

# 共変量情報を用いた偏りのある 標本抽出調査の補正の可能性の検討

星 野 崇 宏  
東 京 大 学  
教養学部/総合文化研究科専任講師

## 1. 問題意識と目的

近年、様々な分野で迅速かつ正確な意思決定が必要とされてきている。正しい意思決定を行うためには、以前にも増して情報を迅速かつ正確に収集することが様々な分野において求められるようになった。しかしこれに並行して皮肉なことに、統計学的に優れているとされる確率抽出法を用いた調査研究の実施はますます難しくなってきた。これまで日本においては住民台帳や選挙人名簿などの利用が比較的簡単であったが、近年、住民台帳・選挙人名簿の閲覧を制限する地方自治体が増えたことで、学術機関や民間による確率抽出標本を用いた調査が難しくなってきた。また近年では、市場のニーズを正しく知るために、市場調査を頻繁に行うことが必要になっている。そこで非常に時間かつコストのかかる訪問調査に代わり、郵送調査、RDD法などの電話調査法、さらにはインターネット調査などが主流を占め、さらに現在では政府関係の調査においても利用されるようになってきている。しかし、多くの研究で、従来型の厳密な確率的標本抽出に基づく既存調査（訪問調査や郵送調査法）とRDD電話調査やインターネット調査の結果の比較から、後者は標本の代表性に欠け、バイアスを持つことが示されている（Couper, 2000；大隅, 2002；鈴木, 2003；前田・土屋, 2001；山岡・林, 1999）

本研究では上記の問題意識から、これまで準実験での因果効果の推定に利用されてきた、傾向スコア（Rosenbaum & Rubin, 1983；星野・繁樹, 2004；Hoshino, Kurata & Shigemasu, in press；狩野, 2002）を用いた共変量調整法を調査の偏りの補正に利用するための統計的な手法を提案する。さらに、市場調査および社会調査を目的として行った、偏りのある抽出方法によって得られた調査データから、より偏りの少ない抽出による調査での結果を予測する方法に関する

検証実験の報告を行う。

## 2. モデル設定

本研究の目的は傾向スコアを用いて有意抽出による調査から抽出の偏りの少ない確率抽出標本調査の結果を予測する方法の開発と具体的な市場調査への応用である。傾向スコアは、RosenbaumとRubinが1983年に提案した概念であり、近年医学や経済学などの無作為割り当てを伴う実験研究が難しいとされてきた分野において、無作為割り当てを行って得られる結果を近似するための手法として活発に利用されている（たとえば星野・繁樹(2004)を参照）。この方法を異なる調査間の共変量調整に利用するためには、ある調査協力者が2つの調査のどちらに協力するかということ、無作為でない割り当ての問題として考えればよい。さらに、無作為でない割り当てを欠測データの問題(Little & Rubin, 1976)として考えると非常に分かりやすい。

ここでは割り当てに2つの条件があるとし、条件1の時の従属変数ベクトルの値を  $y_1$ 、同様に条件2の時の従属変数ベクトルの値を  $y_2$  とする。理論上は、各被験者（ユニット）は2つの従属変数ベクトル  $y_1, y_2$  を有するが、観測されるのはどちらか一方であるとする。

また、被験者が条件1の時は  $z=1$ 、条件2の時は  $z=0$  の値をとる変数  $z$ （これを割り当て変数と呼ぶ）を導入する。また第  $i$  被験者の  $z$  の値  $z_i$  が1となる確率を  $w_i$  とおく。さらに、条件1に割り当てられた被験者の集団を群1、同様に条件2に割り当てられた被験者の集団を群2とする。ここで全ての被験者において、群1の場合の従属変数  $y_1$  も群2の場合の従属変数  $y_2$  も「本来はある」が、観測されるのはどちらか一方だけであるとする（図1-1）。

このとき、割り当てを与えた時の  $y_1, y_2$  の分布を

$$\begin{aligned} \mathbf{y}_1|z=1 &\sim p(\mathbf{y}_1|\boldsymbol{\theta}_{11}, z=1), & \mathbf{y}_1|z=0 &\sim p(\mathbf{y}_1|\boldsymbol{\theta}_{12}, z=0) \\ \mathbf{y}_2|z=1 &\sim p(\mathbf{y}_2|\boldsymbol{\theta}_{21}, z=1), & \mathbf{y}_2|z=0 &\sim p(\mathbf{y}_2|\boldsymbol{\theta}_{22}, z=0) \end{aligned} \quad (1)$$

