

# 価格モデルを考慮した 牛乳の需要予測に関する研究

---

大同大学情報学部情報システム学科 守田真矢  
指導教員 柴田慎一

# 目次

---

- 研究背景
- 研究目的
- 関連研究
- 提案手法
- 使用データ
- 実験概要
- 実験結果
- 考察
- まとめ
- 今後の課題
- 参考文献

# 研究背景 (1/3)

今日のスーパーマーケット, コンビニエンスストアでは、POSシステム (Point Of Sales System) という 商品の販売情報の管理システム が利用されている。 [1]

POSシステムとは、

1982年…セブンイレブンを筆頭に導入 [2]

1989年…消費税制度がPOSシステム普及に拍車をかける



[1] 生田目崇, “マーケティング分野におけるデータ分析の進展”, 2007.

[2] 阿部誠, 近藤文代, “マーケティングの科学-POSデータの解析-“, 2005.

# 研究背景 (2/3)



## POSシステムの機能

商品名 価格 数量などの

**販売実績情報を収集**することができる。



## POSシステムの応用

需要予測 購買行動分析を行い

**販売戦略を立てる**ことができる。



# 研究背景 (3/3)

このPOSシステムを扱うにあたって需要予測はできる



商品の発注業務は全自動化されておらず、  
発注者の経験や勘で発注を行っている…

適切な発注でない場合…

**廃棄による損失**

商品の発注量多い  
廃棄による損失

トレードオフ

**欠品による購入機会の損失**

商品の発注量少ない  
欠品による機会損失

# 研究目的(1/2)

発注者の業務内容に負担が大きく、  
適切な発注業務が行われていると限らない。 [3]



スーパーマーケット, コンビニエンスストアは  
これらの業務を考慮し損失を抑えることが目的となってくる。



[3] 松村直樹[他], ” POSデータに基づく欠品時の顧客行動を考慮した小売店舗の購買シミュレーション”, 2016.

# 研究目的 (2/2)

需要予測に用いるデータとして→牛乳を使用する。



牛乳は

日配品と呼ばれる日持ちしない要冷蔵商品である。

その中でも購入頻度が高いと考えられる。

また、定価日と特価日の価格変動による販売量の変化が顕著である。[4]



予測に用いるデータを定価日と特価日に分けることによって予測精度が向上するかどうか検証する。

[4]高橋幸一[他], ニューラルネットワークモデルによる牛乳販売量予測の検討”, 2000.

# 関連研究 (1/2)

---

牛乳の企業別需要分析-特売情報を含むPOSデータを利用して-[5]

牛乳についてメーカー別に需要分析を行い、  
企業の特売効果を明らかにしていた。



販売金額シェアの7割以上が特売で占められていること  
牛乳の大手メーカーでも他との差別戦略が十分でないこと

[5]外園智史[他], ”牛乳の企業別需要分析-特売情報を含むPOSデータを利用して-”, 2009.

# 関連研究 (2/2)

ニューラルネットによる牛乳販売量の予測 [6]

従属変数としてPOSデータの変数が少ない場合…

新しい変数を作成することが重要な技術であると判断している。

下記のように変数を作成していた。



課題…店舗による商品の違いや期間限定商品の検討

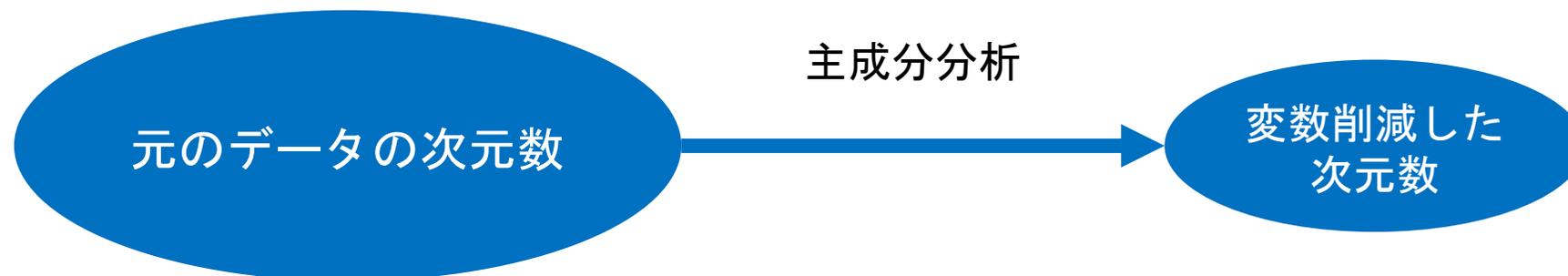
[6] 鈴木督久, ”ニューラルネットによる牛乳販売量の予測”, 2001.

# 提案手法 (1/4)

主成分分析 (Principle Component Analysis)

相関のある一連の変数の変動を一連の新しい相関のない変数によって記述すること。この新しい変数が主成分で、主成分は重要度の降順で導出される。

特徴として…



機械学習における過学習を防ぐ

[7] 大門貴志, “Rによる統計解析ハンドブック”, 2010.

# 提案手法 (2/4)

## ニューラルネットワーク (Neural Network)

神経細胞網を模したモデルであり、様々な分野で応用されている。

データに応じ非線型の関係を比較的簡単に導き出すことができる。[8]

ニューラルネットワークの研究例

顔画像内特徴点の自動抽出[9]

水文観測データに基づいた洪水予測[10]

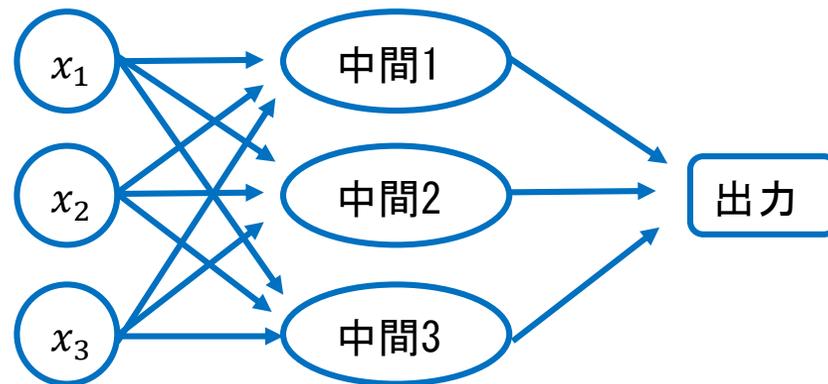


図1. ニューラルネットワークモデル図

[8] 柳貴久男, 垂水共之, “ニューラルネットワークを用いたデータ圧縮”, 1999.

[9] 阿部満[他], “ニューラルネットワークによる顔画像内特徴点の自動抽出”, 2013.

[10] 一言正之[他], “ニューラルネットワークによる洪水予測の精度向上”, 2015.

