すだれの需要予測
– パターン分析による最適発注量の決定 –

2001MM052 長坂 俊希　　2002MM007 深見 尚登
指導教員　鈴木 敦夫

1　はじめに

1.1　背景

ホームセンターは在庫切れを起こさないために多量に在庫を抱えている．ホームセンターの総費用の 10 %から 30 %は在庫費となっている．

在庫問題の一つに「季節品」がある．季節品とは一定の期間内でしか売れない商品の事である．季節品の例として, すだれ, エアコン, 扇風機, ホットカーペット, こたつ等が挙げられる．特に「すだれ」については大量の在庫が問題となっている．すだれは 4 月から 9 月のみ販売している商品である．その発注方法は, 各店舗が必要に応じて, 販売のためにすだれを発注する．その際には, ピークの時期に欠品することを恐れてあらかじめ大量に発注している．売れ残ったすだれは廃棄処分となる．

まとめると, 各店舗が発注する数量を決める事は在庫問題にも関わるため, 重要な問題なのである．

1.2　アプローチ

本研究では, すだれの発注の基礎的データとなる予測手法を提案する．参考文献 [2] ではすだれの販売が開始する 1 週目から全ての売上数データを用い, 1 週先の予測を行う．しかし, その方法ではいくつかの改善点が必要に感じた．理由としては昨年度, 研究室で考案された需要予測のモデル (以後, 昨年度のモデル) において, 時間の経過を伴う需要の変化に対する配慮がなされていなかった点である．4 月から 8 月は気温の上昇が著しい．避暑の役割を果たす, すだれの需要は伸びていく．つまり, 7 月, 8 月のデータを予測する際に, 4 月からの売上数データを用いるより直近の売上数データだけを用いる方が良いのではないかと考えた．そこで, 予測する週の直近の 5 週間の売上数データを用い, 実際に売上数を予測する．

また冷夏となった 2003 年を 2002 年と同じように扱う点が気になった．2004 年の売上数を予測するのに冷夏の 2003 年を用い予測するのは適していないと考えた．

本研究ではすだれの売上数を逐次的に予測する方法を導き出す．以下の 4 つのモデルを考案した．

A　昨年度のモデル
B　直近の過去 5 週の売上数を用いたモデル
C　典型的な気候変動のみを用いたモデル
D　予測期間を延長したモデル

昨年度のモデルより精度の高い新しいモデルを考案することを目標とする．予測との誤差を出来る限り少なくするモデルを考え, どのような改善をする必要があるか考察する．

1.3　データの種類

本研究で扱うホームセンターのデータの種類は次の 3 通りである．

1.　各店舗のすだれの週別販売数
2.　各週におけるすだれの単価
3.　各週における最高気温

今回, データとして扱った店舗数は 2002 年では, 87 店舗, 2003 年は 99 店舗, 2004 年は 113 店舗となっている．**表 1** は 2002 年から 2004 年の第 1 週目から第 19 週目までの店舗のすだれの販売が開始する週の日から販売が終了する週の日を示している．

表 1　すだれの販売期間の週の割り当て

	2002 年	2003 年	2004 年
1 週目	4/1〜7	4/7〜13	4/5〜11
2 週目	4/8〜14	4/14〜20	4/12〜18
⋮	⋮	⋮	⋮
19 週目	8/5〜11	8/11〜17	8/9〜15

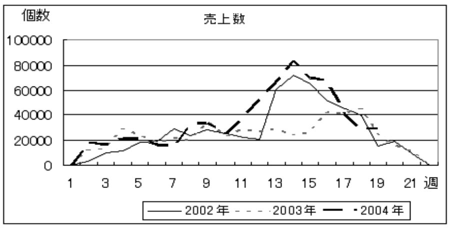

図 1　総売上数の推移図

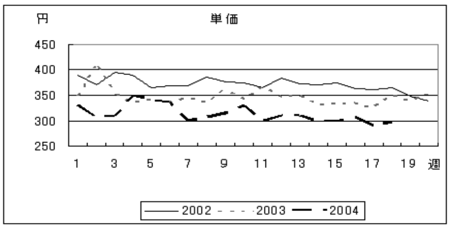

図 2　単価の推移図

図 1 は 2002 年から 2004 年までの全店舗のすだれの売上総数を示している．2002 年, 2004 年は平年通りの夏の天気となったのでピークの時期やピークの売上数がとてもはっきりとしていて似通ったデータである．しかし, 初期の売上数にかなり差がある．2003 年は冷夏となりピークの時期があいまいで売上数も伸びていないものとなっ

