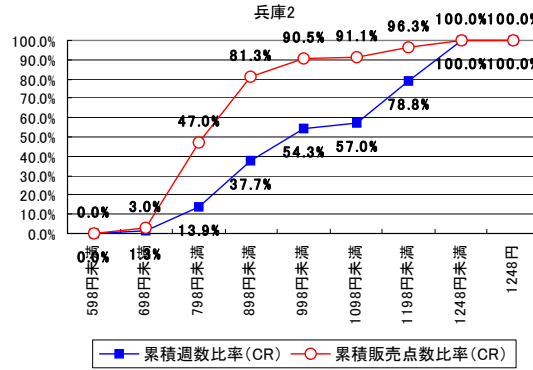
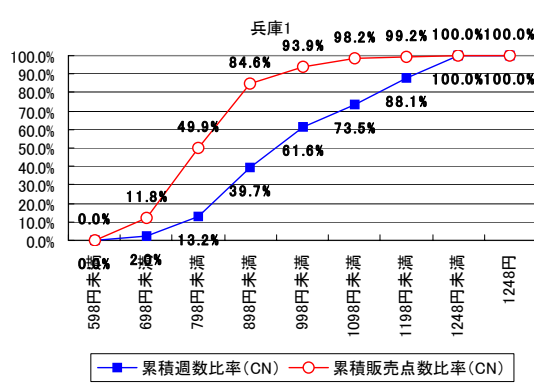
点数PIと売価の系列 (POSデータ)



価格帯別週数比率(店別)



累積販売点数比率(店別)



価格弾力性推定結果

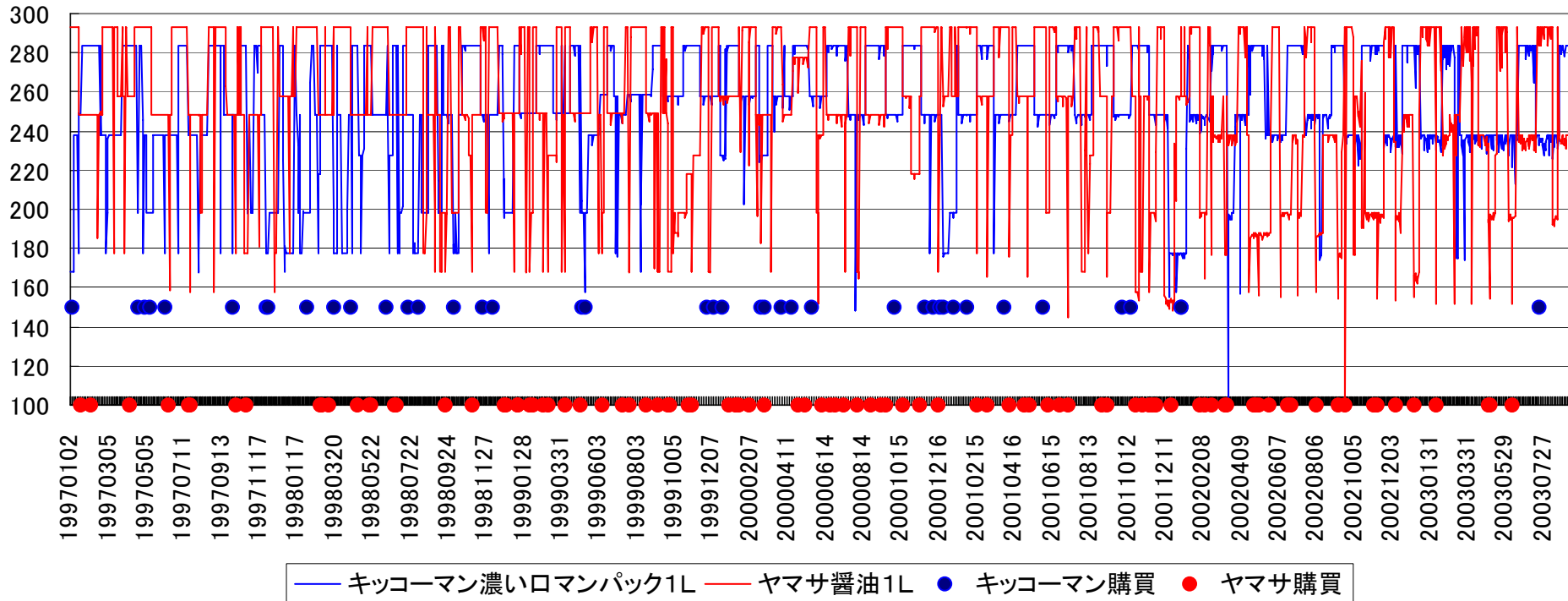
$$\log(\text{点数PI}_n) = C + \alpha \log(\text{売価}_n) + \varepsilon_n, \quad \varepsilon_n \sim N(0, \sigma^2)$$

店舗	定数項	価格弾力性	修正決定係数
兵庫1	53.0381	-7.6740	0.69
兵庫2	37.5815	-5.4343	0.65
東京	24.3339	-3.4729	0.60
滋賀	45.0059	-6.5812	0.65
広島	36.9774	-5.3027	0.60
京都	47.3750	-6.8327	0.65
神奈川	40.9428	-5.9029	0.51
茨城1	30.4673	-4.5928	0.36
茨城2	25.8047	-3.9262	0.32

ここからの知見

- 店舗によって価格政策は異なる
 - 地域の競合状況
 - チェーンの施策
 - メーカーとの関係 等
- 店舗によって消費者の価格反応は異なる
 - 消費者の価格反応は店舗の価格施策に影響され形成される
 - 売上→価格の因果は通常考えないが実はあるかもしれない(売れないから価格を下げる・・・)

醤油の例 (ID付きPOSデータ)



- この2商品を買いまわっている消費者の購買履歴
- 売価の系列はPOSデータより獲得
- 購買のデータは同一の店舗の消費者から得られたデータより獲得
- このようなデータを用いて解析することがマーケティングでは多い

データ取得に関する問題点

- データを購入できる機関が限定的である
 - 無償で提供してくれるところがあるか…
- 小商圈化の影響により、地域によって状況が大きく異なる
 - マーケティングではマイクロ・マーケティングと呼ばれる活動が脚光あびている（個人、あるいは個店、あるいは特定地域などに特化した活動）
 - データをマクロ的観点から分析するには、願わくば、同一地域で複数のお店のデータ、かつそのようなグループをなるべく数多く取得したい
- 小売業によって品揃え、価格施策が異なる

状態空間モデルを活用した 効果的価格戦略実現のための研究

日本における消費財市場の現状

- 値引率の上昇傾向

- 消費者の価格への反応はどのように変化しているか？

- 小売業の価格戦略の転換 (Hi-Lo→EDLP)