# 提案手法 (3/4)

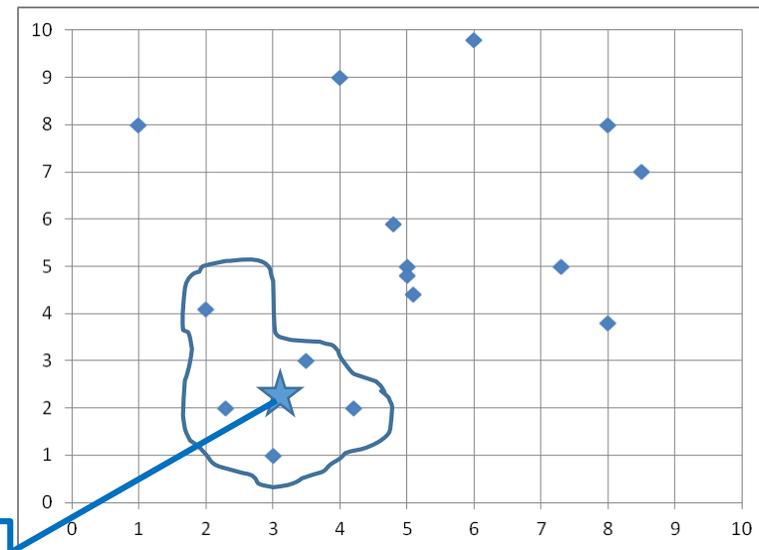
## k近傍法 (k-nearest neighbor algorithm)

数値データとそのときのクラスを保存し、新たな数値データを得るごとにデータから最も近いk個のデータを選び多数決で予測値を選ぶアルゴリズムである。[11]

### k近傍法の研究例

ゲームにおける敵の位置と行動の予測[11]

語義曖昧性解消[12]



予測したい値

図2. k-nn モデル図

[11] 浅山和宜[他], “線形補外とk近傍法を用いた格闘ゲームにおける敵の位置と行動の予測”, 2015.

[12] 新納造幸[他], “k近傍法とトピックモデルを利用した語義曖昧性解消の領域適応”, 2013.

# 提案手法 (4/4)

## 価格弾力性

弾力性とは他の要因を一定としたときに、Xの変化率をYの変化率で割った値と定義される。よって、**価格弾力性**は下記の式で表される。

$$\text{価格弾力性} = \frac{\text{販売量の変化率}[\%]}{\text{価格の変化率}[\%]}$$

$$\frac{\text{変化後の販売量} - \text{変化前の販売量}}{\text{変化前の販売量}}$$

$$\frac{\text{変化後の価格} - \text{変化前の価格}}{\text{変化前の価格}}$$

[13] 碓井 健寛, “ごみ処理サービス需要の価格弾力性—要因分析と予測”, 2003.

[14] 滝沢昌[他], “たばこ需要の特殊性を考慮した税収への影響分析“, 2014.

# 使用データ (1/7) -スーパーマーケットの情報-

使用データとして表1のスーパーマーケット9店舗の  
1年間(2007年4月～2008年3月)のPOSデータをまとめている。

表1. スーパーマーケットの詳細表

データ概要	石川県の同一チェーンスーパーマーケット9店舗の販売データ (以下、「店舗1」～「店舗9」と表記)の販売データ
期間	2007年4月～2008年3月までの1年間 (ただし、店舗2の4月のデータが欠損)
データ項目	日付、顧客ID、部門ID、商品名、JANコード、単価の6項目
レコード数	全店舗の平均レコード数は、約45万レコード/月 (販売数の多い店舗では、約75万レコード/月、 販売数の少ない店舗では、約25万レコード/月)

# 使用データ (2/7) -使用する牛乳の情報-

全体のPOSデータから互いに競合関係にある

「元気になあれ牛乳(以下、元気牛乳と表記)」

「北海道十勝牛乳(以下、北海道牛乳と表記)」

2つのデータを抽出し使用する。

表2. 元気, 北海道牛乳の定価, 特価日の詳細表

	元気牛乳		北海道牛乳	
	定価(～189円)	特価(～168円)	定価(～189円)	特価(～168円)
販売量合計	5,169	53,420	17,927	35,528
月平均販売量	430.75	4451.67	1493.92	2960.67
標準偏差	182.89	411.92	248.93	522.62
合計日数	157	324	251	183
月平均販売日数	13.08	27	20.92	15.25
標準偏差	2.84	2.12	1.26	1.42

どちらも  
定価日の販売価格…～189円  
特価日の販売価格…～168円  
であった。

# 使用データ (3/7) -9店舗の情報-

全9店舗(以下、「店舗1」～「店舗9」)の中から使用する店舗として店舗3と店舗6を使用する。

右の散布図から

- 店舗3の販売個数は平均的
  - 店舗6の販売個数は多い傾向
- であることから使用するデータとして採用した。

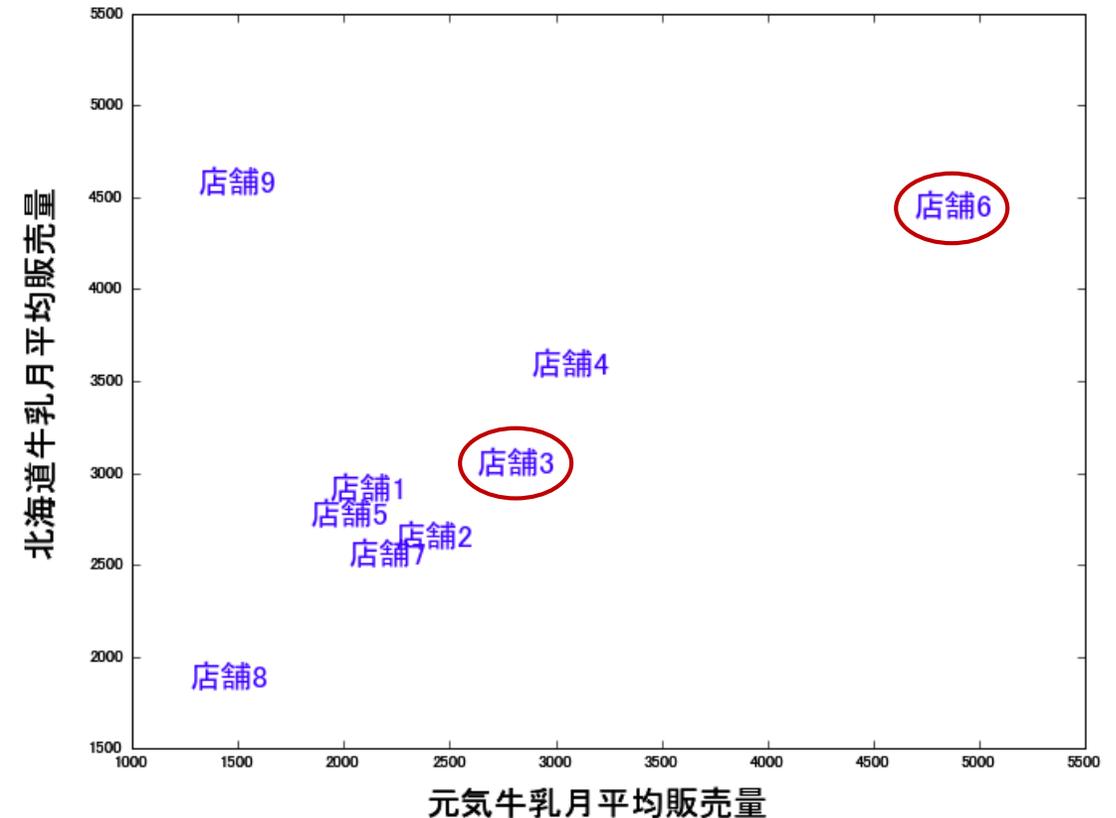


図3. 9店舗分の2種類の牛乳の散布図

表3. 店舗3の元気牛乳に関する情報

	元気牛乳	
	定価 (~189円)	特価 (~169円)
販売量合計	4,036	31,697
月平均販売量	336	2,641
標準偏差	24	94
合計日数	115	246
月平均販売日数	10	21
標準偏差	2.8	2.6