と置く。通常推定できるのは母数  $\boldsymbol{\theta}_{11}, \boldsymbol{\theta}_{22}$ （図1-1の(1)と(4)の部分のモデルの母数）だけであり、割り当てを所与とした  $y$  の周辺分布  $p(\mathbf{y}_1|\boldsymbol{\theta}_{12}, z=0)$ （(2)の部分） $p(\mathbf{y}_2|\boldsymbol{\theta}_{21}, z=1)$ （(3)の部分）の母数推定は通常は不可能である。

ここでの目的は、Rosenbaum & Rubin (1983) の Strong Ignorability の仮定

を置くこと で、一般的なパラメトリックモデルでの  $\theta_{12}$ ,  $\theta_{21}$  (図1-1の (2) と (3) の部分のモデルの母数) の推定を行うことである。ここで  $y_{i1}$  を第  $i$  被験者が第1条件に割り当てられた場合の従属変数ベクトル、同様に  $y_{i2}$  を第  $i$  被験者が第2条件に割り当てられた場合の従属変数ベクトルとし、 $x_i$  を第  $i$  被験者の共変量ベクトルとする。また、合計の被験者数を  $N$  とする。そして

$$p(\mathbf{y}_1, \mathbf{y}_2, z, \mathbf{x}), \quad (2)$$

を従属変数と割り当て変数、共変量の同時分布とする。一般的に行動科学では  $y, x$  の次元が非常に大きいので、この同時分布の真の形状を仮定することは可能ではないとする。また、本論文では、Rosenbaum & Rubin(1983)の strong ignorability の仮定 (割り当ては従属変数の値には依存せず共変量に依存するという仮定)、つまり

$$p(z|\mathbf{y}_1, \mathbf{y}_2, \mathbf{x}) = p(z|\mathbf{x}). \quad (3)$$

が成立しているとする。この仮定をベイズの定理で言い換えると

$$p(\mathbf{y}_1, \mathbf{y}_2|z, \mathbf{x}) = p(\mathbf{y}_1, \mathbf{y}_2|\mathbf{x}), \quad (4)$$

つまり共変量と従属変数の回帰モデルが  $z$  (条件) によらないと仮定する。また、共変量を所与とした時に  $y_{i1}$  が観測される確率を  $w_i = w(x_i, \alpha) = p(z_i = 1|x_i, \alpha)$  とし、 $\alpha$  をそのモデルの母数ベクトルとする。割り当て変数の共変量への回帰モデルは複合二項分布を仮定し

$$p(z_1, \dots, z_N | \mathbf{x}_1, \dots, \mathbf{x}_N, \alpha) = \prod_{i=1}^N w(\mathbf{x}_i, \alpha)^{z_i} (1 - w(\mathbf{x}_i, \alpha))^{1-z_i} \quad (5)$$

とする。一般的にこの回帰モデルはロジスティック回帰モデルなどを仮定することが多いが、プロピットモデルなども利用できる。

### 3 傾向スコアを重みとするM推定量の提案とその性質

$\theta_{12}$  を推定するための目的関数を考える。もし欠測が無い場合の目的関数を

$$Q_N^{comp}(\mathbf{y}, \mathbf{x}, \mathbf{z} | \theta_{12}, \alpha, z = 0) = \frac{1}{N} \sum_{i=1}^N m(\mathbf{y}_{i1} | \theta_{12}, z_i = 0) \quad (6)$$

とする。

欠測を考慮した目的関数を

$$Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z} | \boldsymbol{\theta}_{12}, \boldsymbol{\alpha}, z = 0) = \frac{1}{N} \sum_{i=1}^N \frac{z_i}{w(\mathbf{x}_i, \boldsymbol{\alpha})} \frac{1 - w(\mathbf{x}_i, \boldsymbol{\alpha})}{p(z_i = 0)} m(\mathbf{y}_{i1} | \boldsymbol{\theta}_{12}, z_i = 0), \quad (7)$$

とする。

一般的に  $\alpha$  の真値  $\alpha^0$  は未知であるので、その割り当て変数の共変量への回帰モデル (式 (5)) での最尤推定量  $\hat{\alpha}$  を利用するとき、以下の定理が成立する。

