

# 予測手法

## (4)：選択行動モデル

上田 徹

### 4. 選択行動モデル

これまで述べてきた方法は地域や集団全体の需要を推定・予測するのに有効な方法であったが、マーケティングや交通分析の分野では個人または企業ごとの選択行動に着目した分析が非集計（行動）モデルと称して盛んに検討されてきた[1]。新サービスの導入や販売戦略策定のためには販売額などの量的予測の以前に、顧客が何を望んでいるかを把握できることが望ましく非集計モデルが活用できる場合が多いと考えられる。通信の分野でもその活用が検討され始めている[2]。

そこでここではまずアンケート調査結果や実際の選択結果データから個体（個人、企業または集団を表わす単位であってもよいが以下では「個人」と表わす）ごとの選択構造を明らかにしてくれる方法を概観し、それから量的予測にどう結びつけるかを示す。

#### 4.1 効用と選択行動の関連づけ

多数の選択対象があるときに特定のものは他のものに効用として劣っていないときに選択されるであろうと考えるのはごく自然であろう。極端な例として、いろいろな機能が他よりも劣っているが、自分のまわりでは誰も持っていないということで選ばれた場合、その人にとっての効用は機能ではなく希少性に重きが置かれていたことになる。このように効用は個人ごとに異なるであろうが、それを選択行動なりアンケート調査なりから推定しようとするのである。

選択行動は

- (i) 個人*i*の属性  $Z_i$
- (ii) 選択対象*j*の属性  $X_j$
- (iii) 選択時の環境、いわゆるT.P.O.（時点、場所、場合）などを表わす  $O$

によって変化するであろう[2]。

うえだ とおる NTT通信網総合研究所

〒180 武蔵野市緑町3-9-11

そこで個人*i*にとっての選択対象*j*の効用  $U_{ij}$  を

$$U_{ij} = f(Z_i, X_j, O) \quad (4.1)$$

によって表現する。個人*I*が選択対象*J*を選ぶのは

$$U_{II} \geq U_{ij} \quad (j \neq J)$$

を意味する。対象の選好順位データがあるときにコンジョイント分析を適用でき、選好（購買）結果がわかっているときにはロジット分析を適用できるので、この2方法についてまず述べる。

#### 4.2 コンジョイント分析

個人*i*の選択対象*j* ( $j=1,2,\dots,m$ )の選好度  $P_{ij}$ 、選好順位  $S_{ij}$  がわかっているときには

$$P_{ij} < P_{ik} \text{ または } S_{ij} > S_{ik} \text{ ならば } U_{ij} \leq U_{ik} \quad (j \neq k)$$

となるように式(4.1)を決める。それにはLuce and Tukey [3]から発展したコンジョイント分析を用いることができる。アルゴリズムとしてはKruskalの単調回帰原理[4]を用いる方法(MONANOVA)が発展を促したと言える。このほか

JohnsonのTRADE-OFF [5],[6],

Srinivasan and ShockerのLINMAP[7],

OgawaのRANKLOGIT[8]

などがあり、それらを概説する。なお、個人*i*を特に意識する必要がないときには添字*i*を省略し、選択対象*j*の選好度を  $P_j$ 、選択順位を  $S_j$ 、効用を  $U_j$  と表わす。

##### 4.2.1 効用の推定アルゴリズム

###### (1) MONANOVA

MONANOVAは単調回帰原理にもとづいており、選択対象*i*と*k*のあいだで弱単調性

$$P_i < P_k \text{ または } S_i > S_k \text{ ならば } \hat{U}_i \leq \hat{U}_k \quad (4.2)$$

が成り立つような  $\hat{U}_i$  を媒介変数として、 $\hat{U}_i$  に近い効用  $U_i$  をストレスと呼ばれる非適合度指標

$$\eta = \left\{ \sum_{i=1}^m (U_i - \hat{U}_i)^2 / \sum_{i=1}^m (U_i - E[U])^2 \right\}^{1/2} \quad (4.3)$$