- 消費者の需要構造はどのように変化しているか？
- 小売業はどのような価格戦略を採用すべきか？

- SM・GMSでの需要量の減少傾向

- どうすれば売れるようにできるのか？

- ブランド資産の低下

- 利益のだせるブランドへはどうすれば育成できるか？

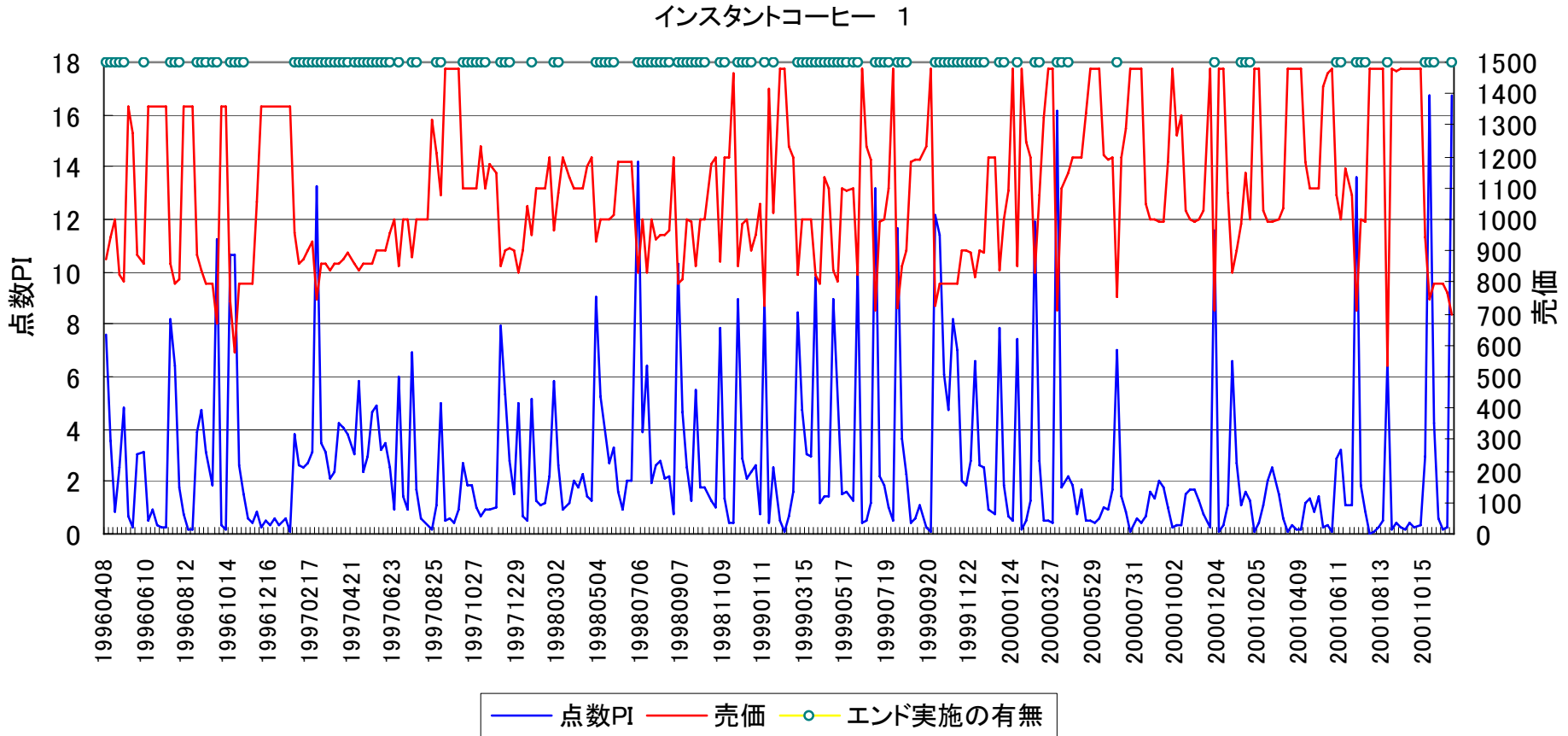
研究の目的

本研究は、個々の店舗レベルで効果的な店頭価格戦略を実現するための示唆を獲得することを目的として行っている。

具体的には、

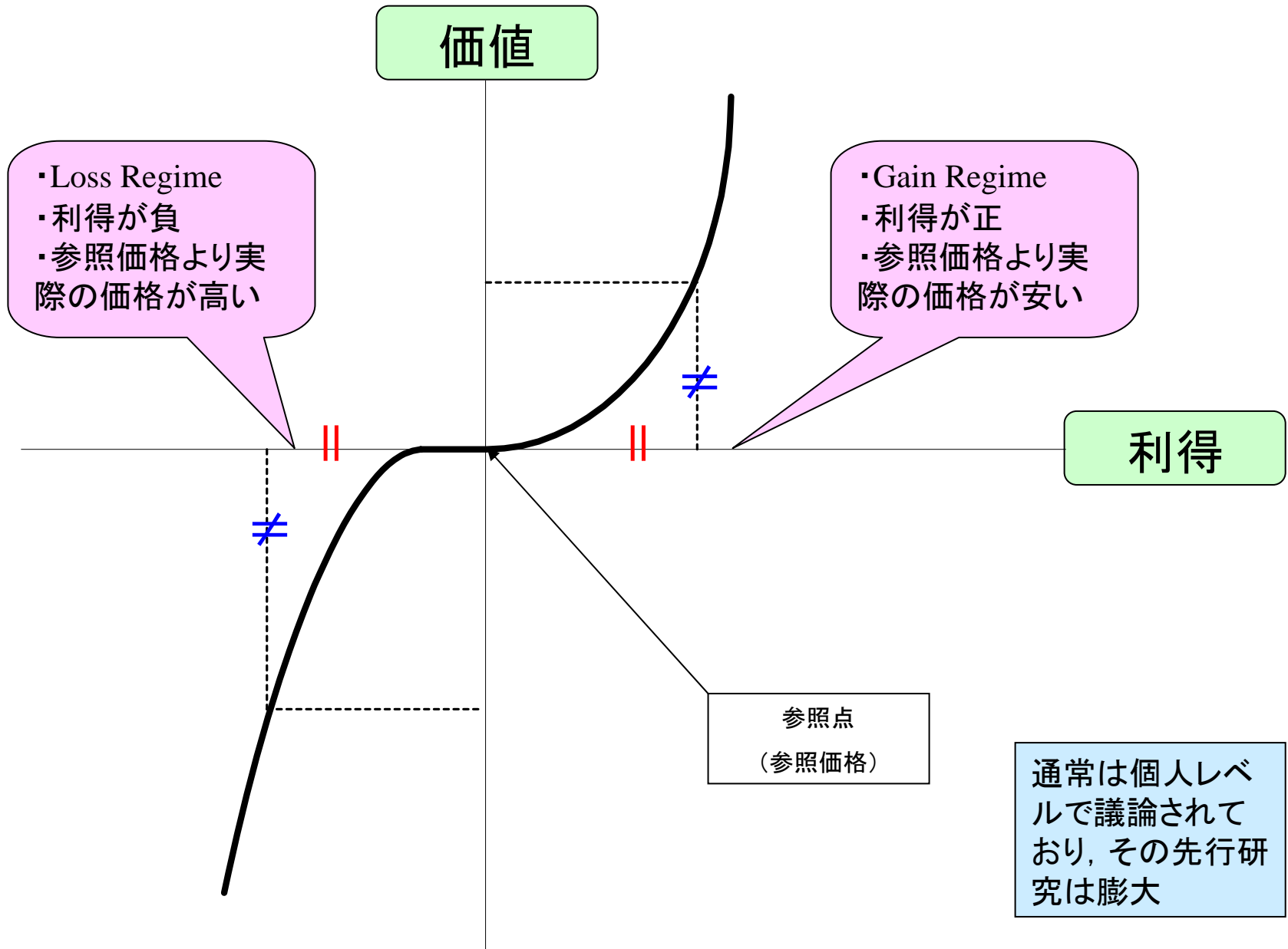
- POSデータを用いた状態空間モデルの構築
- 数値実験による価格戦略の評価を行う。

分析データ例(時系列)