**Theorem 1.**  $Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z} | \boldsymbol{\theta}_{12}, \hat{\alpha}, z = 0)$  について  $\frac{\partial}{\partial \boldsymbol{\theta}_{12}} Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z} | \boldsymbol{\theta}_{12}, \hat{\alpha}, z = 0) = 0$  を満たす推定量  $\tilde{\boldsymbol{\theta}}_{12}(\hat{\alpha}_{12})$  は  $\boldsymbol{\theta}_{12}$  の一致推定量であり、漸近正規性

$$\sqrt{N}(\tilde{\boldsymbol{\theta}}_{12}(\hat{\alpha}) - \boldsymbol{\theta}_{12}^0) \sim N(0, A^{-1}(\boldsymbol{\theta}_{12}^0)(B(\boldsymbol{\theta}_{12}^0) - \Gamma \Sigma_{\hat{\alpha}} \Gamma^t) A^{-1}(\boldsymbol{\theta}_{12}^0)) \quad (8)$$

を有する。但し  $\Sigma_{\hat{\alpha}}$  は  $\hat{\alpha}$  の分散であり、

$$\begin{aligned} \Gamma &= E_{\mathbf{y}, \mathbf{x}, \mathbf{z}} \left( \frac{\partial^2}{\partial \boldsymbol{\theta}_{12} \partial \boldsymbol{\alpha}^t} \frac{z}{w(\mathbf{x}, \boldsymbol{\alpha})} \frac{1 - w(\mathbf{x}, \boldsymbol{\alpha})}{p(z = 0)} m(\mathbf{y}_1 | \boldsymbol{\theta}_{12}, z = 0) \right) \\ &= -E_{\mathbf{x}} \left\{ E_{\mathbf{z}} \left[ \frac{z}{w^2(\mathbf{x}, \boldsymbol{\alpha}) p(z = 0)} \frac{\partial w(\mathbf{x}, \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right] | \mathbf{x} \right\} E_{\mathbf{y}_1} \left[ \frac{\partial}{\partial \boldsymbol{\theta}_{12}} m(\mathbf{y}_1 | \boldsymbol{\theta}_{12}, z = 0) | \mathbf{x} \right] \end{aligned} \quad (9)$$

#### 4 提案された推定量を用いた一般的な検定手法

上記の重み付きM推定量における検定手法を提案する。一般性を考えて、複合仮説の検定を考える。またこの節では、記法の簡便化のため、 $\boldsymbol{\theta}_{12} = \boldsymbol{\theta}$  と置く。帰無仮説  $H_0$  が  $\boldsymbol{\theta}$  の補集合  $\boldsymbol{\theta}_0$  として定義されていて、 $\boldsymbol{\theta}_0$  は  $p - q$  個の制約

$$R_l(\boldsymbol{\theta}) = 0, \quad (l = 1 \cdots p - q), \quad (10)$$

によって与えられているとする。

ここで  $\beta$  を  $H_0$  の下での母数とし、その真値を  $\beta^0$  とし、

$$\boldsymbol{\theta}_l = g_l(\beta_1, \dots, \beta_q) \quad (l = 1 \cdots p), \quad (11)$$

とする。但し  $\boldsymbol{\theta}$  は  $\boldsymbol{\theta}_l$  の第  $l$  要素である。

また  $\tilde{\beta}(\hat{\alpha})$  を  $H_1$  の下での  $\boldsymbol{\theta}$  の重み付きM推定量とし、 $\tilde{\boldsymbol{\theta}}(\hat{\alpha})$  を  $H_1$  の下での母数  $\boldsymbol{\theta}$  の重み付きM推定量とする。

ここで検定統計量  $\Delta^C$  を

$$\Delta^c = -2\{Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z}|\hat{\beta}(\hat{\alpha}), \hat{\alpha}) - Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z}|\hat{\theta}(\hat{\alpha}), \hat{\alpha})\} + o_p(1). \quad (12)$$

とする。但し  $Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z}|\beta, \alpha)$  は  $Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z}|g(\beta), \alpha)$  のことを示す。この時、以下の定理が成立する。

