

## 通信販売業における受注予測モデル

Forecasting Model in the Mail Order Industry

松田 芳雄, 海老原 純二

**要約** 通信販売業では、商品の仕入れを買取りで行っているため、売り切れによる機会損失や売れ残りによる過剰在庫が発生して深刻な問題になっている。そのため、商品発注を適正化するために、受注予測システムの開発に力を入れているが、満足なものは完成していない。

本稿で報告する受注予測は、通販会社でカタログを配布し始めて2~4週間後に、それまでの受注傾向をもとに個々の商品の最終受注数を予測するものである。現在3社で本番または試行を行っているが、いずれも中率で10~20%程度の改善がなされている。

今回開発した予測モデルは、発売日があり、販売の有効期限が決められた場合の途中の受注傾向から最終受注数を予測するためのものである。この販売形態は通信販売だけでなく、鉄道や航空機の予約や旅行の販売などでも同様であり、それらへの適用も可能である。

**Abstract** The mail order companies must sell out all of the commodities purchased for marketing. Serious problems may occur as a result of opportunity loss caused by stock outs and/or excess stock caused by poor demand. In order to avoid these problems, the companies have spent a lot of time and effort in developing a demand forecasting system intended to order the reasonable quantity of goods, and however, there exists and did exist no forecasting system with the sufficient precision.

This report describes a forecasting model that determines the quantity of orders finally received for individual commodities based on ordering trends during the early period of mail order life cycle (2-4 weeks after catalog distribution)

The standard model in this industry has used commodity attributes for sales forecasting in despite of a serious shortcoming until comparatively lately. Our model described in this report resolves this weak point of the traditional model, and forecasts the necessary and sufficient stock more accurately at the initial stage of mail order activities for an individual commodity. This model is now operating in production or test running state at three major mail order companies in Japan, and demonstrates a marked improvement of 10-20% of accuracy in demand forecasting in these customers.

### 1. はじめに

通信販売業は、商品の仕入れを買取りで行うことにより、廉価な商品を提供している。一方、買取りであるがために、売り切れによる機会損失や売れ残りによる過剰在庫が問題になる。顧客にカタログを配布して葉書や電話で注文を取るカタログ通販では、カタログに掲載した商品と全く同じものしか販売できないため、少数の売れ残り商品は再度販売することができない。衣料品ならば裁断して処分することになる。そのため、売れ残りによる過剰在庫は収益に悪い影響を与えることになり、商品発注数の適正化は通販各社にとって重要な課題である。発注数を決めるためには各商品の正確な受注数の予測が必要であるが、十分な精度を持つ予測システムは未だ確立していない。本稿では、今まで通販各社が行ってきたものとは別の考え方に基づいた受注予

測の方法について記述する。

一般に、カタログ通販では、カタログの配布前に商品の1回目の発注がなされ、配布後に受注の傾向を見て追加の発注が行われる。したがって、予測はカタログ配布の前と後に必要になる。カタログ配布前の予測を需要予測、カタログ配布後の予測を受注予測と呼んで区別している。ここでは、カタログ配布後の受注予測について述べる。

## 2. 受注予測

カタログ通販では郵便または宅配便でカタログを顧客に配布し、葉書または電話で受注を受け付けている。カタログの配布は部数が多いと1ヶ月間くらいの日数を要し、カタログ配布開始日から受注締め切り日までの日数、すなわちカタログの有効期間は120~210日くらいである(文献<sup>1</sup>参照)。

このとき、カタログ配布開始後2~4週間後に、個々の商品(色サイズ別)の最終的な受注数を予測することを考える。

### 2.1 受注曲線

カタログ配布開始日からの経過日数と受注数の関係を受注曲線と呼ぶ。図1は受注曲線の例であるが、日々の日計受注数と日計受注数を累計した累積受注数が表示されている。日計受注数の傾向を把握するために、7×3項移動平均線を描き加えてある。7項移動平均により曜日の影響を除去し、さらに3項移動平均を行いカタログや注文葉書の到着日のバラツキを除去したものである。

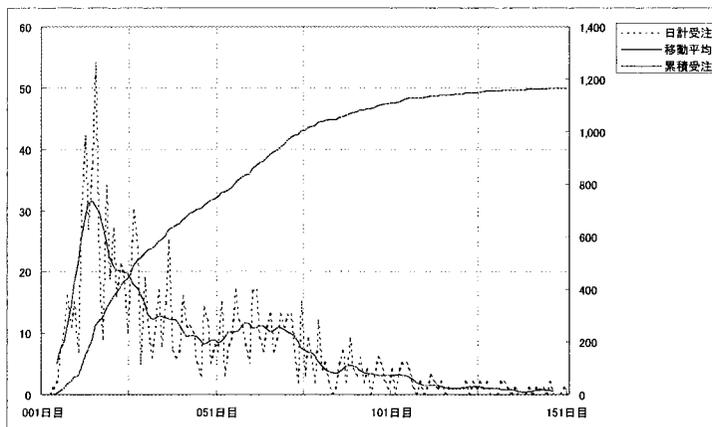


図1 受注曲線の例

予測はカタログ配布開始日から2~4週間後に行われる。予測を行う日を予測実施日と呼ぶ。受注予測で取り扱う受注データには以下の種類がある。

- ① 日計受注数.....日別の受注数
- ② 累積受注数.....カタログ配布開始日からの日計受注数の累積
- ③ 日計受注率.....最終受注数を100としたときの日計受注数の割合
- ④ 累積受注率.....最終受注数を100としたときの累積受注数の割合
- ⑤ 日計受注率2...予測実施日までの累積受注数を100としたときの日計受注数