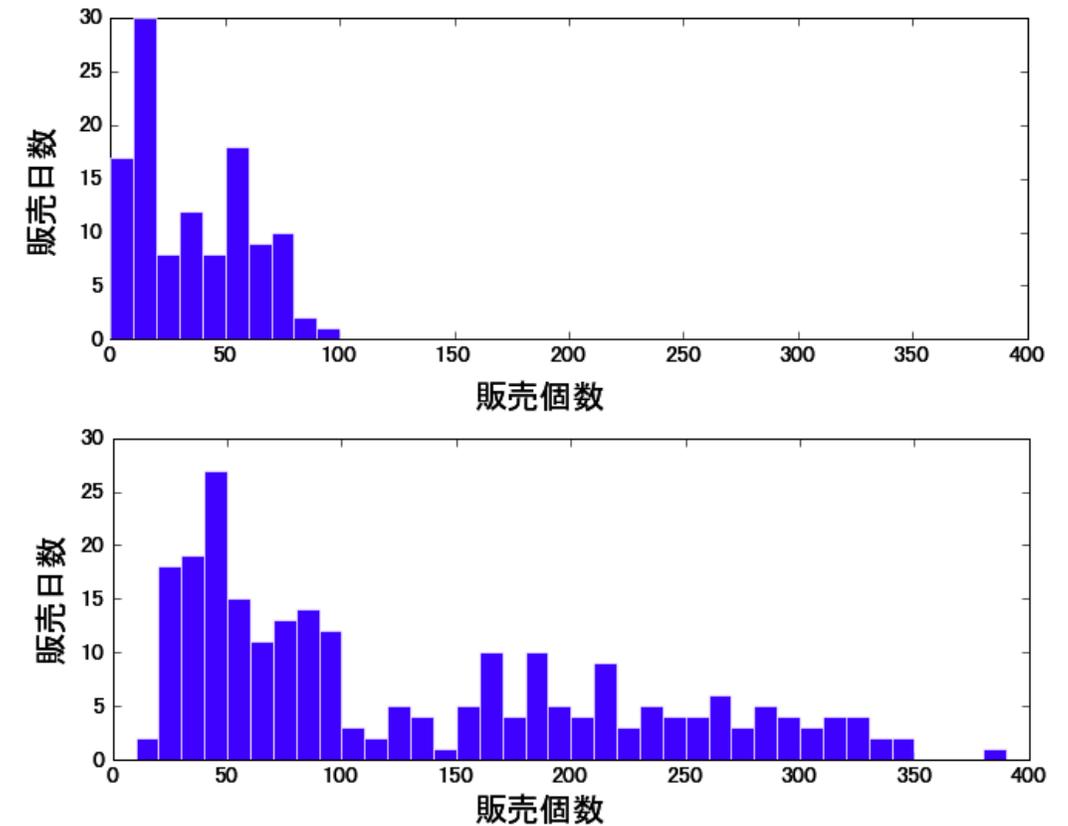


図4. 店舗3の定価と特価のヒストグラム

表4. 店舗3の北海道牛乳に関する情報

	北海道牛乳	
	定価 (~189円)	特価 (~169円)
販売量合計	10,311	26,309
月平均販売量	859	2,192
標準偏差	20	81
合計日数	234	124
月平均販売日数	20	10
標準偏差	1.44	1.37

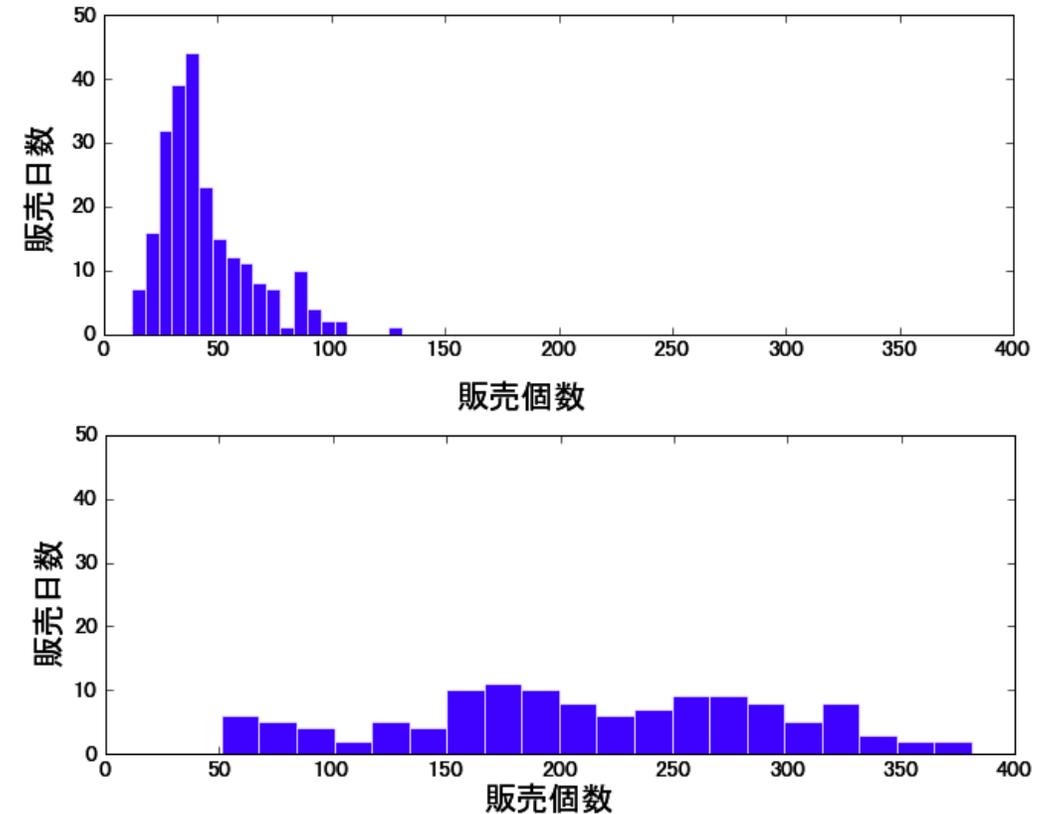


図5. 店舗3の定価と特価のヒストグラム

表5. 店舗6の元気牛乳に関する情報

	元気牛乳	
	定価 (~189円)	特価 (~169円)
販売量合計	5,637	55,155
月平均販売量	470	4,596
標準偏差	29	125
合計日数	88	264
月平均販売日数	7	22
標準偏差	2.92	2.86