Theorem2・ $H_0$ の下での  $\Delta^c$  は  $\sum_{k=1}^p \lambda_k U_k$  の分布に漸近的に等しくなる。但し  $U_k$  は独立な自由度 1 の  $\chi^2$  確率変数であり、 $\lambda_k$  は以下の行列の第  $k$  番目に大きい固有値である：

$$(B(\theta^0) - \Gamma \Sigma_{\hat{\alpha}} \Gamma^t)(A(\theta^0)^{-1} - M A(\beta^0)^{-1} M^t), \quad (13)$$

但しここで

$$A(\beta^0) = E_{\mathbf{y}, \mathbf{x}, \mathbf{z}} \left[ \frac{\partial^2}{\partial \beta \partial \beta^t} m(\mathbf{y}_1 | \beta, \mathbf{z} = 0) \right] \quad (14)$$

## 5 異なる2つの調査間の共変量調整への応用と共変量選択法の提案

### 5.1 偏りのある抽出による調査結果の補正

無作為でない割り当て状況において、傾向スコアを用いて共変量情報を一元化し、無作為割り当てを近似する手法を、異なる調査間の共変量調整に利用できる。

これは、調査協力者が2つの調査に割り当てられる状況を想定し、その割り当てが共変量に依存すると考えることで可能になる。ここで前節までの  $y_1$  を調査1に回答したときの項目の値、 $y_2$  を調査2に回答したときの項目の値、 $z$  を調査1に参加するなら  $z=1$ 、調査2に参加するなら  $z=0$  とするインディケータ変数とする。また、 $z=1$  となるユニットの集合を標本1、 $z=0$  となるユニットの集合を標本2とする。

第2節の仮定と同じく  $\mathbf{y}_j | z = 1, \mathbf{y}_j | z = 0$  ( $j = 1, 2$ ) の分布を式(1)のように置く。ここで例えば  $p(\mathbf{y}_1 | \theta_{12}, z = 0)$  の意味は、「調査2と同じ抽出を行った場合の項目の分布」と言える。これは言い換えると、「共変量の分布を調査2と同じに調整した場合の、調査1の項目の分布」でもある。

従って、標本1が偏りのある標本抽出で得られた標本、標本2が無作為抽出によって得られた標本であるとする、もし調査モードの影響がない、つまり  $y_1 = y_2$  であれば、 $p(\mathbf{y}_1 | \theta_{12}, z = 0)$  の母数を前節までの方法で推定することによ

って、偏りのある標本抽出で得られた結果と共変量情報から「無作為抽出による被験者集団での  $y_2$  の分布の母数」を推定することが可能となる。

## 5. 2 共変量の選択法

傾向スコアを用いた調査データの調整において重要なのは共変量の選択である。因果効果を推定するために傾向スコアを利用する場合は、観測されている共変量群を利用することで Stong Ignorability が成立するかを調べれば良い。これに対して、調査の補正のために本方法を利用する際には、星野 (2003)、星野・繁樹 (2004) が述べるように、本調査で利用する前に、どのような項目を共変量として用いるべきかについての実験調査を行う必要があると考えられる。また、共変量の選別についても、単に傾向スコア算出のモデリングの枠組みで考えることは不十分である。例えば、ロジスティック回帰分析でステップワイズ法などの変数選択などを行って共変量を決めると、概して調整がうまく行われない可能性が高い。これは、「2つの群の群別（ここではインターネット調査と訪問留置調査に参加している被験者の群別）をよりよく説明する共変量」が「調査項目をよりよく説明する共変量」であるとは限らないためである。そこで、本研究では下記の4つの基準を満たす共変量を選択する方法を提案する。

- 1 個人内変動が少なく（つまり各個人内で安定した）、かつインターネット調査と訪問調査で継続的に質問できる可能性が大きい項目を選ぶ。
- 2 訪問留置調査とインターネット調査間で差のある項目を選ぶ。（具体的には t 検定やロジスティック回帰分析を行い、標準偏回帰係数が大きくなる項目や差の検定の p 値が小さくなる項目を共変量候補として選ぶ）。
- 3 調整したい項目を共変量に回帰させた時の偏回帰係数が、二群とも同じ方向に（正なら正、負なら負に）なるものを選ぶ。特に標準偏回帰係数の絶対値が大きいものを選ぶ。
- 4 上記の基準で選択された共変量のセットから、さらに二乗誤差和を減少させる（又はもっとも増加分が小さくなる）ように共変量を減らす。