$$; E[U] = \sum_{i=1}^m U_i / m \quad (4.4)$$

を小さくすることで求める方法である（図 4.1 参

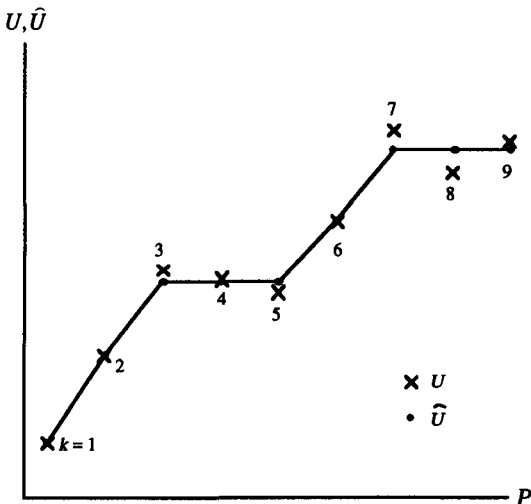


図 4.1  $\{m=9\}$  の場合の単調回帰原理イメージ

照). 初期値または繰り返しの途中で得られている  $U_i$  から弱単調性を持つ  $\hat{U}_i$  を求める手順を以下に示す.

[a] 対象  $i$  は選好度  $P_i$  の小さいもの順に並び変えられているものとする. すなわち

$$i > h \text{ ならば } P_i > P_h$$

とする.

[b]  $\hat{U}_0 = 0, \hat{U}_1 = U_1, k=2$  とする.

[c]  $\hat{U}_k$  に仮の値  $U_k$  を与える.

[d]  $\hat{U}_k \geq \hat{U}_{k-1}$  ならば  $\hat{U}_k = U_k$  と定め,  $k$  に 1 を加えて [c] に戻る.

[e]  $\hat{U}_k < \hat{U}_{k-1}$  ならば

$$\overline{MU}_I = (\hat{U}_k + \sum_{h=1}^k U_{k-h}) / (I+1) \geq \hat{U}_{k-1} \quad (4.5)$$

を満たす最小の  $I$  を求め,

$$\hat{U}_k = \hat{U}_{k-1} = \hat{U}_{k-2} = \dots = \hat{U}_{k-I} = \overline{MU}_I \quad (4.6)$$

とする. たとえば, 図 4.1 で  $\{k=5\}$  のときには

$$\hat{U}_5 = \hat{U}_4 = \hat{U}_3 = \overline{MU}_2 > \hat{U}_2; I=2$$

である.

[f]  $k < m$  ならば,  $k$  に 1 を加えて [c] に戻る.  $k=m$  ならば終了する.

このようにして弱単調性 (4.2) が実現され, ストレス  $\eta$  が計算できる.

$Y$  を, 効用  $U_i$  を規定するパラメタからなるベクトルとすると,

$$Y \rightarrow Y - \Delta \cdot d\eta / dY$$

と更新され, 結果として  $U_i$  も更新される.  $\Delta$  は更新幅であり, [9] と同様に設定すればよい.

MONANOVAでは, データの序列が矛盾を含むときの結果の不安定性, パラメタの初期値のとり方に結果が依存することなどの欠点があり, 選好順位が  $I$  ならそれに選好度として数値  $(m-I+1)$  を与えて被説明変数として数量化I類を適用して初期値を求めたり, パラメタに制約をつけることなどが提案されている. [10]

## (2) TRADE-OFF

TRADE-OFFは, 選好度  $(P_i; P_j)$  間の順序と効用  $(U_i; U_j)$  間の順序との不一致を示す指標

$$E_{ij} = 1 : (P_i - P_j) (U_i - U_j) < 0 \quad (4.7)$$

0: その他

を使った

$$\theta^2 = \sum_{i < j} E_{ij} (U_i - U_j)^2 / \sum_{i < j} (U_i - U_j)^2 \quad (4.8)$$

を基準精度内に収められるように  $U_i$  を求める方法である.  $U_i$  の初期値としては MONANOVA の最後で述べた数量化I類がここでも適用される. パラメタ  $Y$  の修正も MONANOVA と同じく勾配法が用いられ,

$$Y_{r+1} = Y_r - \Delta \cdot d\theta^2 / dY_r \quad (4.9)$$

(添字  $r$  は  $r$  回目の繰り返しの値を示す)

である. ここで, [4] では

$$\Delta = \theta_r \quad (4.10)$$

としているが,  $\theta_r$  が大きい場合はともかく,  $\theta_r$  が小さくなってきた場合には

$$\Delta = \epsilon \theta_r^2 \quad (\text{たとえば } \epsilon=0.1) \quad (4.11)$$

とした方が確実に収束するようである.