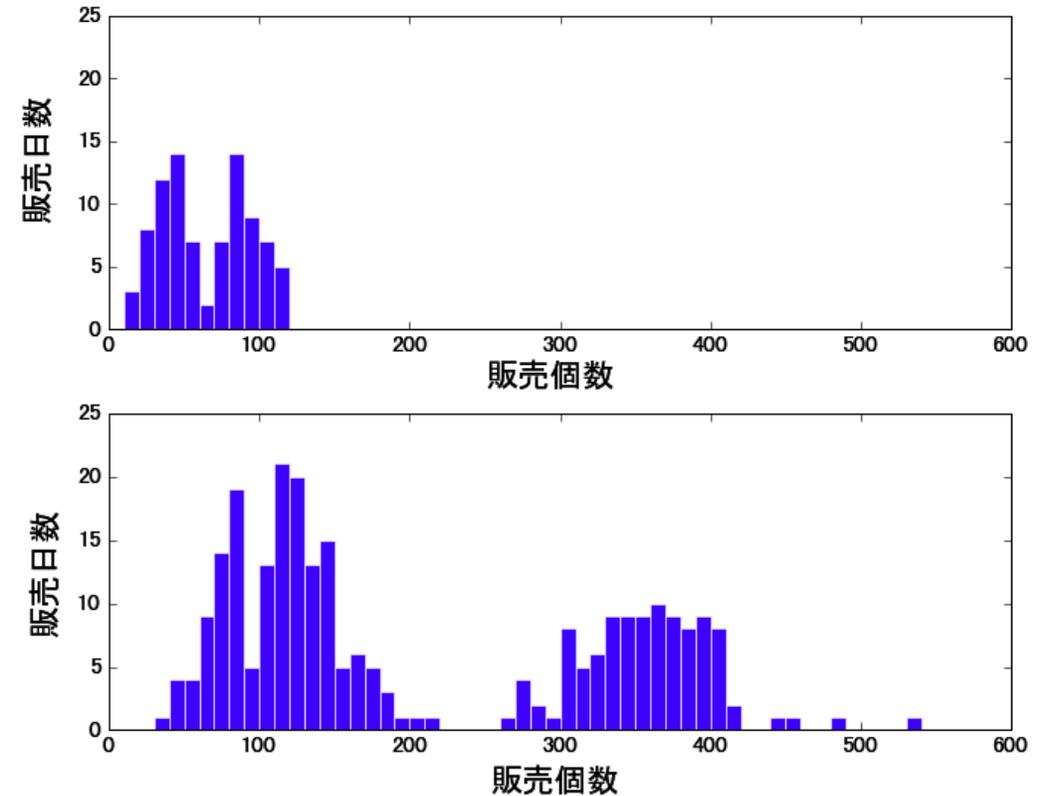


図6. 店舗6の定価と特価のヒストグラム

表6. 店舗6の北海道牛乳に関する情報

	北海道牛乳	
	定価 (~189円)	特価 (~169円)
販売量合計	18,607	32,655
月平均販売量	1,551	2,721
標準偏差	39	84
合計日数	234	118
月平均販売日数	20	10
標準偏差	1.44	1.46

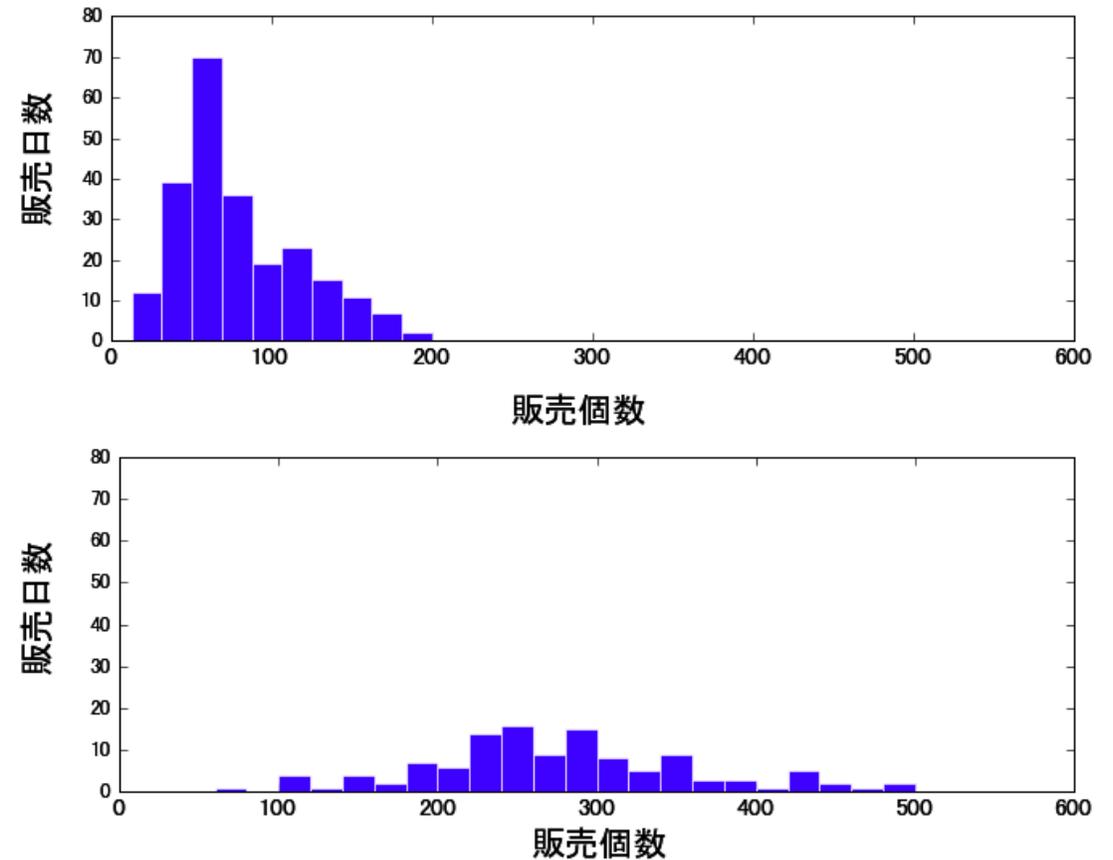


図7. 店舗6の定価と特価のヒストグラム

# 実験概要 (1/8)

---

本実験では、4つの実験を行う。実験内容を以下に示す。

- 実験1…店舗3の元気牛乳の需要予測
- 実験2…店舗3の北海道牛乳の需要予測
- 実験3…店舗6の元気牛乳の需要予測
- 実験4…店舗6の北海道牛乳の需要予測

また、本実験では販売単価, 販売量が「0」の値を欠損値として除外する。

1日の販売個数の需要予測を行い、累積誤差を比較し最適なモデルを検討する。

# 実験概要 (2/8)

手法は以下の2種類を用いる。

次元削減しない  
+  
k近傍法

主成分分析で累積寄与率9割以上まで次元削減  
+  
ニューラルネットワーク

また、価格弾力性を考慮するかどうかも含め、  
定価, 特価別に分けるかどうかも考慮した。

したがって、1つの実験で4つの結果が得られることになる。

# 実験概要 (3/8)

## 各モデルの設定

表7. ニューラルネットワークの設定

ニューラルネットワークの設定	
目的関数	Sigmoid-Sum of Square
隠れ層のユニット数	主成分分析で90%以上の説明変数
繰り返し最大数	1000