- ・売価が下がったときに売上が伸びる
- ・エンドが実施された時に売上が伸びる

プロスペクト理論(概念図)



変数の説明

y_n : 第 n 週の点数PI

$$Z_n^1 = \begin{cases} \frac{AP_n - RP_n}{100}, & AP_n > RP_n \text{の時} \\ 0, & \text{その他} \end{cases}$$

Loss Regime

$$Z_n^2 = \begin{cases} \frac{RP_n - AP_n}{100}, & RP_n > AP_n \text{の時} \\ 0, & \text{その他} \end{cases}$$

Gain Regime

$$Z_n^3 = \frac{AP_n - RP_n}{100}$$

AP_n : 第 n 週の売価

RP_n : 第 n 週の店舗レベル参照価格

E_n : 第 n 週のエンド陳列実施の有無(1:実施,0:非実施)

t_n : トレンド成分 \Rightarrow 長期的な変動傾向を示す潜在変数

p_n : 定常AR成分 \Rightarrow 短期的な循環変動を示す潜在変数

$\alpha_n^1, \beta_n^1, \alpha_n^2, \beta_n^2$: 時変パラメータ

$$\varepsilon_n \sim N(0, \sigma^2)$$

モデル化のための仮定

- 売上は、当該週の売価と店舗レベル参照価格の差によって規定される
- 当該週の売価と店舗レベルの差が正の場合と負の場合で、売上に与える影響度合いが異なる
- それらの効果は時間とともに滑らかに変化する
- またそれらの効果はエンドの実施の有無で増減する

検討観測モデル

〈非対称モデル〉

$$\log(y_n) = t_n + p_n + (\alpha_n^1 + \beta_n^1 E_n) Z_n^1 + (\alpha_n^2 + \beta_n^2 E_n) Z_n^2 + \varepsilon_n$$

〈対称モデル〉

$$\log(y_n) = t_n + p_n + (\alpha_n^1 + \beta_n^1 E_n) Z_n^3 + \varepsilon_n$$

〈参照価格非考慮モデル〉

$$\log(y_n) = t_n + p_n + (\alpha_n^1 + \beta_n^1 E_n) \log(AP_n) + \varepsilon_n$$

•非対称モデル:

参照価格を基準として、売価と参照価格の差が正の場合と負の場合でその影響が異なると仮定したモデル

•対称モデル:

参照価格を基準として、売価と参照価格の差が正の場合と負の場合でその影響が同じであると仮定したモデル

•参照価格非考慮モデル:

参照価格を考慮しないモデル

店舗レベル参照価格と価格変数の定義

- ここでは下記に示すような3つの変数を、店舗レベル参照価格として検討した

$$RP_n^1 = \left\{ (1-a)AP_{n-1} + aRP_{n-1}^1 \right\}^{E_{n-1}} \left\{ (1-b)AP_{n-1} + bRP_{n-1}^1 \right\}^{1-E_{n-1}}$$

ただし、 $0 < a, b < 1$ とする。 E_{n-1} は前週の特別陳列実施の有無

$$RP_n^2 = (1-a)AP_{n-1} + aRP_{n-1}^2$$

ただし、 $0 < a < 1$ とする。

$$RP_n^3 = \frac{\sum_{i=1}^{n-1} unit_i \times AP_i}{\sum_{j=1}^{n-1} unit_j}$$

ただし、 $unit_j$ は第*i*週の販売点数とする

検討モデルの種類

- 本研究では、観測モデルのタイプと店舗レベル参照価格のタイプの組合せで下記に示す7つのモデルを比較・検討した

モデルのタイプ	店舗レベル参照価格のタイプ				モデルID
	RP1	RP2	RP3	無	
非対称	1	2	3	—	
対象	4	5	6	—	
参照価格無	—	—	—		7

モデルのタイプ	店舗レベル参照価格のタイプ				モデルパラメータ数
	RP1	RP2	RP3	無	
非対称	10	9	8	—	
対象	8	7	6	—	
参照価格無	—	—	—		6

モデル比較(AIC)

インスタント・コーヒー1

モデルのタイプ	店舗レベル参照価格のタイプ			
	RP1	RP2	RP3	無
非対称	538.45	536.48	564.00	—
対称	581.10	579.19	619.33	—
参照価格無	—	—	—	709.08

醤油1

モデルのタイプ	店舗レベル参照価格のタイプ			
	RP1	RP2	RP3	無
非対称	440.36	446.70	438.08	—
対称	453.66	468.97	472.75	—
参照価格無	—	—	—	592.93

インスタント・カレー

モデルのタイプ	店舗レベル参照価格のタイプ			
	RP1	RP2	RP3	無
非対称	419.67	418.02	520.65	—
対称	445.90	445.67	528.08	—
参照価格無	—	—	—	611.06

サラダ油・天ぷら油

モデルのタイプ	店舗レベル参照価格のタイプ			
	RP1	RP2	RP3	無
非対称	188.36	196.17	196.52	—
対称	194.86	201.41	196.40	—
参照価格無	—	—	—	258.00

インスタント・コーヒー2

モデルのタイプ	店舗レベル参照価格のタイプ			
	RP1	RP2	RP3	無
非対称	470.15	468.15	475.71	—
対称	515.15	513.15	539.42	—
参照価格無	—	—	—	666.99

醤油2

モデルのタイプ	店舗レベル参照価格のタイプ			
	RP1	RP2	RP3	無
非対称	480.76	479.68	509.04	—
対称	505.72	505.78	526.50	—
参照価格無	—	—	—	507.15

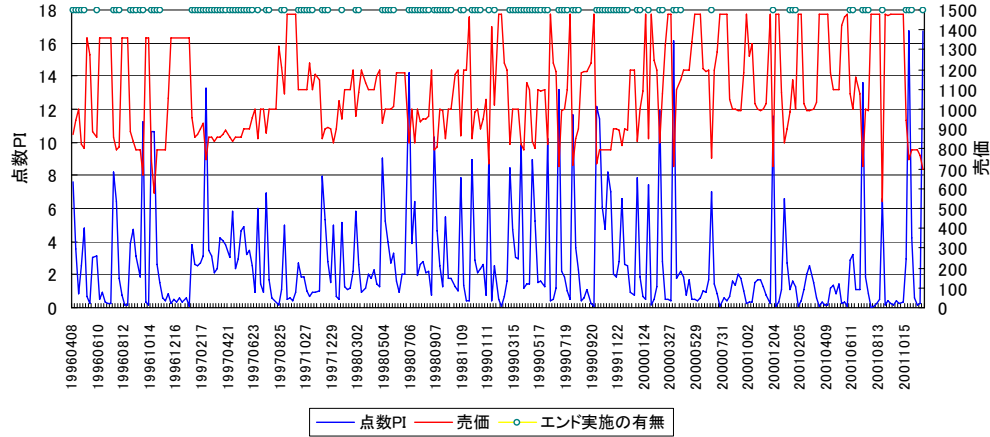
烏龍茶

モデルのタイプ	店舗レベル参照価格のタイプ			
	RP1	RP2	RP3	無
非対称	349.75	355.12	326.00	—
対称	355.77	364.42	337.05	—
参照価格無	—	—	—	458.06