上記の基準は市場調査において提案された調整法を利用するための一種のアドホックな基準である。2段階目及び3段階目などは数理的な根拠があるとはいえ、これらの基準をどのように組み合わせるべきかは、各市場調査ごとに実験調査を行い、探索する必要があるだろう。

上記の基準を満たす項目群を共変量として、傾向スコアを算出し、前節までで提案された調整法を利用して調整を行う。ここで、共変量となる項目の数は、各調査の実施形態によって異なるが、実際の調査に利用するという実用性の観点からは20から30程度がおおよその目安と考えられる。

## 6 インターネット調査の補正の可能性の検討

### 6.1 先行研究と一般的議論

一般的に、無作為抽出による調査とインターネット調査では同じ質問項目に対して結果が大きく異なることが知られている (Couper, 2001 ; 大隅, 2002)。

その原因として、そもそも調査協力者の年齢層・男女構成・生活様式などが本来の母集団から偏っていると考えられるため、ネット調査に対して何らかの調整を行うことが必要であると考えられる。

そこで、提案された推定法をインターネット調査の補正に利用する。前節の「同じ項目を異なるモードで測定する場合」をより具体的に、調査1をインターネット調査、調査2を訪問調査とする。つまり  $y_1$  はインターネット調査群では回答が得られ ( $z=1$ )、訪問調査群では回答が得られていない ( $z=0$ )、同様に  $y_2$  は訪問調査群では回答が得られ ( $z=1$ )、インターネット調査群では回答が得られていない ( $z=0$ ) とする (図 1-2)。

また、標本1を調査会社が管理するクローズドパネル (または大隅 (2002) の言うところのリソース) から得られた標本、標本2を住民基本台帳や選挙人名簿からの無作為抽出標本とする。Strong Ignorability の仮定の下で、提案された手法を利用することで、 $\theta_{12}$  の推定、つまり「標本が無作為抽出された場合のインターネット調査の項目の分布  $p(y_1|\theta_{12}, z=0)$  の母数」の推定が可能になる。

そこで、本節では具体的なデータを用いて補正が可能かを様々な観点から検討する。今回利用したデータは、(株) ビデオリサーチが行った ACR (Audience and Consumer Report) 調査である。この調査は訪問留置調査であるが、今回は

実験調査として、訪問留置調査の項目の一部についてのインターネット調査を行っている。従って、訪問留置群  $z=0$  では  $y_2$ 、インターネット調査群  $z=1$  では  $y_1$  が得られているため、傾向スコアを用いた調整によって得られた  $E(y_1|z=0)$  の推定値が  $E(y_2|z=0)$  とどれくらい近いかを検討することができる。

今回利用した項目のうち、共変量の候補として利用した変数は、年齢などの一部を除き、名義・順序尺度水準であった。また、調整を行った項目は、「様々な金融商品の保有の有無」12項目、「よく読む新聞の記事の種類」34項目の合計46項目であり、全て2値変数である。

本研究ではこの調査データを用いて、(1) 妥当性検証のためのシミュレーション研究、(2) インターネット調査データの調整による訪問留置調査の予測、を行った。

## 6. 2 インターネット調査での平均の調整

訪問留置調査データからのリサンプリングについての結果は省略(本文参照)し、ここではインターネット調査のデータを用いた訪問留置調査の結果の予測がうまく行くかどうかを調べる。インターネット調査が東京・関西・名古屋地域に限定されていることから、今回利用した訪問留置調査の有効回答数は5371であった。

ここで、 $z=1$  をインターネット調査、 $z=0$  を訪問留置調査とする。さらに  $y$  をインターネット調査での項目への回答、 $y$  を訪問留置調査での項目への回答とする。目的は  $p(y|\theta_0, z=0)$ 、より具体的にはここでは  $z=0$  での  $y_j$  の期待値  $\pi_{j0}$  = 「訪問留置調査群でのインターネット調査の回答の期待値」を推定することである。但し、傾向スコアの推定値が極めて小さいときには、重みが極めて大きくなり、推定値が不安定になると考えられるため、以降傾向スコアの推定値が0.01以下の場合は0.01に設定した。

ここで、調整の目的変数は前小節と同じ金融商品の保有に関する12項目である。また、共変量の候補として57項目を取り上げた。

結果として選択された共変量は、「年齢」「性別」「職業」「最終学歴」「結婚年数」「世帯収入」「家の部屋数」「運転免許の有無」「家族構成(4項目)」「自動車保有台数」「ネット利用時間」「旅行(3項目)」「読書頻度」「飛行機利用」「都道府県」の20変数である。