表8. k-nn法の設定

k-nn法の設定	
目的関数	Euclid
k値	5

表9. 使用データの情報

No.	データ数	牛乳	定価日	特価日 (例外特価日)
店舗3	358個	元気牛乳	115日	243日 (19日)
		北海道牛乳	234日	124日 (23日)
店舗6	352個	元気牛乳	88日	264日 (32日)
		北海道牛乳	234日	118日 (19日)

# 実験概要 (4/8)

実験1, 3で使用する説明変数を以下に示す。

表10. 実験1, 3で価格弾力性を考慮しない場合

説明変数	備考	数
曜日	日曜～土曜までを2値で表現	7
来客数	前日,前々日	2
平均気温	当日,前日,前々日	3
最高気温	当日,前日,前々日	3
最低気温	当日,前日,前々日	3
降水量	当日,前日,前々日	3
元気牛乳の販売個数	前日,前々日(当日は目的変数)	2
元気牛乳の販売単価	当日,前日,前々日	3
北海道牛乳の販売個数	前日,前々日	2
北海道牛乳の販売単価	当日,前日,前々日	3
説明変数の合計		31

表11. 実験1, 3で価格弾力性を考慮する場合

説明変数	備考	数
曜日	日曜～土曜までを2値で表現	7
来客数	前日,前々日	2
平均気温	当日,前日,前々日	3
最高気温	当日,前日,前々日	3
最低気温	当日,前日,前々日	3
降水量	当日,前日,前々日	3
元気牛乳の販売個数	前日,前々日(当日は目的変数)	2
元気牛乳の販売単価	当日,前日,前々日	3
北海道牛乳の販売個数	前日,前々日	2
北海道牛乳の販売単価	当日,前日,前々日	3
牛乳の価格弾力性	元気牛乳,北海道牛乳 それぞれ前日,前々日	4
説明変数の合計		35

# 実験概要 (5/8)

実験2, 4で使用する説明変数を以下に示す。

表12. 実験2, 4で価格弾力性を考慮しない場合

説明変数	備考	数
曜日	日曜～土曜までを2値で表現	7
来客数	前日,前々日	2
平均気温	当日,前日,前々日	3
最高気温	当日,前日,前々日	3
最低気温	当日,前日,前々日	3
降水量	当日,前日,前々日	3
元気牛乳の販売個数	前日,前々日	2
元気牛乳の販売単価	当日,前日,前々日	3
北海道牛乳の販売個数	前日,前々日(当日は目的変数)	2
北海道牛乳の販売単価	当日,前日,前々日	3
説明変数の合計		31

表13. 実験2, 4で価格弾力性を考慮する場合

説明変数	備考	数
曜日	日曜～土曜までを2値で表現	7
来客数	前日,前々日	2
平均気温	当日,前日,前々日	3
最高気温	当日,前日,前々日	3
最低気温	当日,前日,前々日	3
降水量	当日,前日,前々日	3
元気牛乳の販売個数	前日,前々日	2
元気牛乳の販売単価	当日,前日,前々日	3
北海道牛乳の販売個数	前日,前々日(当日は目的変数)	2
北海道牛乳の販売単価	当日,前日,前々日	3
牛乳の価格弾力性	元気牛乳,北海道牛乳 それぞれ前日,前々日	4
説明変数の合計		35

# 実験概要 (6/8)

[VMS \(Visual Mining Studio\)](#) を使用する。

[VMS](#) は株式会社NTTデータ数理システムが提供するデータマイニングツールである。データマイニングに必要なデータの前処理から分析, 処理など高機能なツール群を誰でも簡単に利用することができる特徴がある。

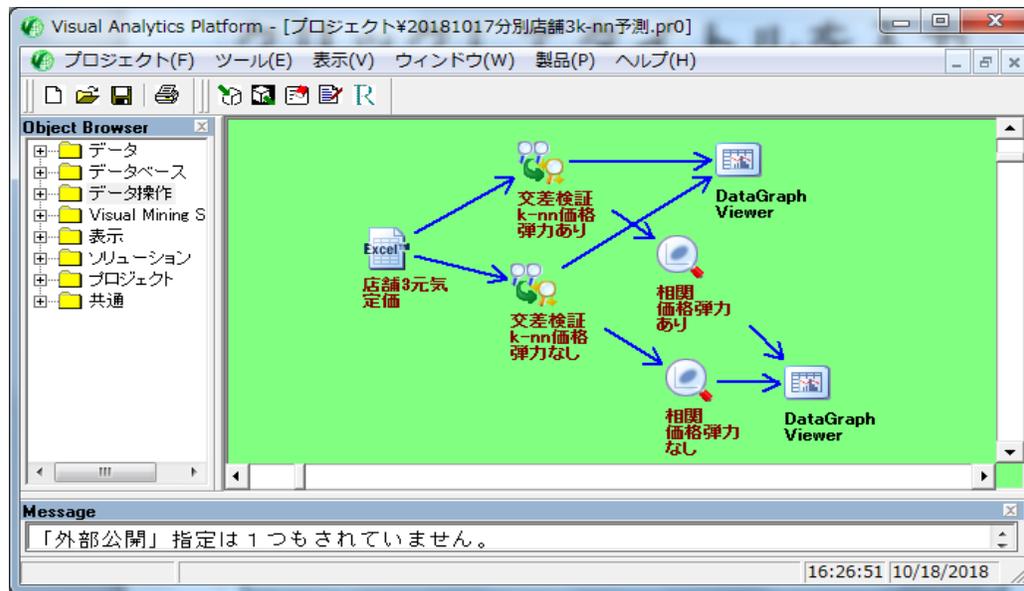


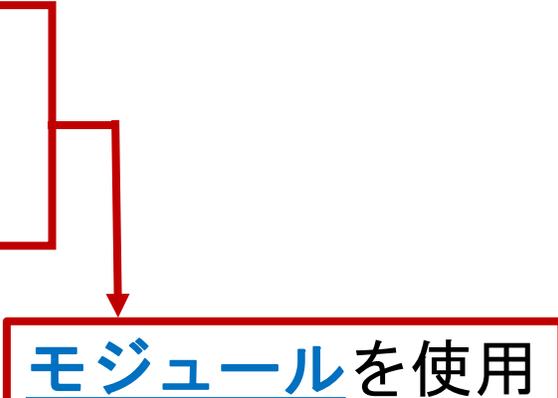
図8. VMSのプロジェクト図

# 実験概要 (7/8)

VMSで使用した機能を以下にまとめる。

表14. VMSで扱った機能

列属性変更	データの列の属性を変更する。
マージ	複数のデータを結合する。
主成分分析	データを主成分で表現する。
相関	データ列の相関を求める。
交差検証法	モデルを構築する際の推定誤差を検証する。
データ & グラフビュー	データをテーブル形式に表示, グラフを作成できる。