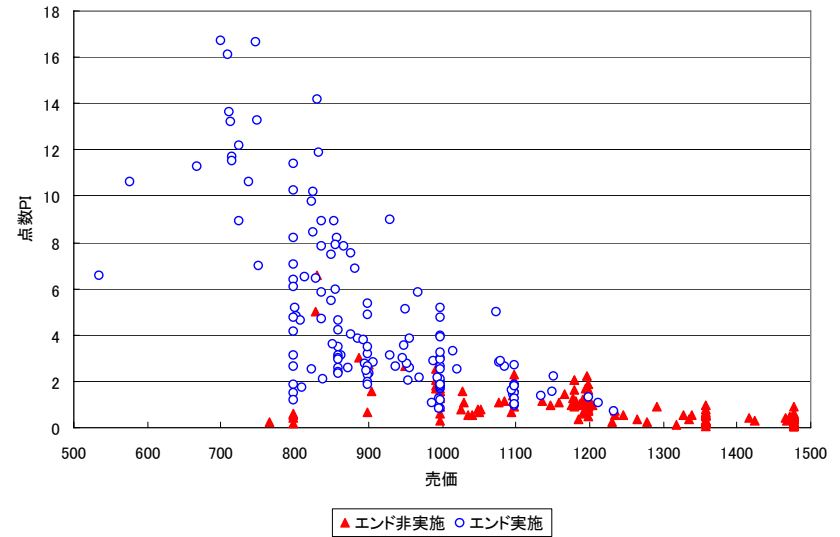
- ・いずれの商品でも、店舗レベル参照価格を考慮したモデルが採択される
- ・また、売価と店舗レベルの差が正の場合と負の場合でその影響度合いが異なると仮定したモデルが全ての商品で採択されている
- ・店舗レベル参照価格の生成メカニズムは、商品毎に異なっている

インスタントコーヒー1

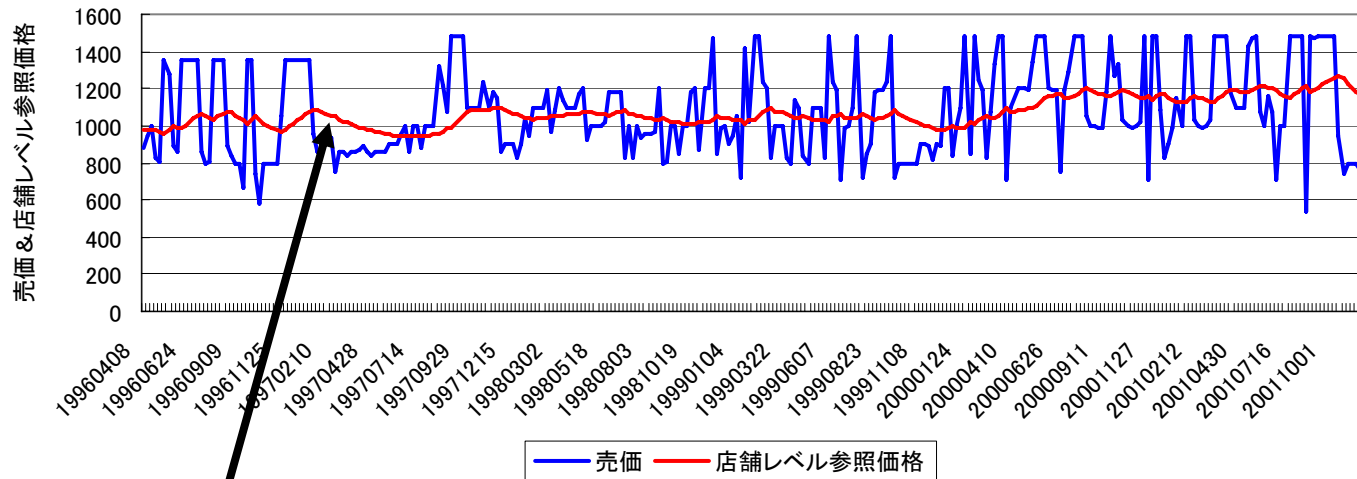
インスタントコーヒー 1



インスタントコーヒー 1



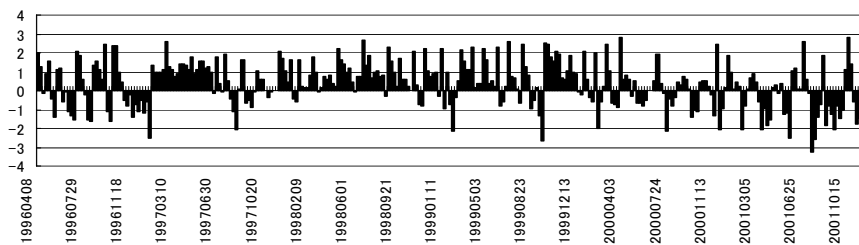
インスタント・コーヒー1



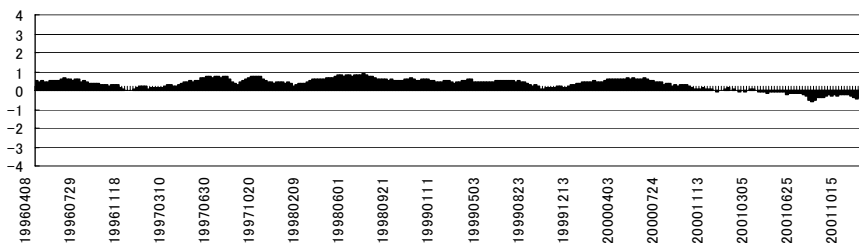
$$RP_n^2 = 0.05 \cdot AP_{n-1} + 0.95 \cdot RP_{n-1}$$

時系列の分解 (インスタントコーヒー1)

Log(点数PI)(観測値: インスタント・コーヒー1)

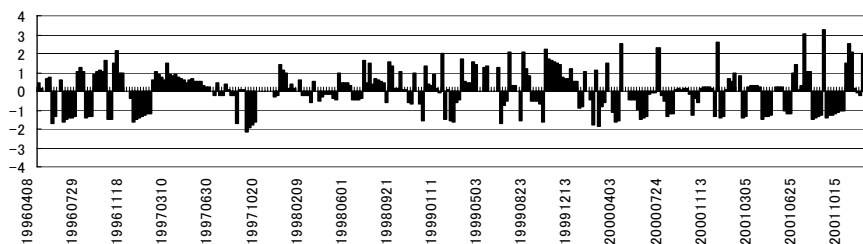


tn+pn(インスタント・コーヒー1)



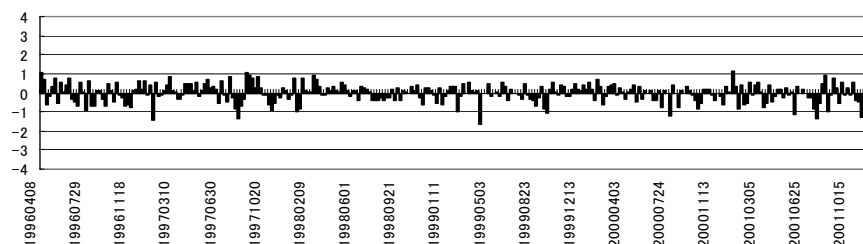
トレンド+定常AR成分低下

回帰成分(インスタント・コーヒー1)