## (3) LINMAP

パラメタ  $Y$  は, 次の線形計画法の解として求められる.

$$\text{人工 (artificial) 変数 } \sum_{i=1}^{m-1} \gamma_i \text{ の最小化} \quad (4.12)$$

$$\text{制約: } U_{i+1} - U_i = (x_{i+1} - x_i)^j Y + \gamma_i \geq 0; i=1, 2, \dots, m-1 \quad (4.13)$$

$$(x_m - x_i)^j Y = 1 \quad (4.14)$$

$$Y \geq 0, \gamma_i \geq 0; i=1, 2, \dots, m-1 \quad (4.15)$$

ただし, MONANOVA の手順 [a] と同じく, 対象  $i$  は選好度  $P_i$  (効用  $U_i$ ) の小さいもの順に並び変えられているものとする.

## (4) RANKLOGIT

$j_h$  を選好順位が  $h$  である選択対象番号,  $J_h$  を選好順位が  $h$  位以降である選択対象番号の集合  $\{j_h, j_{h+1}, j_{h+2}, \dots, j_m\}$  とするとき, 選好全順序データ  $\{j_1, j_2, \dots, j_m\}$  が得られる確率

表 4.1 ボイスメール・サービスの属性

|          | j   | k   |     |
|----------|-----|-----|-----|
|          |     | 1   | 2   |
| 入会金      | I   | 2千円 | 2万円 |
| 月額基本料    | II  | 2百円 | 5千円 |
| 使用料/分    | III | 5円  | 50円 |
| 最大メッセージ数 | IV  | 5   | 50  |
| 保存期間     | V   | 1週間 | 1箇月 |

$$L = P(j_1, j_2, \dots, j_m) = P(j_1|J_1)P(j_2|J_2) \cdots P(j_{m-1}|J_{m-1}) \quad (4.16)$$

が最大となるようにパラメタ  $Y$  を求める。ただし、 $P(j_k | J_k)$  は集合  $J_k$  から  $j_k$  が選ばれる確率であり、

$$P(j_k | J_k) = \exp(U_{j_k}) / \sum_{k \in J_k} \exp(U_k); U_j \text{ は式(4.1)} \quad (4.17)$$

で与えられる。

#### 4.2.2 ボイスメール・サービスの分析

ここでは表 4.1, 表 4.2 で規定されるボイスメール・サービスをとりあげる。サービス  $i$  の選好順位  $S_i$  は

$$S_i = (7, 8, 4, 2, 6, 5, 1, 3)$$

であったとする。サービス  $i$  の効用  $U_i$  は

$$U_i = \sum_{j=1}^5 \sum_{k=1}^2 a_{jk} \delta_{jk}(i)$$

$\delta_{jk}(i) = 1$ : サービス  $i$  が属性  $j$  の分類  $k$  に属するとき  
 0: その他

で推定されるものとする。パラメタ  $a_{jk}$  の間には定性的優位性から

$$a_{11} \geq a_{12}, a_{21} \geq a_{22}, a_{31} \geq a_{32}, a_{41} \leq a_{42}, a_{51} \leq a_{52}$$

の関係があり、これを満たしつつ  $\theta^2$  を最小化すると

$a_{11}=0.33, a_{21}=0.88, a_{31}=0.11, a_{42}=0.33$ , その他  $a_{jk}=0$  となり、

$$U_1=0.33, U_2=0, U_3=0.99, U_4=1.32, U_5=0.44,$$

$$U_6=0.77, U_7=1.54, U_8=1.21$$

が得られ、式(4.2)が完全に満たされている。 $a_{jk}$  の大きさからこの人は保存期間には関心がなく、月額基本料を最も重視していることが読み取れる。

### 4.3 ロジット分析

#### 4.3.1 モデルの概要

個人  $i$  の選択肢  $j$  に対する選好結果あるいは購買データ

$\delta_{ij} = 1$ : 個人  $i$  が選択肢  $j$  を選択

0: その他

表 4.2 8種類のボイスメール・サービス (A-H) と対応する属性

| j | I   | II  | III | IV | V   |
|---|-----|-----|-----|----|-----|
| A | 2万円 | 5千円 | 50円 | 50 | 1箇月 |
| B | 2万円 | 5千円 | 50円 | 5  | 1週間 |
| C | 2万円 | 2百円 | 5円  | 5  | 1箇月 |
| D | 2万円 | 2百円 | 5円  | 50 | 1週間 |
| E | 2千円 | 5千円 | 5円  | 5  | 1週間 |
| F | 2千円 | 5千円 | 5円  | 50 | 1箇月 |
| G | 2千円 | 2百円 | 50円 | 50 | 1週間 |
| H | 2千円 | 2百円 | 50円 | 5  | 1箇月 |