モジュールを使用

## モジュール

- 処理フローの一部をまとめて部品として扱う機能
- 処理フローをまとめて頻繁に使用する処理を繰り返して実行することができる

# 実験概要 (8/8)

VMSのプロジェクト図を下記に示す。

主成分分析を行うにあたりモジュールを使用し作業の簡略化を行った。

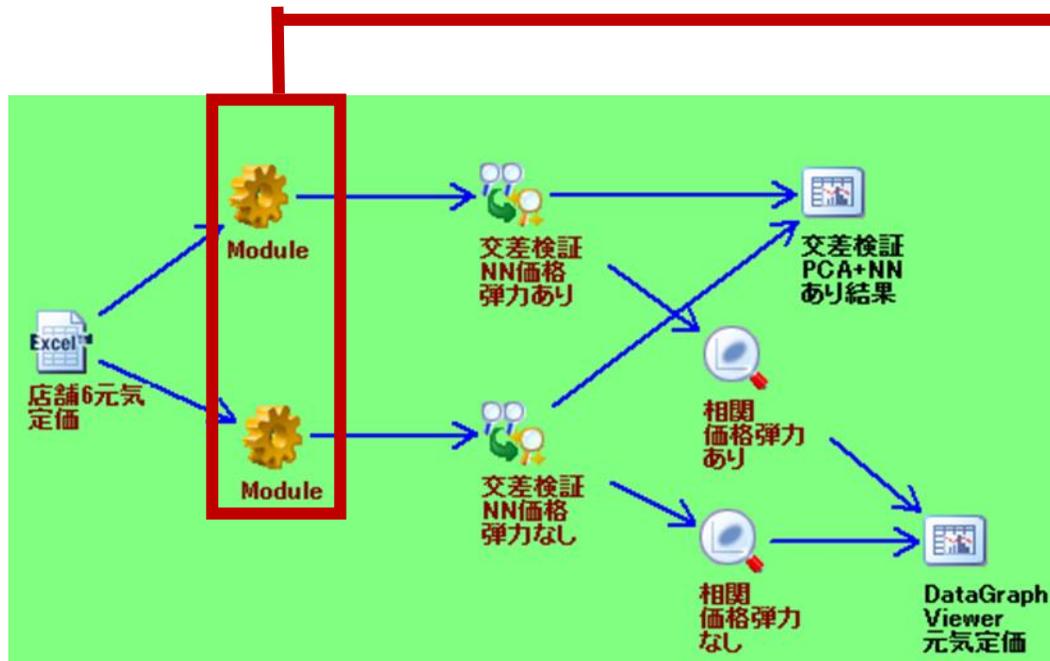


図9. プロジェクト図の一部

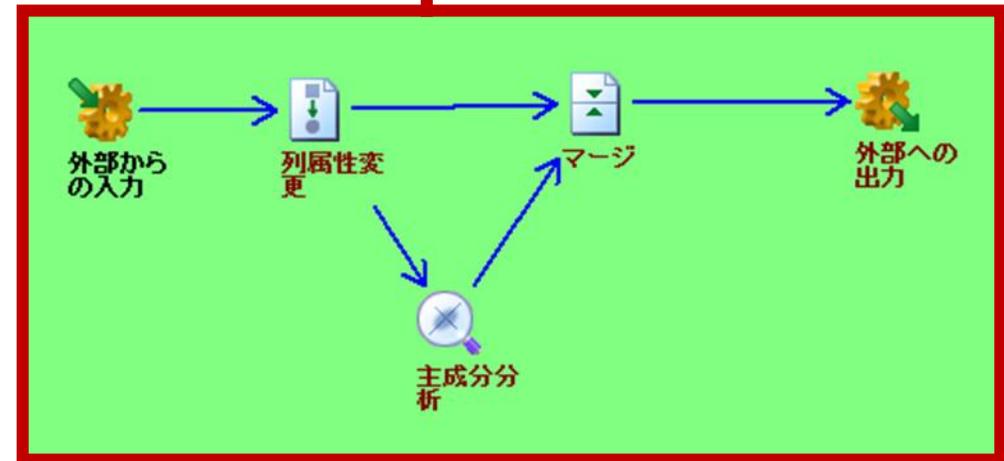


図10. モジュール内部図

# 実験1結果 (店舗3+元気牛乳)

## k-nn法

表15. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,173	4,196	5,368	
	誤差	平均値	10	17	15
		最大値	51	163	163
		標準偏差	11	18	16
あり	累積誤差	1,229	4,230	5,459	
	誤差	平均値	11	17	15
		最大値	47	163	163
		標準偏差	11	19	17

表16. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,085	4,026	5,111	
	誤差	平均値	9	17	14
		最大値	47	163	163
		標準偏差	10	18	17
あり	累積誤差	1,113	4,062	5,175	
	誤差	平均値	10	17	14
		最大値	47	163	163
		標準偏差	10	18	17

### 結果

上の表の累積誤差に着目すると全価格によるモデルより価格別によるモデルの方が精度が良くなる結果が得られた。

# 実験1結果 (店舗3+元気牛乳)

## 主成分分析+NN

表17. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,519	4,340	5,859	
	誤差	平均値	13	18	16
		最大値	61	129	129
		標準偏差	11	16	15
あり	累積誤差	1,258	4,612	5,870	
	誤差	平均値	11	19	16
		最大値	51	145	145
		標準偏差	11	18	17

表18. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,563	4,800	6,364	
	誤差	平均値	14	20	18
		最大値	82	174	174
		標準偏差	16	21	19
あり	累積誤差	1,357	4,502	5,860	
	誤差	平均値	12	19	16
		最大値	54	148	148
		標準偏差	11	19	17

### 結果

表18のマーカ一部分から全価格モデルより価格別によるモデルの累積誤差, 誤差の最大値ともに大きくなっている。

# 実験2結果 (店舗3+北海道牛乳)

## k-nn法

表19. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	2,176	3,361	5,538	
	誤差	平均値	9	27	15
		最大値	53	124	124
		標準偏差	8	24	18
あり	累積誤差	2,144	3,439	5,583	
	誤差	平均値	9	28	16
		最大値	53	124	124
		標準偏差	8	25	19

表20. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	2,145	2,905	5,050	
	誤差	平均値	9	23	14
		最大値	53	105	105
		標準偏差	8	22	16
あり	累積誤差	2,140	2,894	5,034	
	誤差	平均値	9	23	14
		最大値	53	105	105
		標準偏差	8	22	16

### 結果

上の表のマーカ一部分に着目すると全価格によるモデルより価格別によるモデルの方が精度が良くなる結果が得られた。

# 実験2結果 (店舗3+北海道牛乳)

## 主成分分析+NN

表21. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	2,547	3,838	6,385	
	誤差	平均値	11	31	18
		最大値	47	159	159
		標準偏差	9	27	20
あり	累積誤差	2,343	3,653	5,996	
	誤差	平均値	10	29	17
		最大値	48	123	123
		標準偏差	9	24	18

表22. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	2,371	4,075	6,446	
	誤差	平均値	10	33	18
		最大値	52	202	202
		標準偏差	10	32	23
あり	累積誤差	2,492	3,479	5,971	
	誤差	平均値	11	28	17
		最大値	64	170	170
		標準偏差	11	28	21