## の割合

## ⑥ 累積受注率 2... 予測実施日までの累積受注数を 100 としたときの累積受注数の割合

ここで、最終受注数はカタログ有効期限までの累積受注数である。

## 2.2 予測の考え方

受注予測では予測実施日にはいくらかの受注がすでに存在している。したがって、予測実施日までの受注情報をもとに最終受注数を予測するのが自然な考え方である。

## 1) 予測実施日の累積受注数と最終受注数の関係

図 2 は、あるカタログのカタログ配布開始後 28 日目の累積受注数と、150 日目の最終受注数の関係である。

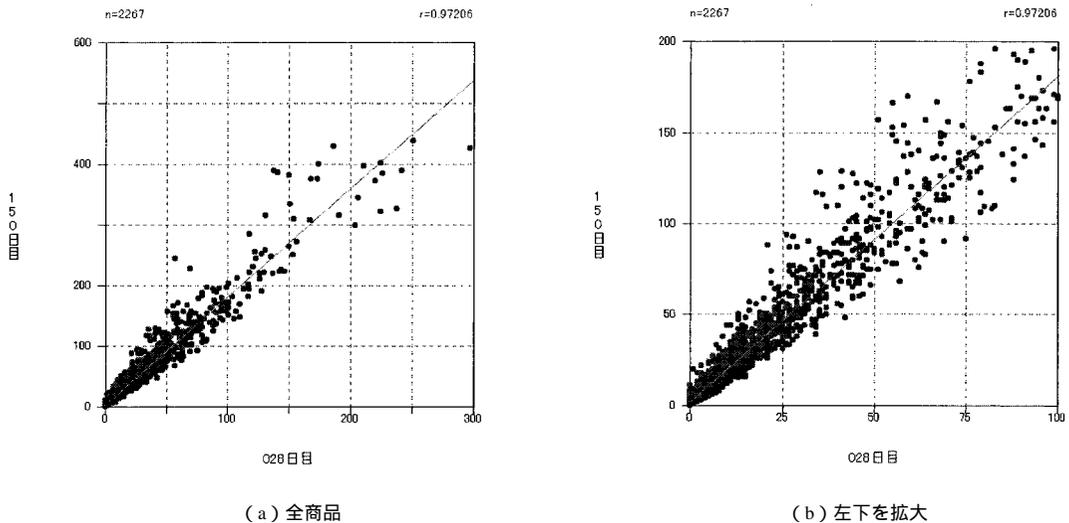


図 2 28 日目対 150 日目の累積受注数の関係

図 2 より、28 日目と 150 日目の累積受注数の相関係数は 0.97 で高い数値を示している。この関係を利用して予測を行うことを考える。

## 2) 予測式

ここで、予測を 28 日目の累積受注数だけで行うと仮定して、予測値を以下の単回帰式で求めることにする。

$$(\text{最終受注数}) = a_0 + a_1 \times (\text{28 日目累積受注数}) \quad (1.1)$$

このとき、予測値は図 2 の斜めの直線上の値になる。たとえば、28 日目の累積受注数が 50 の場合、150 日目の累積受注数は図 2 より約 90 である。150 日目の実績は 160 くらいのもあるし、60 くらいのももあり、予測は大きくはずれることになる。

## 3) 精度向上の考え方

各商品を図 3 のように回帰直線の傾きの違いにより層別し、各層毎に式 (1.1) のような予測式を作成できたとすれば予測精度は向上する。図 3 は各商品を 28

日目の累積受注率で層別したものである。

各層の相関係数は以下のように全体の相関係数 0.97 よりも高くなっており、予測精度の向上が期待できる。

- ・パターン (1)      28日目受注率 (低い)      相関係数: 0.98
- ・パターン (2)      28日目受注率 (高い)      相関係数: 0.99
- ・パターン (3)      28日目受注率 (中間)      相関係数: 0.99

ただし、累積受注率の計算には本来予測すべき最終受注数が必要である。予測実施日にはこの値はわからないので、このような層別は不可能である。そこで、一般には累積受注率の代わりに商品分類などの要因で商品を層別して予測を行っている。

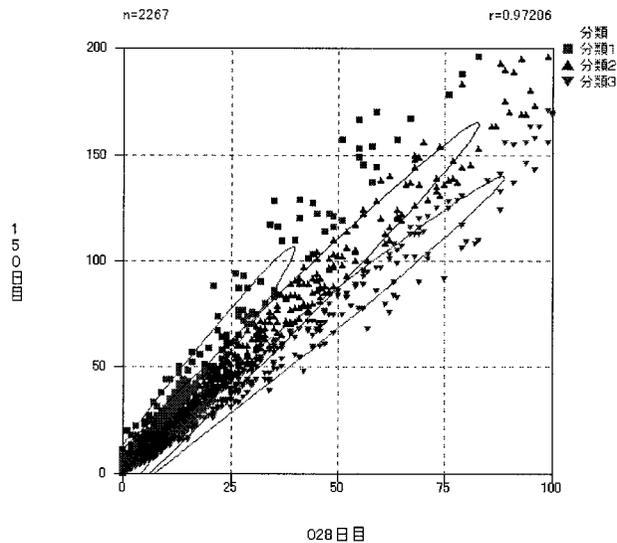


図3 28日目対150日目の累積受注数の関係 (累積受注率で層別)

#### 4) 商品情報による層別

多くの通販会社では累積受注率に代わる層別因子として、商品分類や商品定価などの商品情報を利用している。図4、図5は、1999年と2000年の同カタログ同号 (シーズン) の、28日目と150日目の累積受注数の関係を、商品分類別と商品定価別に示したものである。

層別により予測精度が向上するためには、まず層別したことにより28日目と150日目の相関係数が高くなる必要がある。図4、図5より層別相関は全体相関に較べ大きく改善はされていない。次に、予測が1999年のデータで予測式を推定して2000年に対して行われるとすれば、両年の相関構造は同じである必要がある。すなわち、1999年と2000年の予測式の傾きが同じでなければならないが、年により別々の傾向となっている。以上より、商品情報による層別では、相関係数も改善されないし、年による傾向も異なり、予測精度の改善は期待できない。

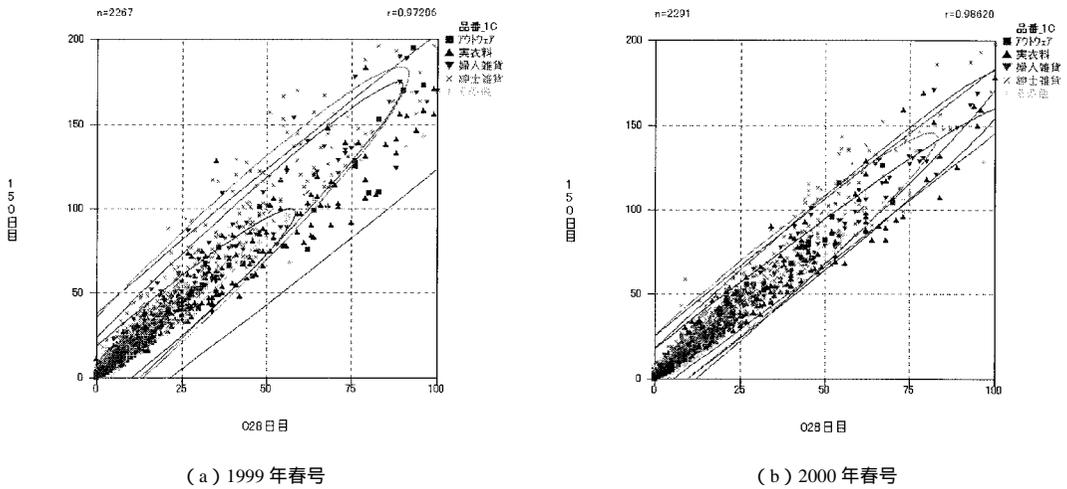


図 4 28 日目対 150 日目累積受注数の関係 (商品分類別)