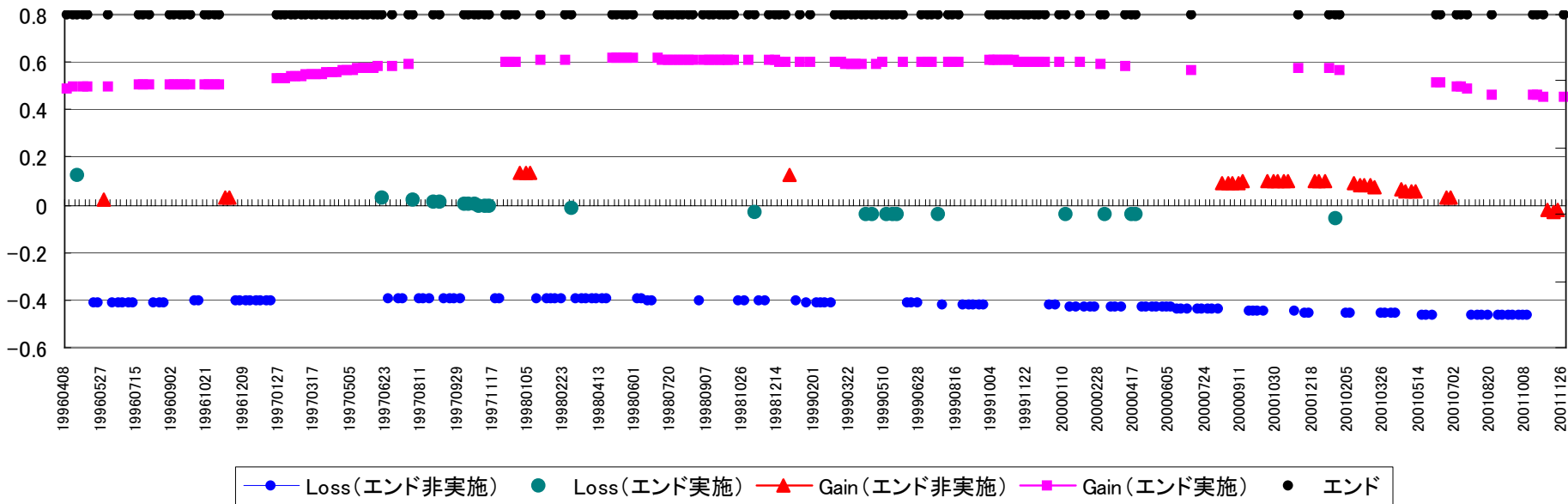
Gain Regime
Loss Regime

誤差(インスタント・コーヒー1)



インスタントコーヒー1 (時変係数: 平滑化)

価格効果(インスタントコーヒー1)



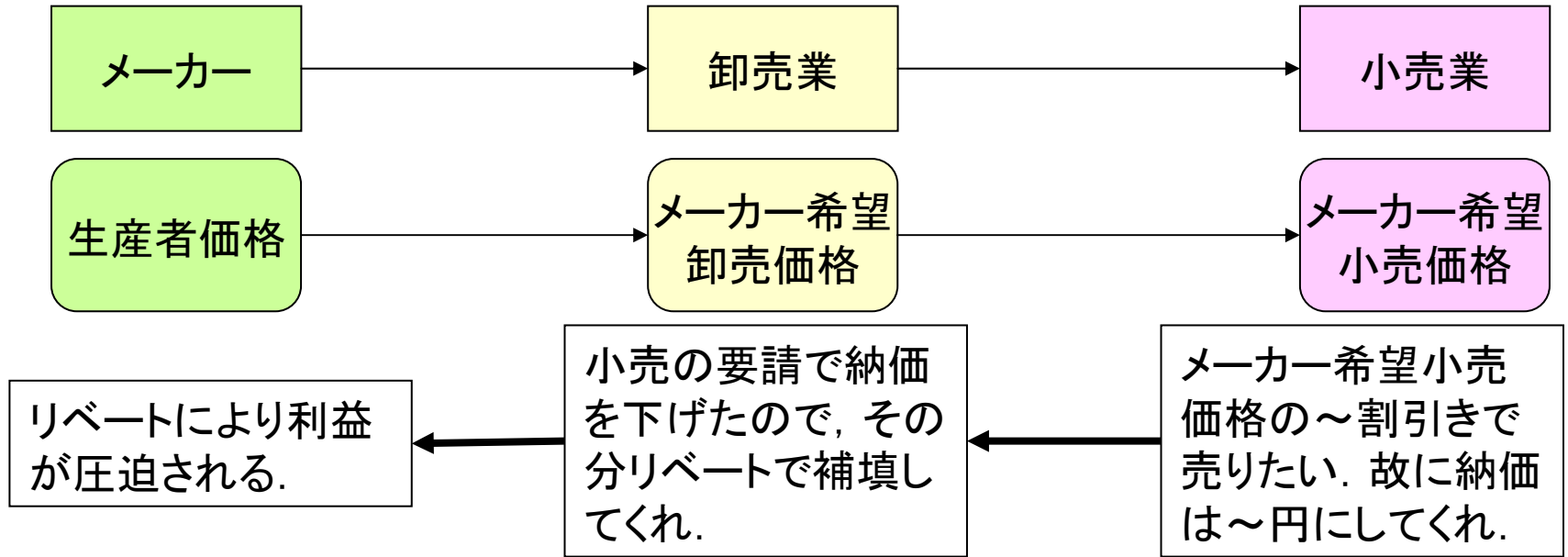
■ エンドでの値引き ($\alpha_n^2 + \beta_n^2$)

▲ 定番での値引き (α_n^2)

● 比較的高価格でのエンド実施 ($\alpha_n^1 + \beta_n^1$)

● 定番販売 (α_n^1)

取引の構造



メーカーはリベートを抑制する価格戦略の実現が必要

- 価格差を現況よりも小さくしたHi-Lo型
- EDLP
- EDMP

メーカーの粗利益算定法(例)