### 結果

表22のマーカ一部分から価格別モデルは精度が悪くなっている。  
しかし、定価の部分に着目すると累積誤差の値は小さくなっている。

# 実験3結果 (店舗6+元気牛乳)

## k-nn法

表23. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,426	7,158	8,584	
	誤差	平均値	16	27	24
		最大値	65	206	206
		標準偏差	14	24	22
あり	累積誤差	1,435	7,152	8,587	
	誤差	平均値	16	27	24
		最大値	65	206	206
		標準偏差	14	24	22

表24. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,064	7,002	8,066	
	誤差	平均値	12	27	23
		最大値	48	181	181
		標準偏差	10	23	22
あり	累積誤差	1,115	7,106	8,221	
	誤差	平均値	13	27	23
		最大値	48	189	189
		標準偏差	10	24	22

### 結果

上の表の累積誤差に着目すると全価格によるモデルより価格別によるモデルの方が精度が良くなる結果が得られた。

# 実験3結果 (店舗6+元気牛乳)

## 主成分分析+NN

表25. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,576	7,156	8,732	
	誤差	平均値	18	27	25
		最大値	113	170	170
		標準偏差	18	24	23
あり	累積誤差	1,821	7,354	9,175	
	誤差	平均値	21	28	26
		最大値	297	173	297
		標準偏差	34	24	27

表26. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,315	6,931	8,246	
	誤差	平均値	15	26	23
		最大値	61	163	163
		標準偏差	12	25	23
あり	累積誤差	1,343	7,152	8,495	
	誤差	平均値	15	27	24
		最大値	45	169	169
		標準偏差	12	25	23

### 結果

k-nn法での結果より累積誤差が増しているが価格別によるモデルの精度が良いという結果が得られた。

# 実験4結果 (店舗6+北海道牛乳)

## k-nn法

表27. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	3,785	4,523	8,308	
	誤差	平均値	16	38	24
		最大値	82	137	137
		標準偏差	14	32	24
あり	累積誤差	3,816	4,507	8,324	
	誤差	平均値	16	38	24
		最大値	82	137	137
		標準偏差	14	32	24

表28. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	3,693	4,813	8,506	
	誤差	平均値	16	41	24
		最大値	83	275	275
		標準偏差	14	43	30
あり	累積誤差	3,666	4,780	8,449	
	誤差	平均値	16	41	24
		最大値	83	275	275
		標準偏差	14	44	30

### 結果

表のメーカー一部分から全価格によるモデルの方が精度が良いことが分かる。  
 表28の最大値の日を調べたところ例外特価日であり実際の値より大きく予測していた。  
 また、全価格モデルのその日は小さい値で予測していた。

# 実験4結果 (店舗6+北海道牛乳)

## 主成分分析+NN

表29. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	3,907	4,203	8,110	
	誤差	平均値	17	36	23
		最大値	95	135	135
		標準偏差	15	27	22
あり	累積誤差	4,437	4,521	8,957	
	誤差	平均値	19	38	25
		最大値	90	155	155
		標準偏差	15	30	23

表30. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	3,951	5,389	9,340	
	誤差	平均値	17	46	27
		最大値	98	366	366
		標準偏差	16	52	35
あり	累積誤差	3,828	4,995	8,823	
	誤差	平均値	16	42	25
		最大値	89	330	330
		標準偏差	16	45	31

### 結果

上の表の累積誤差に着目すると全価格によるモデルより価格別によるモデルの方が精度が悪くなる結果が得られた。

# 考察(1/8)

実験1~4の累積誤差を見ると

価格別に  
分けること

説明変数を削減しない  
k-nn法

価格弾力性を  
考慮しないこと

以上の要素を含んだモデルの精度が良いことが分かった。

価格弾力性を説明変数に含めなくてもよい理由として  
31変数の中に牛乳の前日, 前々日の販売個数, 販売単価の情報が  
含まれていることにより過学習を起こしたと考えられる。

k-nn法がNNより精度が良いのは新しい関数で予測値を求めるより  
訓練データから予測値を多数決で求める手法がよかったからだと考えている。

# 考察 (2/8)

---

表18, 22のマーカ一部分が高い要因を述べる。

## 実験1

誤差の最大値 (174) の日を調べたところ、148円の日曜日であった。日曜日は毎週特価日であるが、148円の日はこの日だけであった。したがって、例外特価日にモデルが対応できなかったのではないかと考えられる。

## 実験2

誤差の最大値 (202) の日を調べたところ、前日が366個でよく販売された日であったため影響を及ぼしたのではないかと考えられる。実際、前日に300個以上売れた翌日の予測精度はよいものではなかった。

# 考察(3/8)

価格弾力性の有用性を確認ために変数を変更して実験を行った。  
変数は以下のように設定した。

表31. 価格弾力性を考慮しない場合

説明変数	備考	数
曜日	日曜～土曜までを2値で表現	7
元気牛乳の販売個数	前日	1
北海道牛乳の販売個数	前日	1
説明変数の合計		9

表32. 価格弾力性を考慮する場合

説明変数	備考	数
曜日	日曜～土曜までを2値で表現	7
元気牛乳の販売個数	前日	1
北海道牛乳の販売個数	前日	1
牛乳の価格弾力性	元気牛乳,北海道牛乳 それぞれ前日	2
説明変数の合計		11

# 考察 (店舗3+元気牛乳) (4/8)

k-nn法

表33. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,592	5,246	6,838	
	誤差	平均値	14	22	19
		最大値	50	157	157
		標準偏差	12	23	20
あり	累積誤差	1,609	5,455	7,064	
	誤差	平均値	14	22	20
		最大値	52	154	154
		標準偏差	12	22	20

表34. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,316	4,026	5,342	
	誤差	平均値	11	17	15
		最大値	49	163	163
		標準偏差	12	18	17
あり	累積誤差	1,219	4,062	5,281	
	誤差	平均値	11	17	15
		最大値	50	163	163
		標準偏差	11	18	17

結果

表34の価格弾力性ありの累積誤差が最も小さいことから価格弾力性が変数として機能したことがわかる。

# 考察 (店舗3+北海道牛乳) (5/8)

k-nn法

表35. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	2,469	5,009	7,478	
	誤差	平均値	11	40	21
		最大値	62	219	219
		標準偏差	10	38	28
あり	累積誤差	2,554	4,996	7,550	
	誤差	平均値	11	40	21
		最大値	62	189	189
		標準偏差	9	35	26

表36. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	2,145	2,905	5,050	
	誤差	平均値	9	23	14
		最大値	53	105	105
		標準偏差	8	22	16
あり	累積誤差	2,140	2,894	5,034	
	誤差	平均値	9	23	14
		最大値	53	105	105
		標準偏差	8	22	16

結果

表36の価格弾力性ありの累積誤差が最も小さいことから価格弾力性が変数として機能したことがわかる。

# 考察 (店舗6+元気牛乳) (6/8)

k-nn法

表37. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	2,116	7,526	9,641	
	誤差	平均値	24	29	27
		最大値	88	136	136
		標準偏差	18	24	23
あり	累積誤差	1,986	7,422	9,408	
	誤差	平均値	23	28	27
		最大値	88	161	161
		標準偏差	19	24	23