ている．

図 2 を見て分かるとおり 2002 年から 2003 年, 2004 年と年々すだれの単価が下がっている売上初期はこの単価が大きく左右されると考えられる．年々すだれが売れ始めた春の時点の売上数の増加がみられる．1 週目の各店舗の売上数をそれぞれ調べていくと 2004 年の売上数が 2002 年の売上数の 10 倍にも達するときがある．また例年では廃棄処分に向けて値下げの傾向が見られる．廃棄処分にはコストが掛かるので値下げをしても売れ残りを起こしたくないためである．

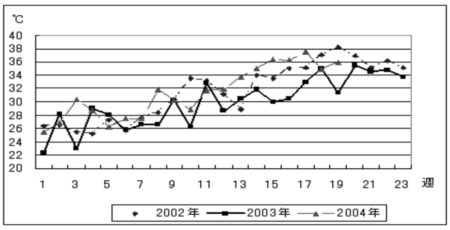

図 3　最高気温の推移図

表 2　35 ℃を超えた日数

2002 年	2003 年	2004 年
21 日	2 日	16 日

表 2 は愛知県名古屋市の 2002 年から 2004 年の最高気温が 35 ℃を超えた日数の表である．データは 2002 年の 4 月 1 日から 9 月 2 日, 2003 年の 4 月 7 日から 9 月 8 日, 2004 年の 4 月 5 日から 8 月 15 日である．すだれは気温に大きく左右される．最高気温が 35 ℃を超え始めるあたりから売上が大きく伸びる．2002 年, 2004 年と例年通りの気温で最高気温が 35 ℃を超えた日数が多くあったため大量にすだれが売れた．しかし 2003 年は 35 ℃を超える日が 2 日しかなく売上に大きな影響を与えた．ピーク時に売上数を伸ばせないと他に大量に売れる時期がないため売上総数がかなり少ないものとなってしまう．

2　すだれの需要予測方法

需要予測はすだれが売れ始めた春の時点の売上数で, 翌週の売上数を予測する逐次的方法である．この方法で売上傾向の特徴を捉える．昨年度のモデルと我々が考案したモデルを比較する．

何週間分かの実際の売上数を用い, それと過去の店舗との売上数を比べ, 差を出し 2 乗する．その差の 2 乗を誤差とする．誤差の最小の店舗を出力する．またロバストな予測値を得るために, 5 店舗の中央値をとる．

2.1　記号の定義

- S :　予測を始める週
- T :　季節の終了する週
- L :　予測する週 ($L = S, S+1, \cdots, T$)
- k :　累積したパターンサンプルの添え字
  (2002 年の売上数データ, モデル C で用いる)
  ($k = 1, 2, \cdots, 87$)
- i :　累積したパターンサンプルの添え字
  (2002,2003 年の売上数データ, モデル A,B,D で用いる)
  ($i = 1, 2, \cdots, 186$)
- j :　週の添え字 ($j = 1, 2, \cdots, 19$)
- x_{ij} :　店舗 i の j 週目の売上数
- y_j :　予測したい店舗の j 週目の実際の売上数
- I :　全体のなかでまだパターン予測の材料として適切でない集合
- K :　パターン予測に適切であるとみなすサンプルが要素である集合
- l :　反復数

2.2　A 昨年度のモデル

昨年度, 参考文献 [2] の松本らが考案したモデルについて説明する．すだれが売れ始めた 1 週目から全ての売上数データを用い, 1 週先の予測を行う．例を**図 4** で示す．**図 4** では店舗 (1) の 8 週目の売上数を予測したい．1 から 7 週目までの売上数の誤差の最も少ないパターンサンプルを見つけ, そのパターンサンプルの 8 週目の売上数を予測値とする．