が得られているときにはロジットモデルを適用することができる。[1],[10]

個人  $i$  が選択肢  $j$  に対して感じている効用  $U_{ij}$  は、

$$U_{ij} = V_{ij} + \epsilon_{ij} \quad (4.18)$$

すなわち、確定要素  $V_{ij}$  と確率の変動要素  $\epsilon_{ij}$  とから成るものとし、 $\epsilon_{ij}$  が選択肢相互に独立な Gumbel 分布に従うものとする、個人  $i$  が選択肢  $j$  を選ぶ確率は多項ロジットモデル

$$P_{ij} = \exp(V_{ij}) / \sum_k \exp(V_{ik}) \quad (4.19)$$

$$\sum_j P_{ij} = 1 \quad (4.20)$$

で与えられる。 $V_{ij}$  は式(4.1)の右辺で与えられるものとする、式(4.1)に含まれるパラメタは尤度関数

$$L = \prod_i \prod_j P_{ij}^{\delta_{ij}} \quad (4.21)$$

が最大となるように決められる。このモデルには「無関係な代替案からの独立 (Independence from Irrelevant Alternatives)」という性質がある。すなわち、選択肢  $j$  と  $k$  の選択確率の比は

$$P_{jk} / P_{ij} = \exp(V_{jk}) / \exp(V_{ij}) \quad (4.22)$$

となり、 $j, k$  以外の選択肢とは無関係になる。選択肢間に相関がある場合にはこの性質は望ましくなく、これを解決するためにいろいろなモデルが提案されているが、本誌でも [11] で報告されているので、ここではこれ以上言及しない。

#### 4.3.2 ホームテレホン所有世帯の分析

全部で 1002 世帯中ホームテレホン所有する世帯数が 110 のときの世帯属性について検討する。この場合

に選択肢は所有する(j=2)か、しない(j=1)かの2つだけである。そこで選択肢の属性 $x_j(j=1,2)$ に関する情報は使用しない。世帯 $i$ の属性 $Z_i$ としては

- A: 世帯主年齢( $Z_{i1}$ )
- B: 職業( $Z_{i2} \sim Z_{i9}$ )
- C: 家族数( $Z_{i,10}$ )
- D: 収入( $Z_{i,11} \sim Z_{i,22}$ )
- E: 床面積( $Z_{i,23}$ )
- F: 1カ月間の電話料金( $Z_{i,24}$ )

を用いる。このうち職業は8分類、収入は12分類されたカテゴリカルデータであり、その他は数値をそのまま用いた。世帯 $i$ にとっての所有者であること(j=2)の効用の確定要素 $V_{ij}$ は $j$ にかかわらず

$$V_{ij} = \sum_{k=1}^{24} a_k Z_{ik} + a_{25}$$

ただし、

$$a_9 = a_{22} = 0$$

で与えられるものとする。数量化II類(質的データと量的データ混在)を用いた判別分析的中率は74.4%でありあまり高くない。そのパラメタを初期値として尤度関数を最大とするパラメタを求める。ロジットモデルでは

$$V_{i1} = 0,$$

$$V_{i2} = \sum_{k=1}^{24} a_k Z_{ik}$$

すなわち

$$P_{i1} = 1 / \{1 + \exp(V_{i2})\} = \exp(-V_{i2}) / \{1 + \exp(-V_{i2})\}$$

$$P_{i2} = \exp(V_{i2}) / \{1 + \exp(V_{i2})\} = 1 / \{1 + \exp(-V_{i2})\}$$

と考える。このとき、

$$\frac{\partial \log L}{\partial a_k} = \sum_i (\delta_{i2} Z_{ik} - P_{i2} Z_{ik})$$