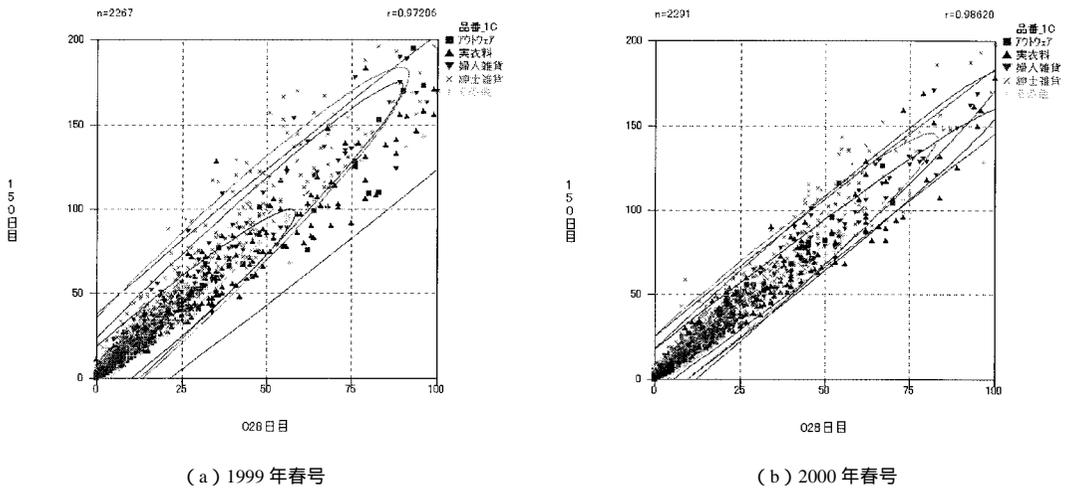


図 5 28 日目対 150 日目累積受注数の関係 (商品定価別)

実際に、1999 年のデータで予測式を作成し、2000 年を予測した結果を以下に示す。

- ・単純予測 (層別なし) 的中率 55.4%
- ・層別予測 (商品分類別) 的中率 47.9%
- ・層別予測 (商品定価別) 的中率 48.7%

的中率は誤差率が ±20% 以内の割合である。誤差率は以下の式で定義する。今後、的中率により予測の精度を評価するものとする。

$$(\text{誤差率}) = \{(\text{実績値}) - (\text{予測値})\} \div (\text{実績値}) \times 100 \quad (1.2)$$

的中率は層別を行わない単純予測よりも悪くなっており、商品情報による層別は予測精度向上に効果がないことがわかる。

ただし、商品定価などは予測の精度向上に全く効果がないわけではない。単価



表1のデータの受注パターン別予測的中率は61.5%であった。単純予測的中率が55.4%であるから、受注パターン別予測の効果はあるといえる。

以下に、受注パターン別予測の詳細について述べる。

### 3. 受注パターン別予測

#### 3.1 分析用データ

##### 1) 分析用データの種類

受注予測に使用する分析用データは推定用データと予測用データの二つを用意する。推定用データは予測式や受注パターンの判別式など予測モデルを推定するためのデータで、各商品のカタログ配布開始日からカタログ有効期限までの受注データである。予測用データは予測対象のデータで、予測対象商品のカタログ配布開始日から予測実施日までの受注データである。一般には、推定用データは予測用データの同種カタログの前年同号とする。受注データには日計および累積の受注数と受注率がある。また、各データにはカタログ配布開始日や配布部数などの媒体情報、商品分類や商品定価などの商品情報が付随している。

##### 2) 受注データの補正

図1の受注曲線で見たとおり受注データは曜日の影響を強く受けている。土曜日や日曜日に葉書で到着した受注は月曜日にまとめて処理されることが多く、土曜日、日曜日の受注は少なく月曜日の受注が多くなる傾向がある。そのため、推定用データと予測用データのカタログ配布日の曜日が異なる場合には両者を直接比較することができない。この場合、曜日の影響を取り除く必要がある。曜日の影響を除去するために日計の受注データに7項移動平均を行う。また、カタログ到着日のずれなど不規則な変動を取り除くために7項移動平均値にさらに3項移動平均を行う（移動平均については文献<sup>4</sup>参照）。この7×3移動平均値により、受注傾向の把握と推定用データと予測用データの比較が可能になる。ただし、7日単位の累積受注数（率）は曜日や不規則変動の影響を受けないのでそのままの値を用いることにする。

#### 3.2 受注パターンの作成

受注パターン別予測は、図6のように受注の立ち上がりが早いか遅いかにより商品进行分类して予測する方法である。

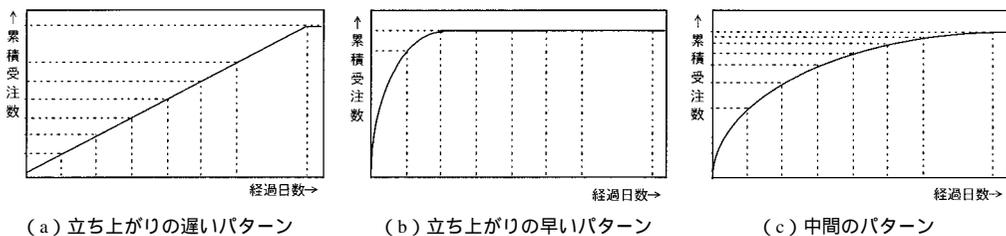


図6 受注パターン

図6の(a)のような立ち上がりの遅いパターンでは、予測実施日の累積受注率は

小さな値になる。(b)のような立ち上がりの早いパターンでは累積受注率は大きな値になる。これを予測実施日の累積受注数対最終受注数の散布図で見ると、立ち上がりの遅いパターン(a)の傾きは大きくなり、立ち上がりの早いパターン(b)の傾きは小さくなっている。