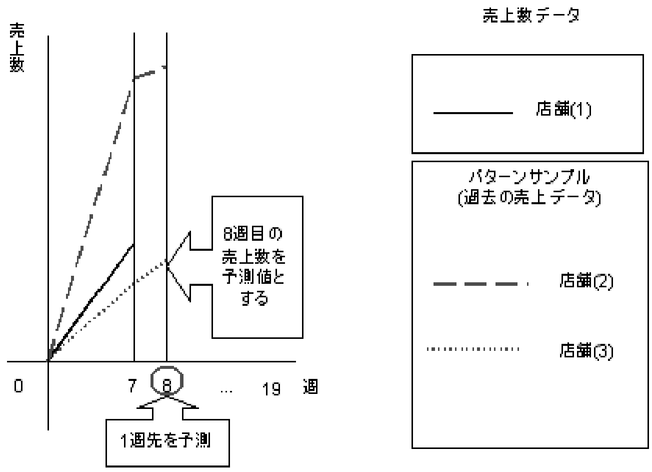

図 4　昨年度のモデル

2.2.1　モデル化

1. $L = S$ とする．
2. 誤差の最も少ないパターンサンプルの出力
$$i^*(L) = \underset{i \in I}{\arg\min}\left\{\sum_{j=1}^{L-1}(y_j - x_{ij})^2\right\}$$
3. 得られた $x_{i^*(L)}$ を $\hat{y}_{(L)}$ とみなし予測値とする．
4. もし $L = T$ ならば予測終了．
もしそうでないならば $L = L + 1$ として 2 に戻る．

2.3　B 直近の過去 5 週の売上数を用いたモデル

B のモデルは予測する週の直前の 5 週間分の売上数データを用い予測を行う．例を**図 5** で示す．**図 5** では店舗 (1) の 8 週目の売上数を予測したい. 3 から 7 週目までの売上数の誤差の最も少ないパターンサンプルを見つけ, そのパターンサンプルの 8 週目の売上数を予測値とする.

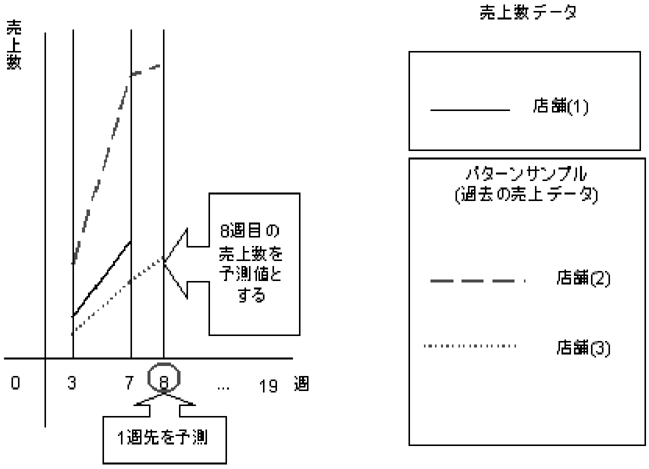

図 5　直近の過去 5 週の売上数を用いたモデル

2.3.1　モデル化

1. $L = S$ とする.
2. 誤差の最も少ないパターンサンプルの出力
$$i^*(L) = \underset{i\,\in\,I}{\arg}\ \min\Big\{\sum_{j=L-5}^{L-1}(y_j - x_{ij})^2\Big\}$$
3. 得られた $x_{i^*(L)}$ を $\hat{y}_{(L)}$ とみなし予測値とする.
4. もし $L = T$ ならば予測終了.
もしそうでないならば $L = L + 1$ として 2 に戻る.