となり、

$$\sum_i \delta_{i2} Z_{ik} = 0$$

すなわち、 $k$ 番目の属性 $Z_{ik}$ はダミー変数でその属性を持つ世帯には1世帯も所有者がいなるときには尤度は $a_k$ に関して単調減少となるので意味をなさない。そのようなときには $k$ 番目の属性を持つ世帯を分析対象からは必ず必要がある。ホームテレホンの場合には所得が最も低い世帯層( $k=11$ )でこの現象が起きており、その世帯層は除去した。その結果、パラメタは表4.3のようになった。なお数量化II類の場合の $a_k$ の値も表4.3に示したが、 $\{\sum_k a_k^2\}$ は等しくなるように調整してある。

カテゴリカルデータで部分効果 $a_k$ の大きい( $k=4$ )は

表4.3 パラメタ $a_k$ の値

|   | $k$ | $a_k$  |        |
|---|-----|--------|--------|
|   |     | II類    | ロジット   |
| A | 1   | -0.032 | -0.005 |
| B | 2   | -1.458 | -0.498 |
|   | 3   | -1.217 | -0.528 |
|   | 4   | 0.717  | 0.323  |
|   | 5   | -0.946 | -0.471 |
|   | 6   | -0.898 | -0.479 |
|   | 7   | -1.201 | -0.662 |
|   | 8   | -1.208 | -0.776 |
|   | 9   | 0      | 0      |
|   | C   | 10     | 0.438  |
| D | 11  | ---    | ---    |
|   | 12  | 1.177  | 0.924  |
|   | 13  | 1.012  | 0.714  |
|   | 14  | 1.236  | 1.048  |
|   | 15  | 1.220  | 1.171  |
|   | 16  | 0.363  | 0.450  |
|   | 17  | 0.719  | 0.833  |
|   | 18  | 0.916  | 0.926  |
|   | 19  | 1.645  | 1.210  |
|   | 20  | 2.038  | 1.416  |
|   | 21  | 2.364  | 1.298  |
|   | 22  | 0      | 0      |
| E | 23  | 0.030  | 0.012  |
| F | 24  | 0.009  | 0.004  |
|   | 25  | -3.696 | -5.383 |

小企業役員であり、 $\{k=20,21\}$ は高所得者である。なお $\{k=9,k=22\}$ は無回答者である。各アイテムA~Fの効果をレンジの大きさと比較してみる。ただし、 $Z_{ik}$ が数値データの場合はアイテム $j$ に対応する変数番号を $k(j)$ とするとレンジ $R_j$ は

$$R_j = \left| a_{k(j)} \left\{ \max_i Z_{i,k(j)} - \min_i Z_{i,k(j)} \right\} \right|$$

で与えられる。表4.4に示すように $R_j$ で見ると床面積が最も大きく、ホームテレホンの特徴(何台かの子機をつけられる)から考えて納得できる結果である。

#### 4.4 集計法

節4.2, 4.3のモデルは個人の選択行動を予測するのに使うことができる。しかし、サービス戦略や投資計画を立てるためにはエリア単位や集団単位での需要予測が必要であり、節4.3までに述べてきた方法を用い

表 4.4 レンジ  $R_j$

| $j$ | $R_j$ |
|-----|-------|
| A   | 0.35  |
| B   | 1.10  |
| C   | 2.17  |
| D   | 1.42  |
| E   | 4.57  |
| F   | 1.50  |

得るが、節 4.2, 4.3 と同じデータが入手されているときにそれを用いてエリアや集団単位の需要推定、予測ができれば、データを一層、有効に利用できたことになるであろう。ここでは節 4.3 のロジット分析を用いた場合について、エリアや集団単位の需要推定、予測をするための集計法について述べる。

あるエリアや集団  $I$  を構成する個人情報すべてわかっており、ロジット分析により、効用  $V_{ij}$  や式(4.19)の  $P_{ij}$  が得られるならば  $I$  における  $j$  の需要  $D_j(I)$  および  $j$  が選ばれる確率  $P_j(I)$  は

$$D_j(I) = \sum_{i \in I} P_{ij} \quad (4.23)$$

$$P_j(I) = D_j(I) / N(I), \quad N(I): I \text{ 内人数} \quad (4.24)$$

として求めることができるが、 $I$  が非常に小さいか、個人データベース整備に相当の費用をかけ得る商品、サービスでない限り、通常はそういうデータを期待することはできない、そのため次のような方法が提案されている。[1]