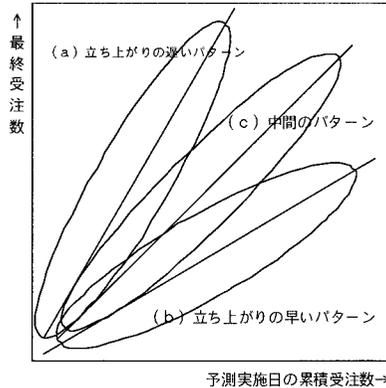


図 7 予測実施日の累積受注数対最終受注数と受注パターンの関係

受注パターンを作成するためには、図7のような散布図上でパターン別回帰直線の傾きの差がなるべく大きくなることと、パターン間の境界が明確になる必要がある。そのためには、予測実施日の累積受注率を使って、各商品の点が楕円状になるように分類すればよい。ここでは、各パターンに商品进行分类したときの累積受注率の級内分散(文献<sup>2</sup>参照)が最小になるように商品の分類を行う。すなわち、式(21)の級内分散が最小になるように、商品进行分类してを決定することになる。受注パターンの数は3程度が適当と実験により確かめられている。

$$\text{級内分散} : V_w = \sum_{j=1}^p \sum_{i=1}^{n_j} (r_{ij} - \bar{r}_{.j})^2 / (n - p) \quad (21)$$

$$\text{ここで } \bar{r}_{.j} = (\sum_{i=1}^{n_j} r_{ij}) / n_j$$

$r_{ij}$  :  $j$  番目のパターンの  $i$  番目の商品の累積受注率

$n_j$  :  $j$  番目のパターンの商品数

$n$  : 全体の商品数

$p$  : 受注パターンの数

### 3.3 受注パターンの判定

3.2節のような方法で商品を受注パターンに分類できるのは推定用データに対してだけである。予測用データに対しては最終受注数がわからないのでパターン分類は不可能である。したがって、予測用データの各商品に対してはどの受注パターンになるかを判定する必要がある。

予測用データの受注パターンの判定は予測実施日までの受注傾向をもとに、各パターンになる確率を判定する以下の三つの線形判別関数(文献<sup>2</sup>参照)により行う。判別関数は、パターン1か否か、パターン2か否か、パターン3か否かを外的基準にし

てそれぞれパターン別に3本作成する．確率へはマハラノビスの平方距離をもとに変換するが(事後確率；文献<sup>3</sup>参照)，三つの確率の合計は1.0にはならないので何らかの正規化を行う．

$$\begin{aligned} \text{パターン (1) になる確率} & : P_1 = \alpha_0 + \alpha_1 \times (\text{予測実施日までの受注傾向}) \\ \text{パターン (2) になる確率} & : P_2 = \beta_0 + \beta_1 \times (\text{予測実施日までの受注傾向}) \\ \text{パターン (3) になる確率} & : P_3 = \gamma_0 + \gamma_1 \times (\text{予測実施日までの受注傾向}) \end{aligned} \quad (2.2)$$

ここで，予測実施日の受注傾向とは以下のようなものである．一つは，カタログ配布開始日から予測実施日までの日計の受注数の推移により商品をいくつかのパターンに分類したものである．分類は各日の受注数を変数にしてクラスター分析( $k$  means法)(文献<sup>5</sup>参照)により行う．このとき，受注数は $7 \times 3$ 移動平均値に対数変換を行ったものを使用する．判別関数へはグループ番号をダミー変数化して説明変数として用いる．これを受注数パターンと呼ぶ．もう一つは，日計の受注率(予測実施日の累積受注数を100としたときの受注率)を使って受注数と同様にパターン化したものである．これを受注率パターンと呼ぶ．

受注数パターンは主に予測実施日までの受注の大きさをパターン化したものであり，受注率パターンは受注の変動の形をパターン化したものである．図8に受注数パターンと受注率パターンの例を，経過日毎の受注数と受注率の箱ひげ図(文献<sup>7</sup>参照)で示す．パターンの数は実験により5くらいが適当である．さらに，判別関数の説明変数に商品分類や定価などの商品情報を付加することにより，判別の精度を向上させることができる．

### 3.4 予測値の計算

#### 1) 予測式の作成

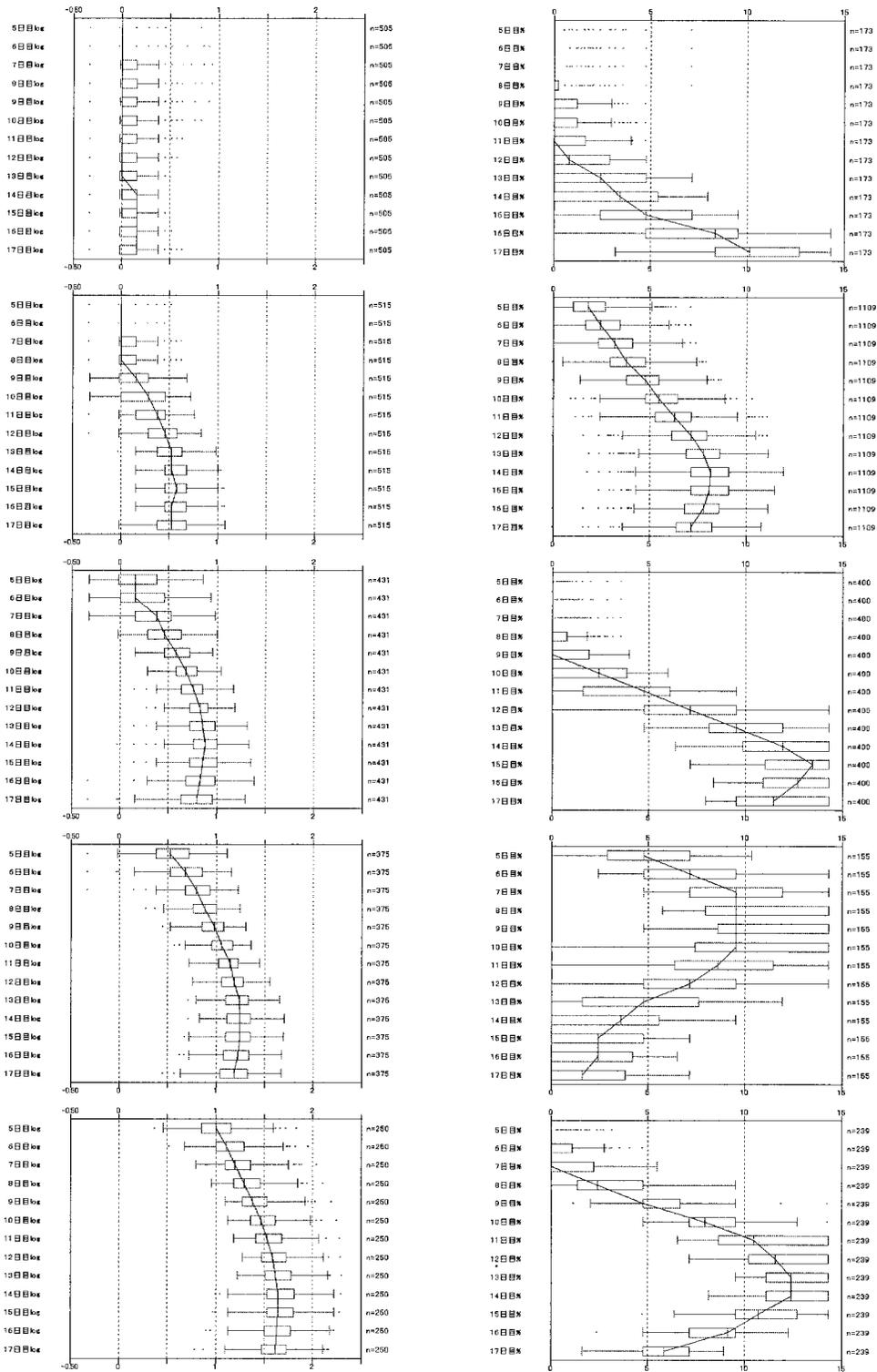
予測値は予測実施日の累積受注数から計算する．予測値を計算する式を予測式と呼び，以下のような単回帰式を受注パターン別に作成する．