生産者価格

メーカー希望
卸売り価格

メーカー希望
小売価格

$$0.713x\text{円} \xleftarrow{5\%} 0.750x\text{円} \xleftarrow{25\%} x\text{円}$$

$$1054\text{円} \xleftarrow{\quad} 1109\text{円} \xleftarrow{\quad} 1478\text{円}$$

—

828円

30%引きで売って、粗利を20%とりたいから納価を下げてください。

— 281円

リベートで補填

例えばこの額をメーカーと小売業で折半するなどする

(メーカーの純売上) =

(販売個数) × { (生産者価格) - (負担率) × (1個あたり補填額) }

売上増大のための 効果的店頭価格実現の方向性

① 参照価格を低下させない価格戦略をとる

- 過度な値引きをせずともAP-RPを負で大きくできる
- すなわち、高い価格でもある程度売上が獲得し易くなる

② 店頭価格を低く設定する

- コストを抑え、低価格を実現することで、AP-RPを負で大きくできる
- しかし、RPもそれに伴って低いレベルに推移する

現実的には、①の戦略を指向することが肝要である

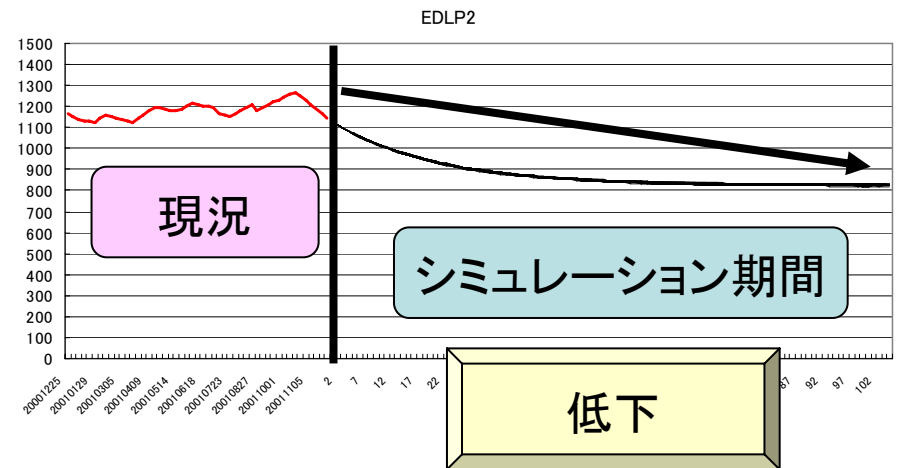
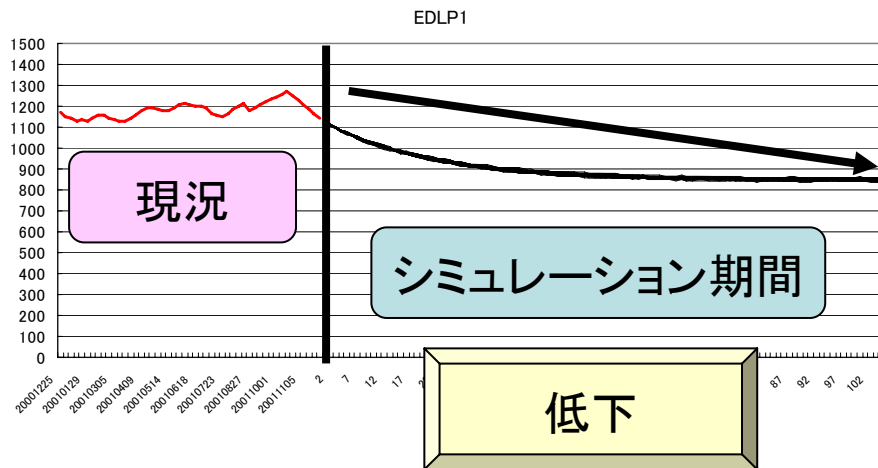
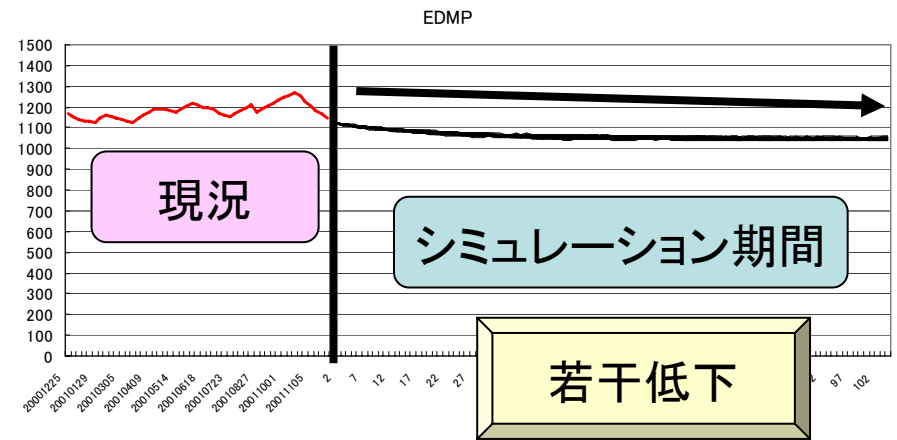
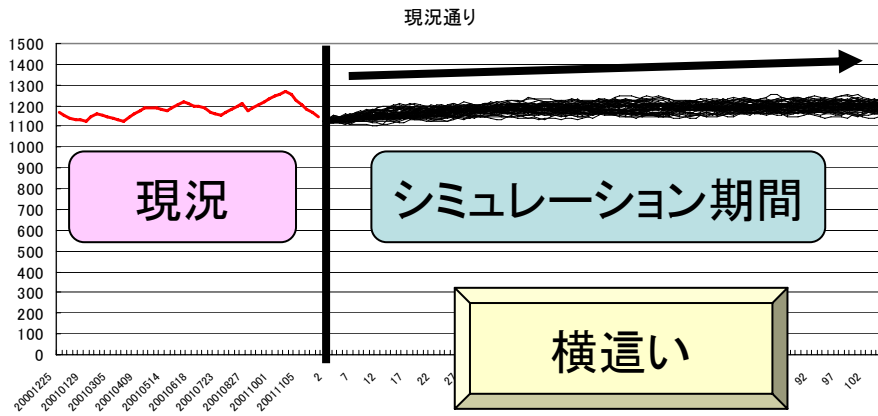
設定したシナリオ

シナリオ NO	価格戦略	値引き額の メーカー側の 補填率	非エンド時		エンド時		生産者 価格	メーカー 希望卸 売価格	メーカー 希望小 売価格
			最低売 価	最大売 価	最低売 価	最大売 価			
1	現況通り	0.80	1,098	1,478	998	1,198	983	1,035	1,478
2	EDMP	0.45	998	1,098	998	1,098	866	911	1,098
3	EDLP1	0.40	798	898	798	898	768	808	898
4	EDLP2	0.40	798	848	798	848	741	780	848

- シナリオ1は、現況通りの価格戦略を採用しつづけた場合を想定
- シナリオ2は、中間的な価格で価格差をつけない価格戦略を採用した場合を想定
- シナリオ3は、比較的低い価格設定の価格戦略を採用した場合を想定
- シナリオ4は、シナリオ3より最低売価と最高売価の差が小さい場合を想定