2.4　C 典型的な気候変動のみを用いた予測

C のモデルは前節と同じ予測方法である. しかし, 2002 年の売上データのみをパターンサンプルとして用いている．2003 年は冷夏であったため, 平年並の気温である 2004 年を予測するデータとして扱うのには適していないと考えられる. 以上より, 2002 年の売上数データのみを用いて予測を行った方が良い結果を得られるという仮説を立てた. 長期天気予報の情報を用いれば, 予測したい年が冷夏であるか否かはある程度推測出来る．よって, この方法を応用すれば実際の問題にも対応出来る.

2.4.1　モデル化

1. $L = S$ とする.
2. 誤差の最も少ないパターンサンプルの出力
$$k^*(L) = \underset{k\,\in\,I}{\arg}\ \min\Big\{\sum_{j=L-5}^{L-1}(y_j - x_{kj})^2\Big\}$$
3. 得られた $x_{k^*(L)}$ を $\hat{y}_{(L)}$ とみなし予測値とする.
4. もし $L = T$ ならば予測終了.
もしそうでないならば $L = L + 1$ として 2 に戻る.

2.5　D 予測期間を延長したモデル

今までは 1 週先を予測していた. しかし, それでは発注が間に合わないことが判明した. 発注を間に合わせるために, B のモデルを用い 2 週先の需要を予測する. 例を**図 6** で示す．**図 6** では店舗 (1) の 9 週目の売上数を予測したい. 3 から 7 週目までの売上数の誤差の最も少ないパターンサンプルを見つけ, そのパターンサンプルの 9 週目の売上数を予測値とする.

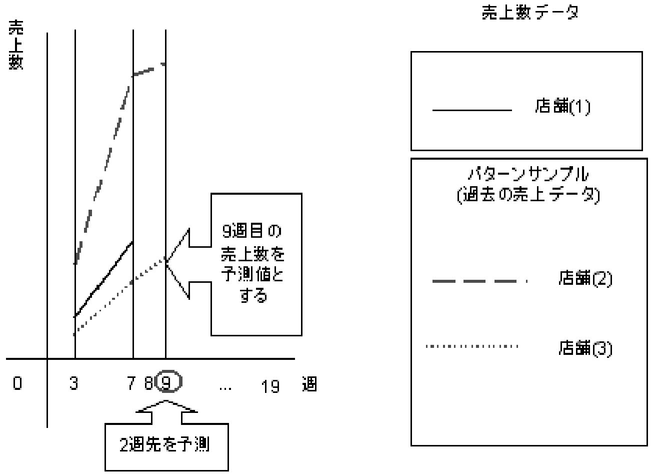

図 6　予測期間を延長したモデル

2.5.1　モデル化

1. $L = S$ とする.
2. 誤差の最も少ないパターンサンプルの出力
$$i^*(L) = \underset{i\,\in\,I}{\arg}\ \min\Big\{\sum_{j=L-6}^{L-2}(y_j - x_{ij})^2\Big\}$$
3. 得られた $x_{i^*(L)}$ を $\hat{y}_{(L)}$ とみなし予測値とする.
4. もし $L = T$ ならば予測終了.
もしそうでないならば $L = L + 1$ として 2 に戻る.

2.6　中央値を予測値とする手順

A,B,C,D のモデルを用い, 中央値を予測値とする手法について説明する. **図 7** で例を示す. 先ほどまでは誤差の最も少ないパターンサンプルのみを出力していたが, 中央値では誤差の少ない 5 店舗を出力する．その 5 店舗で中央値の計算を行い予測値を得る.

モデル A の場合を例に中央値を用いて予測値を求めてみる.