$$\begin{aligned} \text{パターン (1) の予測式} & : (\text{最終受注数}) = a_0 + a_1 \times (\text{予測実施日の累積受注数}) \\ \text{パターン (2) の予測式} & : (\text{最終受注数}) = b_0 + b_1 \times (\text{予測実施日の累積受注数}) \\ \text{パターン (3) の予測式} & : (\text{最終受注数}) = c_0 + c_1 \times (\text{予測実施日の累積受注数}) \end{aligned} \quad (2.3)$$

予測式の係数( $a_0, a_1, \dots$ )の推定は，一般の最小二乗法よりも異常値の影響を受けにくいロバスト推定法(文献<sup>9</sup>参照)で行うほうが予測精度は高くなる．また，途中の経過日の情報も取り入れて重回帰式にすることも考えられるが，これらの情報は受注パターンの判定に取り入れられているので，単回帰式でも予測精度はほぼ同じである．

#### 2) 最終予測値の計算

式(2.3)により予測値は受注パターン別に三つ得られるので，これらから最終的な予測値を計算する必要がある．受注パターンの判定確率が正確ならば最大確率を持つパターンの予測値を採用すればよいが，予測が早い段階に行われる場合には，判定のための受注数パターンや受注率パターンの精度が悪く，判定確率も信頼性が低くなっている．その場合には，各判定確率を重みとする加重平均に



(a) 受注文パターン

(b) 受注率パターン

図 8 予測実施日までの受注傾向

より最終受注数を求める方が高い精度が得られる．

$$\begin{aligned} (\text{最終受注数}) = & \{ P_1 \times (\text{パターン (1) の予測値}) \\ & + P_2 \times (\text{パターン (2) の予測値}) \\ & + P_3 \times (\text{パターン (3) の予測値}) \} / (P_1 + P_2 + P_3) \end{aligned} \quad (24)$$

$P_1, P_2, P_3$  はそれぞれ式 (22) で得られるパターン (1), (2), (3) の判定確率である．

### 3.5 受注パターン別予測の手順

以上の受注パターン別予測の手順をまとめると以下ようになる．

#### 1) 推定用データの選択

どのカタログのデータから予測するか決定する．通常は同カタログの前年同号を使用する．カタログ発行日や配布方法，商品構成などが似ているカタログのデータが望ましい．

#### 2) 推定用データ/予測用データの補正

曜日補正，カタログ到着日補正のため受注データに  $7 \times 3$  項移動平均を行う．

#### 3) 受注パターンの作成

推定用データの各商品を受注パターンに分類する．予測実施日の累積受注率をもとに各パターン内の分散が最小になるように各商品を分類する．受注パターンの数は経験的に 3 パターンが適当である．

#### 4) 受注パターンの判定

予測用データの各商品がどの受注パターンになるか判定する．各パターンへの帰属判定確率を予測する判別関数を推定する．判別関数の説明変数には予測実施日までの受注傾向（受注数パターン，受注率パターン）を用いる．有効ならば商品情報などを付加して精度を向上させることができる．受注パターン別予測の精度はこの判別関数の精度で決定される．

#### 5) 予測式の作成

予測値の計算式を受注パターン別に作成する．通常は予測実施日の累積受注数と最終受注数の単回帰モデルとして作成するが，他の日の受注データやその他の要因を加えて重回帰モデルとすることもできる．回帰係数の推定は異常値の影響を受けないロバスト推定法を用いる．

#### 6) 予測値の計算

受注パターン別予測値と各パターンへの帰属確率から最終受注数を求める．受注パターンへの帰属確率を重みとする加重平均値を最終予測値とする．

## 4. 予測値の補正

予測誤差は予測値のバラツキと偏りにより発生する．予測値のバラツキは，100 を 80 と予測したり 120 と予測したりする不確実性であり，予測誤差の分散の大きさで表現される．これは，予測に関連があると思われる要因をモデルに取り込むことにより誤差分散を小さくすることができる．受注パターン別予測では，受注パターンの判定式に必要な要因を取り入れて精度の向上を計ることができる．

一方、予測値の偏りは、常に大きめに予測したり常に小さめに予測したりする予測値のずれである。これは、1年前のデータを使って予測を行っているため、景気の動向やカタログなどの違いにより引き起こされると思われるが、偏りの原因は多様であり特定は困難である。そこで、予測値自身を補正して偏りをなくす方法を考える。

#### 4.1 予測値補正の考え方

##### 1) 予測値の偏りの原因

予測値の偏りの原因として、推定用データや推定方法の影響でモデル自身に偏りがある場合と、推定用データと予測用データの違いにより予測値が偏る場合の二つおりが考えられる。前者を予測モデルによる偏りと呼び、後者をカタログの違いによる偏りと呼ぶ。