この設定はあくまでも例であり、自社の問題に実際に適応させるためには、この設定を変更することが必要である

各シナリオ下での店舗レベル参照価格の推移

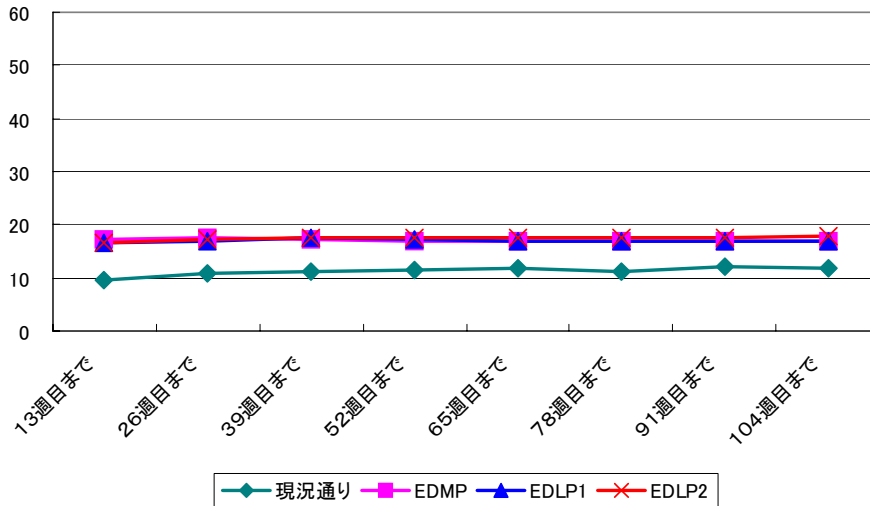


・店舗レベル参照価格の観点からは、現況通りの価格戦略を採った場合は横這い、EDMPを採用した場合は若干低下、EDLPを採用した場合は大きく低下する

・前述したように、消費者は当期の価格と参照価格の差により購買する傾向が強いため、EDLPを採用した場合は中長期的にみると、困難な事態を招く可能性がある

各シナリオ下での平均販売点数

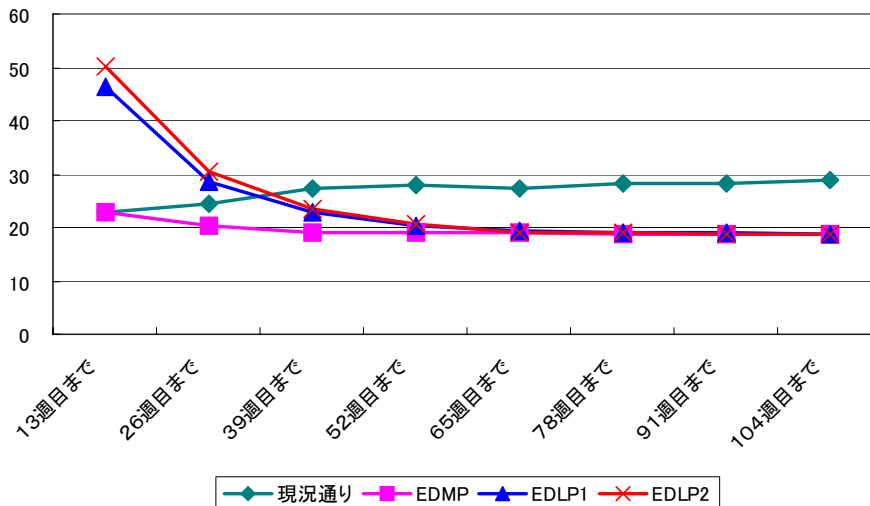
平均週販(エンド非実施)



エンド非実施時

- ・現況通りの価格戦略を採用しつづけた場合の販売点数の推移が最も低い
- ・その他の価格戦略は差が小さい

平均週販(エンド実施)

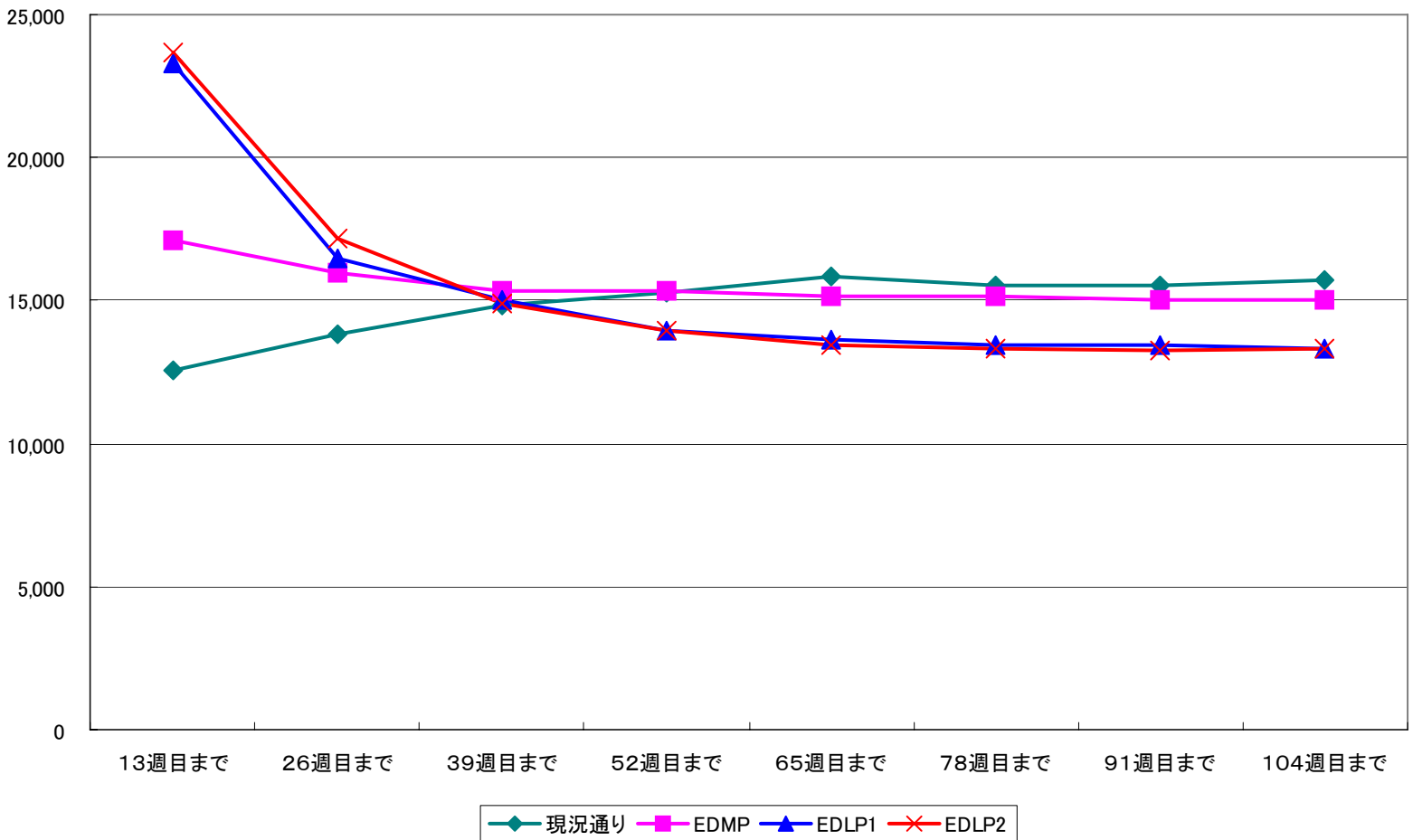


エンド実施時

- ・価格戦略変更当初はEDLP型の価格戦略は売上が増大するが、その後EDMP型の価格戦略と同程度まで低下する
- ・現況通りの価格戦略をとりつづけた場合の販売点数のレベルが最も高いレベルにある

各シナリオ下での平均週当たりメーカー純売上

メーカー純売上(売上-販促費)(シミュレーション結果の平均)



- 本シナリオの設定においては、価格戦略変更半年までは、EDLP型の価格戦略が利益を増大させる可能性がある
- しかし、その後は現況通りかあるいはEDMP型の価格戦略の方が、利益を向上させえることが示唆されている

本数値実験からの知見

- EDLP型の価格戦略は導入当初は売上・利益とも出しやすい
 - しかし、中長期的にみると、参照価格の低下に伴い、売上が伸びなくなる可能性が高い
- 一方EDMP型及び現況型の価格戦略は導入当初こそ、EDLP型の戦略に売上・利益とも劣るが、参照価格がそれほど低下しないため、中長期的には売上が出しやすい
 - しかし、現況通りの価格戦略をとり続ける限りにおいては、販促費問題を解決できない



本シミュレーションの結果からは、EDMP型の価格戦略に変更していくことが望ましいことが示唆される

- 設定が変われば、本結論は変わる可能性はあるが、このフレームで考えれば、より現実に即したように将来の戦略変更を評価できる。このことが重要。

まとめ

- マーケティングについて
- 小売業のデータについて
- 状態空間モデルを活用した効果的価格戦略実現のための研究の紹介

Thank you!

佐藤忠彦: 筑波大学大学院ビジネス科学研究科

Contact: sato@gssm.otsuka.tsukuba.ac.jp

**参考：状態空間モデルを活用した
効果的価格戦略実現のための研究
のモデル**

状態空間モデル

[システムモデル]

$$x_n = Fx_{n-1} + Gv_n, \quad v_n \sim N(0, Q)$$

[観測モデル]

$$y_n = H_n x_n + w_n, \quad w_n \sim N(0, R)$$

•観測モデル:

時系列 y_n が観測される仕組みを表現するモデルであり, x_n はその回帰係数.