表38. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	1,093	7,483	8,576	
	誤差	平均値	12	28	24
		最大値	51	136	136
		標準偏差	10	24	23
あり	累積誤差	1,081	7,406	8,487	
	誤差	平均値	12	28	24
		最大値	51	161	161
		標準偏差	10	25	23

結果

表38の価格弾力性ありの累積誤差が最も小さいことから価格弾力性が変数として機能したことがわかる。

# 考察 (店舗6+北海道牛乳) (7/8)

## k-nn法

表39. 全価格によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	4,531	6,753	11,284	
	誤差	平均値	19	57	32
		最大値	89	222	222
		標準偏差	18	43	34
あり	累積誤差	4,464	6,821	11,285	
	誤差	平均値	19	58	32
		最大値	109	222	222
		標準偏差	19	43	34

表40. 価格別によるモデル結果

価格弾力性		定価	特価	合計	
なし	累積誤差	3,989	6,520	10,509	
	誤差	平均値	17	55	30
		最大値	89	222	222
		標準偏差	15	45	34
あり	累積誤差	3,815	6,680	10,495	
	誤差	平均値	16	57	30
		最大値	89	222	222
		標準偏差	15	46	35

### 結果

表40の価格弾力性ありの累積誤差が最も小さいことから価格弾力性が変数として機能したことがわかる。

## 考察(8/8)

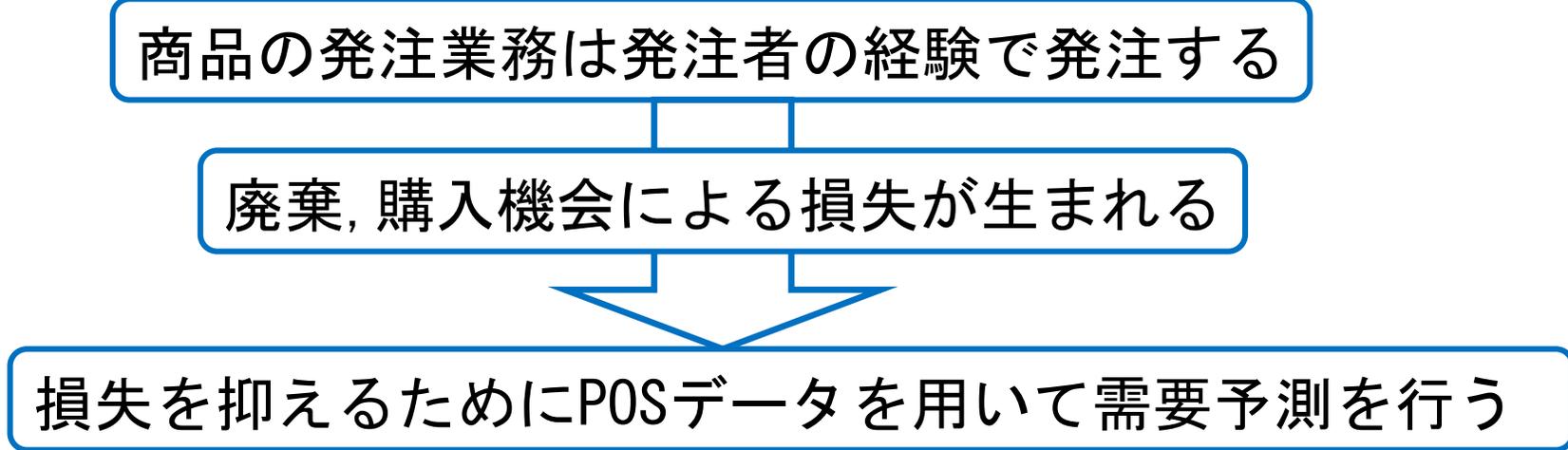
---

実験結果から変数が少ない場合、価格弾力性を考慮すると精度が良くなることが分かった。

したがって、弾力性を考慮し価格別に分けたモデルの精度がよい

しかし、累積誤差が実験1~4より大きくなったことから来客数や気温などの変数も重要であると考えられる。

# まとめ



需要予測としてk-nn法, ニューラルネットワークを用いて精度の比較を行った。

結果として、価格弾力性を考慮せず価格別に分けたk-nn法が

最良のモデルといえる。

また、価格弾力性を変数に取り入れるると例外特価日に有効であることも分かった。

# 今後の課題

---

- 牛乳の販売金額シェアの7割以上が特売で占められている。
- 定価日と特価日と分けてモデルを作るだけでなく例外特価日も考慮したモデル作成を考えたい。
- 今回はデータマイニング手法が2種類だけであったので別の手法を取り入れた実験を行う。

# 参考文献

- [1] 生田目崇, “マーケティング分野におけるデータ分析の進展” 横幹連合コンファレンス予稿集, 2007.
- [2] 阿部誠, 近藤文代, “マーケティングの科学-POSデータの解析-“, 2005.
- [3] 松村直樹, 和泉潔, 山田 健太, ” POSデータに基づく欠品時の顧客行動を考慮した小売店舗の購買シミュレーション”, 人工知能学会論文誌, Vol. 31, No. 2, F-F13, pp. 1-8, 2016.
- [4] 高橋幸一, 石川征郎, ” ニューラルネットワークモデルによる牛乳販売量予測の検討”, 日本ユーザー会, pp. 235-242, 2000.
- [5] 外園, 智嗣, 狩野秀之, 前田幸嗣, ” 牛乳の企業別需要分析-特売情報を含むPOSデータを利用して-”, フードシステム研究, 第16巻3号, pp. 15-23, 2009.
- [6] 鈴木督久, “ニューラルネットによる牛乳販売量の予測”, 品質管理, Vol. 52, No. 3, pp. 249-254, 2001.
- [7] Brian S. Everitt, Torsten Hothorn・著者, 大門貴志, 吉川俊博, 手良向聡・訳者, RIによる統計解析ハンドブック, 株式会社メディカル・パブリケーションズ, 2010.
- [8] 柳貴久男, 垂水共之, “ニューラルネットワークを用いたデータ圧縮”, 日本計算機統計学会大会論文集, 13巻, pp. 84-85, 1999.
- [9] 阿部満, 皆川祐太郎, 趙強福, “ニューラルネットワークによる顔画像内特徴点の自動抽出”, 電気関係学会東北支部連合大会, 2013.
- [10] 一言正之, 清雄一, 桜庭雅明, “ニューラルネットワークによる洪水予測の精度向上“, 人工知能学会全国大会論文集, pp. 29, 2015.
- [11] 浅山和宜, 森山甲一, 福井健一, 沼尾正行 “線形補外とk近傍法を用いた格闘ゲームにおける敵の位置と行動の予測”, pp. 29 人工知能学会全国大会論文集, 2015.
- [12] 新納浩幸, 佐々木稔 “k近傍法とトピックモデルを利用した語義曖昧性解消の領域適応”, 自然言語処理, Vol. 20, No. 5, pp. 707-726, 2013.
- [13] 碓井健寛, “ごみ処理サービス需要の価格弾力性—要因分析と予測”, 環境科学会誌, 16巻4号, pp. 271-280, 2003.
- [14] 滝沢昌, 嶋崎善章, “たばこ需要の特殊性を考慮した税込への影響分析“, 行動経済学, 7巻, pp. 96-99, 2014.

ご清聴ありがとうございました

---