- ・予測モデルによる偏り.....推定用データや推定方法の影響でモデル自身に偏りがある。
- ・カタログの違いによる偏り.....推定用データと予測用データの違いにより予測値が偏る。

##### 2) 予測モデルによる偏りの尺度

予測モデルによる偏りの尺度として、推定用データのカタログ全体の平均受注数(1商品当り)の実績値と予測値の比を考える。予測値の補正は偏りの尺度を使って行うものとし、これを補正係数1とする。

$$(\text{補正係数1}) = (\text{推} \cdot \text{実績平均受注数}) / (\text{推} \cdot \text{予測平均受注数}) \quad (3.1)$$

ここで、

(推・実績平均受注数): 推定用データの実績受注数の平均値(1商品当り)

(推・予測平均受注数): 推定用データの予測受注数の平均値(1商品当り)

##### 3) カタログの違いによる偏りの尺度

カタログの違いによる偏りの尺度として、推定用データと予測用データの累積受注率の比を考える。これを補正係数2とする。

$$(\text{補正係数2}) = (\text{推} \cdot \text{受注率}) / (\text{予} \cdot \text{受注率}) \quad (3.2)$$

ここで、(推・受注率): 推定用データの予測実施日の累積受注率

(予・受注率): 予測用データの予測実施日の累積受注率

注) 実際には(予・受注率)は予測実施日には不明である。

##### 4) 全体の偏りの尺度

予測モデルによる偏りの尺度(補正係数1)とカタログによる偏りの尺度(補正係数2)は独立の事象とみなせるので、全体の偏りの尺度はそれぞれを掛け合わせたものとする。

$$(\text{補正係数}) = (\text{補正係数1}) \times (\text{補正係数2}) \quad (3.3)$$

##### 5) 予測値の補正

予測値の補正は補正前の予測値に補正係数を乗じて行う。

$$(\text{補正後予測値}) = (\text{補正前予測値}) \times (\text{補正係数}) \quad (3.4)$$

#### 4.2 補正係数の推定

予測用データの予測実施日の累積受注率(予・受注率)がわかっているものとして計算した補正係数を理論補正係数とする。実際には、(予・受注率)は不明であるの

で推定する必要がある。(予・受注率)はカタログ全体の予測実施日の累積受注数や媒体情報(配布部数, 配布月, 掲載商品数, 平均定価, 等)を説明変数として, 以下のような回帰モデルにより推定することにする。

$$\begin{aligned} (\text{予・受注率}) = & a_0 + a_1 \times (\text{予測実施日の累積受注数の平均値}) \\ & + a_2 \times (\text{配布部数}) \\ & + a_3 \times (\text{配布月}) \\ & + a_4 \times (\text{掲載商品数}) \\ & \vdots \end{aligned} \quad (35)$$

(予・受注率)の推定値から補正係数2を計算し, さらに全体の補正係数を計算することができる。推定された(予・受注率)から計算した補正係数を推定補正係数と呼ぶ。予測値の補正は推定補正係数を使って行うことになる。表2に補正前と補正後の的中率の例を示す。「99春 00春」は99年春号のデータを使って00年春号の予測を行ったことを示す。予測値補正によりの中率はいずれも改善されており補正の効果があることがわかる。また, 推定補正係数による結果は理論補正係数による結果とほぼ同じであり, 予測用データの受注率(予・受注率)の推定もうまく行っていることがわかる。

表2 予測値補正による的中率の変化

番号	予測	補正前	理論補正係数による補正	推定補正係数による補正
1	99春→00春	61.5%	66.9%	63.1%
2	99夏→00夏	58.0%	67.1%	65.3%
3	99秋→00秋	62.8%	72.1%	70.9%
4	99冬→00冬	57.1%	64.8%	64.5%

注) 理論補正係数:(予・受注率)の実績値から計算した補正係数

推定補正係数:(予・受注率)の推定値から計算した補正係数

## 5. カタログ配布の問題

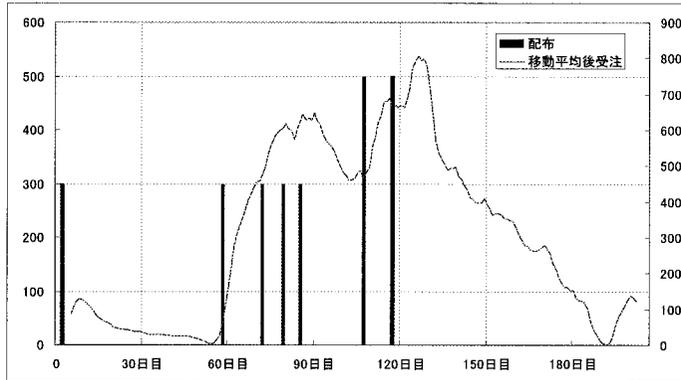
### 5.1 カタログの配布方法の違い

通常, 予測モデルは同カタログの前年同号のデータを使って作成される。この場合, 予測用データと推定用データのカタログ部数や配布方法, 商品内容などが同一であることが前提である。同種のカタログであれば商品の内容は大きく変わらないが, 配布の部数や配布方法は年毎に変えることがある。カタログの配布には以下の3種類がある。

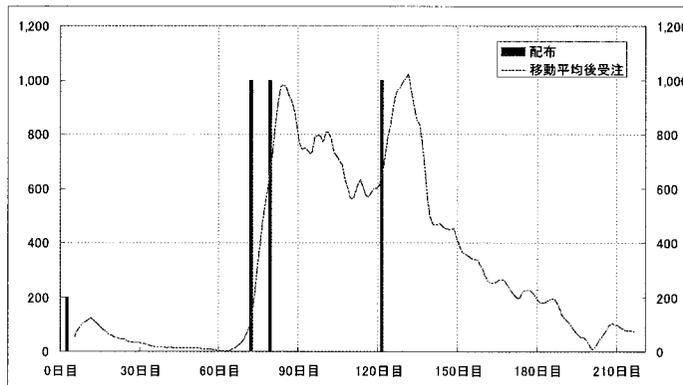
- ① 先行号.....本番カタログ配布の2~3ヶ月前に受注の調査等のためにテスト的に配布されるカタログである。本番号の5~10%程度の部数が配布される。内容は本番号とほぼ同じであるが5~7%の値引きがされることもある。
- ② 本番号.....本来の配布カタログである。部数が多いときには1~2ヶ月くらいの期間で配布される。基本的には先行号と同じカタログであるが, 表紙や一部商品に変更がある場合がある。