#### (1) 平均的な個人を用いる方法

個人  $i$  が  $j$  を選ぶ確率  $P_{ij}$  は効用  $V_{ij}$  の関数であり、 $V_{ij}$  は個人  $i$  の属性  $Z_i$  などの関数なので  $P_{ij}$  は  $Z_i$  の関数となり、

$$P_{ij} = f_j(Z_i) \quad (4.25)$$

と表現できる。そこで  $I$  での  $Z_i$  の平均値  $\bar{Z}_I$  を用いて  $I$  で  $j$  を選ぶ人の割合を

$$P_j(I) = f_j(\bar{Z}_I) \quad (4.26)$$

により求める方法であるが、 $f_j$  が  $Z$  に関して線形に近くない場合には誤差が大きくなる。特に  $N(I)$  が大きくなると用いることはできない。

#### (2) 似たもの同士のグループごとに平均的な個人を用いる方法

方法(1)は  $N(I)$  が大きくなると精度が落ちるので  $I$  を  $G$  個のグループに分け、グループ  $g$  の平均的な個人  $\bar{Z}_g$  を用いて

$$P_j(I) = \sum_{g=1}^G \frac{N(I_g)}{N(I)} f_j(\bar{Z}_g) \quad (4.27)$$

$N(I_g)$ : グループ  $g$  の人数

により求める。

グループ分けでは個人属性  $Z$  の全要素を使うことは通常できず、効用に大きな影響を持つ要素を取り出してグループ分けすることが实际的である。

#### (3) $Z$ の分布の平均値まわりのモーメントを用いる方法

属性  $Z$  を持つ個人が  $j$  を選ぶ確率  $f_j(Z)$  は、 $Z$  の平均値まわりで Taylor 展開し、2次微分で打ち切ることにより

$$f_j(Z) \cong f_j(\bar{Z}) + \sum_r \frac{\partial}{\partial Z_r} f_j(Z) \Big|_{Z=\bar{Z}} (Z_r - \bar{Z}_r) + \frac{1}{2} \sum_r \sum_q \frac{\partial^2}{\partial Z_r \partial Z_q} f_j(Z) \Big|_{Z=\bar{Z}} (Z_r - \bar{Z}_r)(Z_q - \bar{Z}_q) \quad (4.28)$$

で近似できる。そこで  $Z$  の確率密度関数  $p(Z)$  とすると  $I$  で  $j$  を選ぶ人の割合は

$$P_j(I) = \int_I f_j(Z) p(Z) dZ \cong f_j(\bar{Z}_I) + \frac{1}{2} \sum_r \sum_q \frac{\partial^2}{\partial Z_r \partial Z_q} f_j(Z) \Big|_{Z=\bar{Z}} \text{cov}(Z_r, \bar{Z}_q) \quad (4.29)$$

で近似できる。これは第2項が上記(1)の修正項として作用していることを意味するが、(1)よりも精度向上をもたらす保証はなく、第2項内の共分散の推定に必要なデータが充分でない場合には用いることはできない。

#### (4) $Z$ の分布の近似を用いる方法

$Z$  の確率密度関数  $P(Z)$  の近似として既知の分布  $\tilde{P}(Z)$  を用いる方法があり、式(4.18)の  $\varepsilon_{ij}$  が正規分布すなわち個人  $i$  が  $j$  を選ぶ確率は Probit モデルで表現できる場合で  $\tilde{P}(Z)$  も正規分布とできるときには  $P_j(I)$  は求めることができ[13]、特に対象が  $i$  と  $j$  の2個しかない場合には

$$P_j = \Phi \left( \frac{\beta'(\bar{Z}_j - \bar{Z}_i)}{\sqrt{1 + \sigma^2}} \right)$$

$$\sigma^2 = \beta' V \beta$$

$V$ :  $Z$  の共分散行列

により求めることができる。しかし  $\varepsilon_{ij}$  が Gumbel 分布すなわちロジットモデルの場合には対象が2個の場合以外は計算が困難なようである。([1], P145)

#### (5) 標本の直接利用

標本が無作為抽出されておれば

$$\hat{P}_j(I) = \frac{1}{N(I)} \sum_{i=1}^{N(I)} P_{ij}$$

により  $P_j(I)$  を推定することができる。  $I$  が (2) に示された  $G$  個のグループに分けられるときには

$$\hat{P}_j(I) = \sum_{g=1}^G \frac{N(I_g)}{N(I)} \frac{1}{N(I_g)} \sum_{i=1}^{N(I_g)} P_{ij}$$