ここで、傾向スコア算出のためのロジスティック回帰分析における正判別率は98.9%であった。

以降、 $\tilde{\pi}_{j_0}$ を訪問平均、 $\tilde{\pi}_{j_1}$ をネット平均、 $\tilde{\pi}_{j_0}^w$ を提案推定値と表記する。

各項目での訪問平均、ネット平均及び提案推定値を表1に記載した。

さらに、横軸を訪問平均、縦軸をネット平均及び提案推定値とするプロットをそれぞれ記載したものを図2、図3とした。表1や図2、図3からも分かるように、インターネット調査の平均値は訪問調査での平均と大きく異なっていることが分かる。また、図に示したように「コンピュータ記事」や「デジタルメディア記事」のようなインターネット調査の参加者にとって訪問留置群よりも明らかに関心が高いような項目を除くと、おおむね調整がうまく作用していることが分かる。

金融商品の12項目の訪問平均とネット平均の二乗誤差和は0.034781、また訪問平均と提案推定値の二乗誤差和は0.011098であった。従って、提案された手法によって二乗誤差和は68.1%減少したことになる。また、新聞記事の34項目では訪問平均とネット平均の二乗誤差和は0.277918、また訪問平均と提案推定値の二乗誤差和は0.095302であった。従って、提案された手法によって二乗誤差和は65.7%減少したことになる。

一方、共変量の探索を行わず、一般的な属性変数（性別・年齢・職業・最終学歴・都道府県）だけを用いて傾向スコアを算出した場合の金融商品の12項目の訪問平均と提案推定値の二乗誤差和は0.025597、新聞記事の34項目では訪問平均と提案推定値の二乗誤差和は0.250303であり、46変数での二乗誤差の減少は12%に過ぎなかった。

### 6.3 リサンプリングによる一貫性の検討

さらに妥当性を確かめるため、2つの調査からそれぞれ500人づつを無作為抽出し、補正を行うという処理を10000回行った。リサンプリングデータによる平均値の分布や、検定における棄却比率などからも、単純なネット調査で得られた推定値は、訪問調査の推定値から標本誤差を遥かに上回って異なっているのに対して（平均の差の有意水準5%検定の全項目での平均棄却率は30.7%）、一部の項目を除いて提案手法での推定値と訪問調査の推定値の差は標本誤差の程度に収まっている（平均の差の有意水準5%検定の全項目での平均棄却率は

8.7%)。このように、提案された手法が安定した補正を可能にしていることがわかる。

#### 6. 4 潜在クラス分析での調整

平均の補正だけでなく、一般的なパラメトリックモデルにおいても補正が可能であれば、市場調査において非常に有用であると考えられる。そこで、マーケティングリサーチにおいて重要な「マーケットセグメンテーション」のために頻繁に利用される統計手法の一つである、潜在クラス分析についての補正が可能かどうかを検討した。詳細は本文を参照されたい。結果として、所有率のプロファイルからは、潜在クラス自体は調査によってあまり変わらないが、3つのクラスの比率は大きく変わっていて、クラス1は訪問留置調査では48%であるが、インターネット調査では33%に減少し、逆にクラス2は訪問留置調査では30%ほどであるが、インターネット調査では51%に増大している。これに対して、提案手法を利用した調整結果では、比率、プロファイルともに訪問留置調査により近い結果を与えていることがわかった。この結果からも、単なる平均についての補正だけでなく、一般的な統計モデルにおいても本研究で提案された手法が利用できることが示唆される。

### 7 市場調査での再現性の検証

(株) ビデオリサーチが定期的に行っているCMに関する訪問留置調査と同時期に、同じ内容のインターネット調査を行うという実験調査を2回行うことで、提案手法の再現性を検証した。第一回調査のおよそ4カ月後に第二回の調査を行った。調整の目的とした項目は、様々なジャンルのCMの認知を尋ねる項目であり、第一回調査、第二回調査とも10のCMについて、「見た」「見たような気がする」「知らない」の3件法で答えさせている。これまでこの市場調査をCMの評価に利用する際は、「見た」と「見たような気がする」への回答率を合わせた比率をそのCMへの認知率としており、関心はこれまでの調査で得られた各CMの認知率との比較である。今回はこの変数についてインターネット調査での回答の訪問留置調査への補正がどれくらい可能かどうかを検討した。まず第一回調査において、提案された共変量選択法を利用し、共変量として以下の23項目を得た：