- step0
$K = \{\phi\}$
$l = 0$
- step1 誤差の最も少ないパターンサンプルの出力
$$i^*(L) = \underset{i\,\in\,I}{\arg}\ \min\Big\{\sum_{j=1}^{L-1}(y_j - x_{ij})^2\Big\}$$
$K = K \cup \{i^*(L)\}$
$I = I - \{i^*(L)\}$

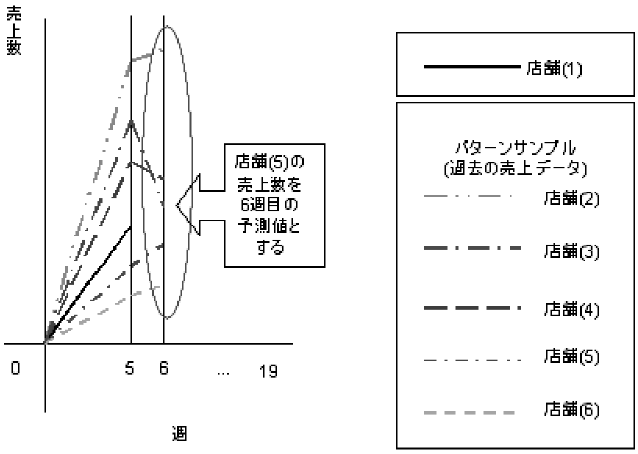

図 7　中央値を用いた予測値

$l = l + 1$

- step2
  $l = 5$ ならば step3 に
  $l < 5$ ならば step1 に
- step3 $\hat{y_L} = \underset{i \in K}{\mathrm{median}} \left\{ x_{iL} \right\}$

  で計算終了

同様の方法でモデル B,C,D も中央値を用いて予測値を求めることが出来る.

3　実行結果

3.1　予測の実行結果

2004 年の全ての店舗について予測を行った結果を**表 3**, **表 4** に纏めた.

表 3　最小値を用いた予測結果

	平均誤差 (%)	範囲 (%)	標準偏差 (%)
A	40.8	23.5〜93.5	10.4
B	36.9	21.8〜63.0	9.3
C	35.7	17.2〜87.5	11.8
D	41.2	20.2〜85.6	11.1

表 4　中央値を用いた予測結果

	平均誤差 (%)	範囲 (%)	標準偏差 (%)
A	37.2	24.4〜62.9	7.3
B	32.7	18.3〜55.1	7.4
C	32.7	18.2〜92.2	10.8
D	44.8	18.4〜231.9	24.4

4　考察

A のモデルの最小値は平均誤差が 40.8 %と多いものとなっている. また, 中央値は誤差を少なくすることが出来, 標準偏差が小さくなった.

B のモデルでは中央値の平均誤差が 1 番少ない. 誤差の範囲もある程度抑えることが出来, 標準偏差の小さい結果が出せた. この方法は A のモデルより 4 %以上誤差を少なくすることができ, より良い結果を残すことができた.

C のモデルは愛知の店舗などの売上数の多い店舗は予測の精度が上がった. しかし売上数の少ない店舗では売上傾向の近いパターンがないため, 誤差が多くなった. 中央値の平均誤差は B のモデルと変わらなかったが, 標準偏差が大きくなった.

D のモデルは平均誤差が最小値, 中央値ともには 40 %以上となり, 大きいものとなった. さらに中央値の誤差の範囲がとても大きく, 標準偏差が非常に大きなものとなった.

5　おわりに

結論として, 最も精度の高い結果が得られるモデルは B のモデルである. また, 中央値を用いた予測をすることが良いと考えられる.

昨年度より精度の高いモデルを考案することができたが, 需要予測はまだまだ難しい.

過去の蓄積するデータによって精度が異なる. よって予測には 10 年程度の長期間のすだれの売上数データを集め, 様々なパターンが必要と考えられる. 精度を上げるためには気温や天候等のモデルの考案をしていく必要がある. 本研究が今後の需要予測に少しでも役立てれば幸いである.

6　謝辞

本研究を進めるにあたり, さまざまなご指導をして下さいました南山大学数理情報学部数理科学科の鈴木敦夫教授には深く感謝いたします. 共に研究をした鈴木ゼミの皆様, 多くの助言をしてくださった石川総一郎先輩に深く感謝いたします. カーマホームセンターの方々にはお忙しい中, 時間を割いていただき感謝いたします.

参考文献

[1] 気象庁電子閲覧室, http://www.data.kishou.go.jp/
[2] 松本京子, 坂口祥子：ホームセンターの季節品の需要予測, 南山大学数理情報学部数理科学科卒業論文, (2004).