- ③ リメール...受注が好調のときや逆に低調のときに本番号配布終了後1~2ヶ月後に追加で配布されるカタログである。内容は本番号とほぼ同じである。

先行号とリメールの有無や時期/部数および本番号の配布期間/部数など毎年同じにはなっていない。配布の方法が異なれば受注曲線が変わり、予測精度に影響を及ぼすことになる。



(a) 1999年春号



(b) 2000年春号

図9 日計配布数と日計受注数(7×3移動平均値)

図9は1999年と2000年の春号の配布数と受注数である。棒グラフが日計配布数である。本番号配布開始日の2ヶ月前に先行号が配布され、2ヶ月後にリメールが配布されている。先行号の部数、本番号の配布パターン、リメールの回数などが異なっている。日計受注数の7×3移動平均値が鎖線で示されている。配布数に対応して受注数が変動していることがわかる。

ここで、配布の違いは以下の3通りがある。

1) 総部数の違い

本番号または先行号、リメールを合わせた総配布部数が推定用データと予測用データで異なる場合である。受注予測では予測実施日の累積受注数をもとに予測

を行うので、総部数が変化しただけでは予測式（式 2 3）の説明変数と目的変数が同じように変化するため、大きな影響は受けない。

## 2) 配布率の違い

予測実施日以前に配布された部数と以後に配布する部数の比率の違いである。総配布部数に対する予測実施日までの累積配布部数の割合を配布率とする。配布率は受注率に影響を与え、予測値の偏りとなって現われる。したがって、予測値の補正において受注率（予・受注率）を推定するときの要因に配布率を取り入れることによりこの違いを除去することができる。

## 3) 配布パターンの違い

経過日数別の配布数の分布を配布パターンと呼ぶ。推定用データと予測用データの配布パターンが異なる場合がある。受注パターン別予測では予測実施日までの受注のパターンをもとに予測を行うので、配布のパターンが異なると予測は大きくはずれることになる。配布パターンの違いの補正が必要になる。

## 5.2 配布数と受注数の関係

配布パターンの違いを補正するために配布数と受注数の関係を定式化することを考える。カタログの配布と受注の関係を以下のように仮定する。

- ・受注数はカタログの配布数に比例して増加する。
- ・日々の受注数は配布からの経過日数に従って変化する。

経過日数に伴う受注数の変化をカタログの効力が変化したと考え、カタログ効力を経過日数  $t$  の関数  $f(t)$  で表すことにする。カタログ効力は、時刻におけるカタログの寿命であり、顧客がカタログを保有している確率とも考えられる。このとき、1 回（1 日）のカタログ配布数を  $d$  とし、配布から  $t$  日目の受注数を  $O_t$  とすると、 $d$  と  $O_t$  の関係は以下のように表すことができる。レスポンス率はカタログ有効期間中一定であると仮定する。

$$O_t = rd f(t) \quad (4.1)$$

ただし、 $O_t$ ：カタログ配布開始日から  $t$  日目の受注数

$d$ ：カタログ配布数

$r$ ：レスポンス率（定数）

$f(t)$ ：カタログ効力曲線（カタログ保有確率）

ここで、カタログの効力（カタログの保有確率）は配布からの経過日数に従って、最初は増大し、ある時期を境に減少すると仮定する。このような分布としてワイブル分布（文献<sup>[10][11]</sup>参照）を仮定する。

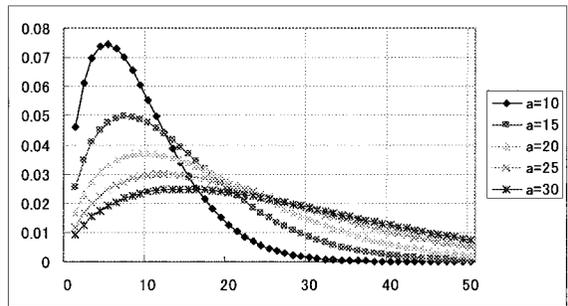
$$f(t) = \left(\frac{b}{a^b}\right) t^{b-1} \exp\left[-\left(\frac{t}{a}\right)^b\right] \quad t > 0 \quad (4.2)$$

ただし、 $a$ ：尺度パラメータ

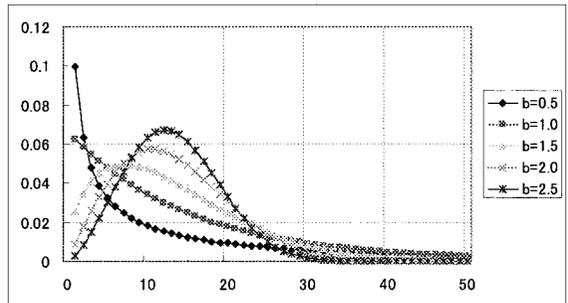
$b$ ：形状パラメータ

式（4.2）はワイブル分布の確率密度関数で、形状パラメータが 1.5 の場合と、尺度パラメータが 15 の場合を図 10 に示す。

次に、カタログの配布を複数の日に分けて行う場合を考える。 $i$  日目のカタログ配



(a)  $b = 1.5$  の場合



(b)  $a = 15$  の場合

図 10 ワイブル分布の確率密度関数

布数を  $d(i)$  とすると,  $i$  日目の配布から  $t$  日目の受注数  $o_t^{(i)}$  は以下ようになる.

$$o_t^{(i)} = rd^{(i)}f(t) \tag{4.3}$$

ただし,  $o_t^{(i)}$ :  $i$  日目のカタログ配布開始日から  $t$  日目の受注数

$d^{(i)}$ :  $i$  日目のカタログ配布数

このとき, 1 日目の受注数は, 1 日目の配布の 1 日目の受注  $o_1^{(1)}$  だけである. 2 日目の受注数は, 1 日目の配布の 2 日目の受注と, 2 日目の配布の 1 日目の受注の和  $o_2^{(1)} + o_1^{(2)}$  となる. 以下同様に,

- ・ 1 日目の日計受注数  $o_1^{(1)}$
- ・ 2 日目の日計受注数  $o_2^{(1)} + o_1^{(2)}$
- ・ 3 日目の日計受注数  $o_3^{(1)} + o_2^{(2)} + o_1^{(3)}$
- ⋮

となる. 以上のような方法で配布と受注の関係を定式化することができる. 1999 年春号の日計受注数 (7×3 移動平均値) に対し, 配布数から受注数を推定した結果を図 11 に示す. 点線が日計受注数の実績で鎖線が推定値である. ワイブル分布のパラメータ, は非線形計画法のひとつであるシンプレックス法 (文献<sup>12</sup>参照) で推定し, レスポンス率は実績値を用いた.