年齢、性別、結婚、家事について、家族人数、ある世代の子供の有無（3項目）、職業、テレビを見る時間帯（6項目）、1日あたりテレビ視聴時間、利用交通機関（2項目）、家電新製品類所有有無、携帯メール使用頻度、3ヶ月以内の飲酒（3項目）

そして、第一回調査、第二回調査でのインターネット調査のCM認知率を補正した。

第一回調査での10のCMの認知率について、訪問留置調査、インターネット調査、及び提案手法による調整後の推定値を表2に記載した。すべてのCMで、ネット調査と訪問調査の差は大きく、単純にインターネット調査を利用するだけではこれまでの訪問留置調査で得られた結果と比較することは難しいことがわかる。これに対して、提案手法を用いた調整によって、すべてのCMについて補正が成功していることがわかる。また図4は横軸に訪問留置での認知率、縦軸にインターネット調査及び補正後の認知率をプロットしたものである。図4からもわかるように、訪問留置調査において認知度の高いCMはネット調査においても高く、差があまりないために調整もあまり効果がない。しかし、訪問留置調査の認知率が低い（0.6以下の）項目ではインターネット調査との差が非常に大きく、また調整によって「ネット調査＝訪問調査」の直線に近づくことがわかる。

ここで、第一回調査から得られた共変量の組み合わせは、第一回調査においては調整を可能にする（二乗誤差は58.45%減少）のは当然であるといえることができる。しかし、第二回調査（しかも異なるCMに対して）においても、これらの共変量が有用であった（二乗誤差は42.31%減少）ということは、提案された調整法と、共変量選択法の効果が再現性を有するということを示している。

## 8 JGSSデータを用いた社会調査での傾向スコア調整法の妥当性の検討

提案された補正法の社会調査データへの利用の妥当性と補正結果の再現性を確かめるために、JGSS2001、2002及び2003調査データを利用した。具体的にはJGSS調査の一部の項目についてインターネット調査を実施し、インターネット調査から各JGSS調査の各項目の平均を傾向スコアを用いた補正法により予測し、補正前のインターネット調査での平均よりもどれくらいJGSS調査での平

均に近いかを二乗誤差の観点から検討した。スペースの都合上、この部分の結果については割愛する（本文を参照いただきたい）。結果として、インターネット調査を用いて2001年のJGSSデータの結果をよりよく予測するように共変量を選択し、2002年と2003年のデータに対しては選択された共変量をそのまま利用して傾向スコアを算出し、補正を行ったが、その結果はおおむね良好であった。このことは、本補正法が再現性を有することと、共変量の選択法が妥当性を有することを示している。インターネット調査は2001年～2003年のJGSS本調査とはそれぞれ3年から1年の時期的なズレがあるが、そのような時期的なズレに対しても本方法が頑健であることを示している。さらに妥当性を検証するため、各年度の調査ごとに2つの調査からそれぞれ500人づつを無作為抽出し、得られたサンプルサイズ1000の標本からロジスティック回帰の母数の推定と傾向スコアの計算直行い、補正を行うという処理を1000回行った結果も、本方法の妥当性を示すものであった。

## 9 2つのインターネット調査間の調整への提案手法適用

これまでの節では提案手法を「偏った標本抽出によるインターネット調査を、無作為抽出に基づく訪問調査へ近似する手法」として利用したが、提案手法はこのような目的だけではなく、2つの異なった標本からの調査を、もう一方の調査に向けて調整する手法として利用することができる。そこで、本節では2つの調査会社に実施を委託したインターネット調査間の調整を行った。標本抽出を行うモニター集団は各社ごとに異なることがこれまでも言われており、モニターの差異に伴う変動を除去すること自体もインターネット調査を用いた研究では今後重要な問題になると考えられる。（詳細は割愛。本文参照）。ここではB社の調査結果をA社の調査結果に近づけるように補正を行ったが、提案手法による補正值を用いた場合の二乗誤差は補正せずにB社の結果を利用する場合より全体として3分の2に減少しており、B社の結果をそのまま利用するよりもよいことを示している。