•システムモデル:

回帰係数 x_n の時間的な変化の様子を表現するモデル

システムモデル(非対称モデルのケース)

平滑化事前分布の導入

$$t_n = t_{n-1} + v_n^1, \quad v_n^1 \sim N(0, \tau_1^2)$$

$$p_n = a_1 p_{n-1} + a_2 p_{n-1} + v_n^2, \quad v_n^2 \sim N(0, \tau_2^2)$$

$$\alpha_n^1 = \alpha_{n-1}^1 + v_n^3, \quad v_n^3 \sim N(0, \tau_3^2)$$

$$\beta_n^1 = \beta_{n-1}^1 + v_n^4, \quad v_n^4 \sim N(0, \tau_4^2)$$

$$\alpha_n^2 = \alpha_{n-1}^2 + v_n^5, \quad v_n^5 \sim N(0, \tau_5^2)$$

$$\beta_n^2 = \beta_{n-1}^2 + v_n^6, \quad v_n^6 \sim N(0, \tau_6^2)$$

- 平滑化事前分布を導入することにより, 回帰係数の時間変化を実現できる
- ベイズ統計の言葉で言えば t, p, α, β はPriorということになる
- システムノイズの分散 τ はhyper Parameterになる
- 本モデルのフレームでは τ は最尤法により推定される(**Empirical Bayes**)

状態空間表現(非対称モデルのケース)

<システムモデル>:

$$\mathbf{x}_n = \mathbf{F} \cdot \mathbf{x}_{n-1} + \mathbf{G} \cdot \mathbf{v}_n, \quad \mathbf{v}_n \sim N(0, \mathbf{Q})$$

$$\begin{bmatrix} t_n \\ p_n \\ p_{n-1} \\ \alpha_n^1 \\ \beta_n^1 \\ \alpha_n^2 \\ \beta_n^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_1 & a_2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_{n-1} \\ p_{n-1} \\ p_{n-2} \\ \alpha_{n-1}^1 \\ \beta_{n-1}^1 \\ \alpha_{n-1}^2 \\ \beta_{n-1}^2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_n^1 \\ v_n^2 \\ v_n^3 \\ v_n^4 \\ v_n^5 \\ v_n^6 \end{bmatrix},$$

$$\mathbf{Q} = \begin{pmatrix} \tau_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \tau_2^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau_3^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tau_4^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_5^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \tau_6^2 \end{pmatrix}$$

<未知パラメータ>:

$$\theta = \{a, b, a_1, a_2, \tau_1^2, \tau_2^2, \tau_3^2, \tau_4^2, \tau_5^2, \tau_6^2\}$$

<観測モデル>:

$$y_n = \mathbf{H}_n \mathbf{x}_n + \varepsilon_n, \quad \varepsilon_n \sim N(0, \sigma^2)$$

$$\mathbf{H}_n = \begin{bmatrix} 1 & 1 & 0 & Z_n^1 & E_n \cdot Z_n^1 & Z_n^2 & E_n \cdot Z_n^2 \end{bmatrix}$$

モデルパラメータの推定は、a,bはグリッドサーチ、その他のパラメータは数値的最適化により、入れ子型の形で同時に推定する

Kalman Filter / 固定区間平滑化

<一期先予測>:

$$\mathbf{x}_{n|n-1} = \mathbf{F}\mathbf{x}_{n-1|n-1}$$

$$\mathbf{V}_{n|n-1} = \mathbf{F}\mathbf{V}_{n-1|n-1}\mathbf{F}^t + \mathbf{G}\mathbf{Q}\mathbf{G}^t$$

<フィルタ>:

$$\mathbf{K}_n = \mathbf{V}_{n|n-1}\mathbf{H}_n^t \left(\mathbf{H}_n\mathbf{V}_{n|n-1}\mathbf{H}_n^t + \mathbf{R} \right)^{-1}$$

$$\mathbf{x}_{n|n} = \mathbf{x}_{n|n-1} + \mathbf{K}_n \left(\mathbf{y}_n - \mathbf{H}_n\mathbf{x}_{n|n-1} \right)$$

$$\mathbf{V}_{n|n} = \left(\mathbf{I} - \mathbf{K}_n\mathbf{H}_n \right) \mathbf{V}_{n|n-1}$$

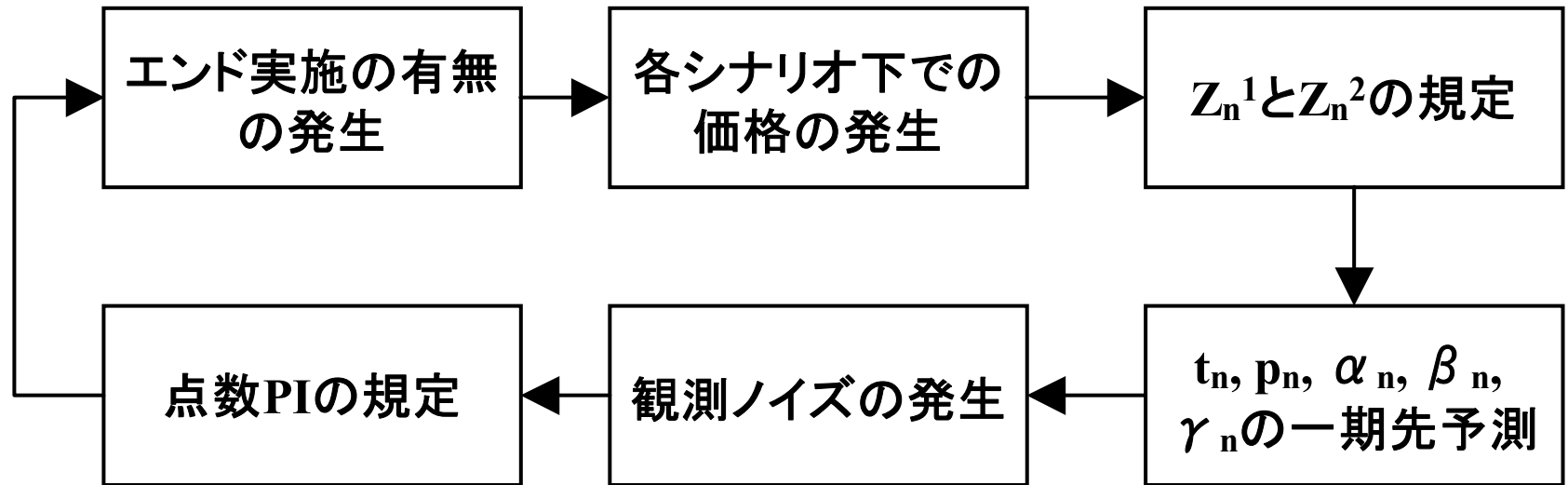
<固定区間平滑化>:

$$\mathbf{A}_n = \mathbf{V}_{n|n}\mathbf{F}^t\mathbf{V}_{n+1|n}^{-1}$$

$$\mathbf{x}_{n|N} = \mathbf{x}_{n|n} + \mathbf{A}_n \left(\mathbf{x}_{n+1|N} - \mathbf{x}_{n+1|n} \right)$$

$$\mathbf{V}_{n|N} = \mathbf{V}_{n|n} + \mathbf{A}_n \left(\mathbf{V}_{n+1|N} - \mathbf{V}_{n+1|n} \right) \mathbf{A}_n^t$$

数値実験の手順



- いくつかの価格変化のシナリオ設定し, 上記に示す手順でシミュレーションを行う
- 各設定50ケースずつ系列を発生させる
- 参照価格の推移, 売上, 粗利益から各シナリオを比較する