### 5.3 配布の違いの補正

式(4.1)からは 1 回のカタログ配布についての受注曲線を求めることができるが,

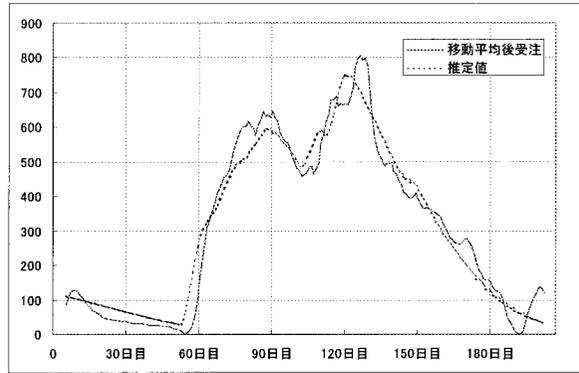


図 11 日計受注数(7×3 移動平均)の推定結果(99 年春号)

カタログを何日かに分けて配布している場合、日計受注数は各配布からの受注の合成値となっており、正確な受注パターンは求められない。本来、受注パターンは1回毎のカタログ配布について求める必要があるので、何らかの調整を行う必要がある。

たとえば、推定用データと予測用データの受注数を単位部数(たとえば100万部)を1回だけ配布した場合に変換して予測を行えばよい。そのためには、以下のような方法が考えられる。

- ① 受注数を単位部数当りの受注数に変換する。  

$$(\text{単位部数当り受注数}) = (\text{受注数}) / (\text{配布部数}) \quad (4.5)$$
- ② ただし、配布部数の代わりにカタログの効力を考慮した有効部数を使用する。  

$$(\text{有効部数}) = (\text{配布部数}) \times f(t) \quad (4.6)$$
 $f(t)$  は式(4.2)のようなカタログ効力曲線である。  

$$(\text{単位部数当り受注数}) = (\text{受注数}) / (\text{有効部数}) \quad (4.7)$$
- ③ 単位部数当り受注数を使って受注パターンの作成や判定を行う。

以上のように配布の違いを補正して予測した結果を表3に示す。夏号を除いて的中率は大幅に改善されている。したがって、配布数と受注数の関係を定式化し受注データを補正することにより、配布の違いを考慮して予測することも可能であることがわかる。

表 3 配布の違いの補正による的中率の変化

番号	予測	補正前	配布の違いの補正後
1	99春→00春	24.5%	61.0%
2	99夏→00夏	55.3%	36.8%
3	99秋→00秋	22.7%	47.5%
4	99冬→00冬	26.1%	56.8%

## 6. おわりに

現在、不況の影響やインターネットなど新しい媒体の出現などに伴い、カタログの統廃合や配布方法の変更などが頻繁に行われている。そのため予測が不安定になって

きている。カタログの変化に追隨できる汎用的なモデルの開発が今後の課題となる。配布と受注の関係の定式化は、汎用モデル開発の重要な基礎技術であり、これを確立することにより他の媒体などと組み合わせた予測なども可能になる。

本稿で取り上げたモデルは、発売日があり、販売の有効期限が決められた場合の途中の受注傾向から最終受注数を予測するためのモデルである。この販売形態は通信販売だけではなく、鉄道や航空機の予約や旅行の販売なども同様と考えられる。これらの分野への今後の適用が期待される。

- 参考文献**
- [ 1 ] 日本通信販売協会編 (1994): 「通信販売ファクトブック'94年版」日本通信販売協会.
  - [ 2 ] 奥野忠一, 久米均, 芳賀敏郎, 吉沢正 (1981): 「多変量解析法」日科技連出版社.
  - [ 3 ] 杉山高一 (1985): 「多変量データ解析入門」朝倉書店.
  - [ 4 ] 中村隆英, 新家健精, 美添泰人, 豊田敬 (1983): 「経済統計入門」東京大学出版会.
  - [ 5 ] 大隅昇, L. ルパール, 他 (1994): 「記述的多変量解析法」日科技連出版社.
  - [ 6 ] H.Charles Romesburg (1989): CLUSTER ANALYSIS FOR RESEACHERS, Robert E.Krieger.
  - [ 7 ] John W. Tukey (1977): EXPLORATORY DATA ANALYSIS, Addison Wesley.
  - [ 8 ] 渡部洋, 鈴木則夫, 山田文康, 大塚雄作 (1985): 「探索的データ解析入門」朝倉書店.
  - [ 9 ] D. C. Hoaglin, F. Mosteller, J. W. Tukey (1985): EXPLORING DATA TABLES, TRENDS, AND SHAPES, John Wiley & Sons.
  - [ 10 ] 奥野忠一監訳 (1988): 「寿命データの解析」日科技連出版社.
  - [ 11 ] 真壁肇 (1985): 「信頼性工学入門」日本規格協会.
  - [ 12 ] 今野浩, 山下浩 (1978): 「非線形計画法」日科技連出版社.
  - [ 13 ] 芳賀敏郎, 橋本茂司 (1980): 「回帰分析と主成分分析」日科技連出版社.
  - [ 14 ] 日本銀行調査統計局編 (1975): 「計量経済学の基礎と応用」東洋経済新報社.

**執筆者紹介** 松田 芳雄 (Yoshio Matsuda)

1974年慶応義塾大学工学部管理工学科卒業。同年日本ユニシス(株)入社。オペレーションズ・リサーチ, 統計解析関係のシステム開発に従事。現在, 第一ソフトウェアサービスセンターソリューションサービス室長。日本オペレーションズ・リサーチ学会, 日本シミュレーション学会, 情報処理学会会員。

海老原 純二 (Junji Ebihara)

1989年九州大学理学部数学科卒業。1992年日本ユニシス(株)入社。多変量解析等を利用したデータ解析業務と客先向け解析システムの構築業務に従事。現在は, データマイニングソリューション「MiningPr 21」の開発業務とその客先サポートに携わる。第一ソフトウェアサービスセンターソリューションサービス室に所属。