$N(I_g)$ : グループ  $g$  の標本サイズ

として層化抽出による精度向上を期待できる。

以上、5つの方法を紹介したが、筆者としては実用性、精度の観点から層化抽出（事後層別でもよい）による(5)の方法を推奨したい。なお、節4.3.2では所得でのクラス分けの場合の  $\{k=11\}$  に相当する世帯を除去したが、集計ではその世帯が  $j=2$  を選択する確率を零として集計しなければならない。

コンジョイント分析の場合には節3.2.2（前回）の数量化理論におけるエリア集計と同じ方法をとれるが、需要数にするためには別途、工夫しなければならない。

## あとがき

4回にわたって予測手法を紹介したが、筆者が使用したことのある手法に限定したため、重要な予測手法が欠落しているであろう。また各手法の専門家から見ると随分もの足りないものになっていると思う。とりあげなかったがよく知られた手法としては

計量経済学（同時方程式）モデル  
システム・ダイナミックス  
産業連環（投入算出）分析  
アルファイ法

などが上げられる。

長期予測と新サービス需要予測が予測困難な2大課題であり、特に長期予測については環境変化をどこまで考えるかが問題である。たとえばバブル崩壊などは果たして予測で考慮できたであろうかという点に関しては、長期の景気循環を考慮していれば上げ一片倒ということはあり得ないので予測できたはずだと言えるかも知れない。しかしそのためには長期景気循環を表わす長期データの存在が前提になるであろうし、データがなければ外部機関のマクロ予測データをうまく利用することを考えなければならない。それは他人に説明できる高度な格好いい手法といったものではなく、直観を働かせた泥くさい手法であろう。

予測でもう一つ考慮しなければならないのは過剰適合の問題である。予測手法の精度を論議するときには過去のデータへの適合のよさが一つの尺度となりうる。しかし、回帰分析の場合から類推できるように説明変

数をふやしたり、モデルを複雑にすれば適合度は上がるが、予測力は一般に下がるものである。そこで一方では簡明な予測モデルを求めつつ、他方では将来起こりうる事態をどうモデルに組み込んでおくかを考えなければならない。

このように予測手法には過度の期待はできないが、意思決定の基礎データとして予測は欠かせず、予測はずれを意思決定の中にどう反映していくかが最も重要である。

## 参考文献

- [1] M. Ben-Akiva and S.R. Lerman: "Discrete Choice Analysis", MIT Press, 1985.
- [2] 井上, 山本: 「通信サービスにおけるユーザの選択行動分析法」, 電子情報通信学会誌, Vol.76, No.5, pp.510-517, 1993.
- [3] R.D. Luce and J.W. Tukey: "Simultaneous conjoint measurement: A new type of fundamental measurement", J. Math. Psychology, 1, pp.1-27, 1964.
- [4] J.B. Kruskal: "Analysis factorial experiments by estimating monotone transformations of the data", J. Royal Statistical Society, Series B-27, pp.251-263, 1965.
- [5] R.M. Jonson R. M.: "Trade-off analysis of consumer values", J. Marketing Research, 11, May, pp.121-127(1974).
- [6] R.M. Johnson: "A simple method for pairwise monotone regression", Psychometrika, Vol.40, No.2, pp.163-168, 1975.
- [7] V. Srinivasan and A.D. Shocker: "Linear programming techniques for multidimensional analysis", Psychometrika, Vol.38, No.3, pp.337-369, 1973.
- [8] K. Ogawa: "An approach to simultaneous estimation and segmentation in conjoint analysis", Marketing Science, Vol.6, No.1, pp.66-81, 1987.
- [9] 林 知己夫, 鮑戸 弘: "多次元尺度解析法", p.101, サイエンス社, 1976.
- [10] 中西正雄, 阿部周造, 池尾恭一, 片平秀貴, 小島健司: "消費者行動分析のニュー・フロンティア", pp.177-179, 誠文堂新光社(1984).
- [11] 林恒一郎: 「交通需要予測における非集計モデルによるアプローチ」, オペレーションズ・リサーチ, Vol.30, No.3, pp.191-196, 1985.
- [12] 片平秀貴, 杉田善弘: 「マーケティング, サイエンスの最近の動向: 米国を中心として」, オペレーションズ・リサーチ, Vol.39, No.4, pp.178-188, 1994.
- [13] C. Daganzo: "Multinomial Probit: The Theory and Its Applications to Demand Forecasting", Academic Press, New York, 1979.