## 10 議論

本論文では無作為割り当てがなされていない状況での、観測されていない群での周辺分布の母数推定と検定を行う方法を提案した。さらに、この方法を、

異なる抽出法による標本間の共変量調整や比較に応用する枠組みを提示し、市場調査のためのインターネット調査の無作為抽出標本への補正の可能性を検討した。また、補正に利用するための共変量の選択法を提案した。実際の市場調査データから、インターネット調査と既存型の訪問調査の結果が非常に異なることを示した。また今回利用したデータと大規模なリサンプリングによるシミュレーション研究からは、提案手法を慎重に用いれば、市場調査において必要とされる程度のレベルでは、インターネット調査の補正がある程度可能であることが示唆された。より具体的には金融商品保有に関する項目において共変量として有用な項目を選択して傾向スコアを算出し、調整を行った結果、新聞記事に関する項目というそれほど関連が無い項目についても調整がうまく行えたことを報告した。この結果は提案手法の内的な妥当性を示したものであるが、外的妥当性、具体的にはこの手法の再現性を確認することで、さらなる妥当性の検証が可能となる。そこで、2回の調査を行い、第一回調査で選択され、調整に有効とされた共変量を用いて、第二回調査においても調整を行った結果、再現性を示唆する良好な結果が得られた。

また、提案された補正法を有効に用いるためには、傾向スコアを利用した補正法の利用の前提となる Strong Ignorability の仮定を満たすように、提案された共変量選択法を利用する必要があるということが示唆された。

また、Strong Ignorability の定義からも分かるように、目的項目によって適する共変量が異なる可能性が大きい。そこで本研究では簡便な共変量選択手法を提案し、実験的検討を行えば、目的項目が同様のものであれば同じ共変量を用いても調整がうまく行くことが示唆された。今後、今回利用した定型的な市場調査以外の様々なタイプの市場調査に利用するためには、実験調査を通じて調査項目ごとに共変量を最適化することが必要とされよう。具体的にはブランドのイメージ調査なら、各ブランドごとではあまりに汎用性がないので、少なくともブランドのグループごとに共通して利用できる共変量の探索などは求められるだろう。

社会調査データに対しても、今回利用した調査データからは、提案手法を慎重に用いれば、ある程度のレベルには、インターネット調査の補正がある程度可能であることが示唆された。より具体的には JGSS の過去3回の調査データに対して、提案された補正法を用いてインターネット調査を補正した結果、それ

ぞれ単純にインターネット調査を用いた場合に比較して誤差を半分程度に減少させることを示した。この結果は提案手法が3年間の調査時期のズレに対しても再現性を有することを示している。

本論文のJGSS調査データの応用においては、各項目に対して最適な共変量を探索するのではなく、「ロバストな」共変量を探索して利用することで、すべての項目に対して補正が作用することを報告した。しかし今後、今回利用した以外の様々なタイプの社会調査に対して今回の補正法を利用するためには、実験調査を通じて調査項目ごとに共変量を最適化することが必要とされよう。

マーケティングリサーチにおいては、「意思決定に役立つリサーチが優れたリサーチ」(古川・守口・阿部, 2003)であり、たとえ正確であってもコストやスピードの面で実行が容易ではない調査方法より、正確さについてはロスがあってもプラクティカルな方法を改善していくことは重要であると考えられる。その点で、本方法はインターネット調査という、市場調査にとって極めて有望な手法において、その問題点である標本の偏りを調整するという改善を行っているという点で有意義であると考えられる。

本方法を実際に利用してインターネット調査の補正を継続的に行っていく際には、星野・繁樹(2004)が述べるように、継続的にバイアスのより少ない調査(例えば台帳からの無作為抽出による訪問調査)を繰り返し行い、補正が正しく働くことを確かめ、もし補正がうまく行かなくなってきたら利用する共変量を変えていき、さらによりよいものにしていく、といった継続的努力が必要であることも注意点として述べ添えておきたい。

本研究の応用例で取り上げられた調査研究での補正に関する問題意識は、今後社会的にもますます重要となっていくであろう。様々な試行錯誤的な実験研究が望まれる。

## 参考文献

- Berrens, P., Bohara, A. K., Jenkins-Smith, H., Silva, C. & Weimer, D. L. (2003) The Ad-Vent Of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples. *Political Analysis*, 11, 1-22.
- Couper, M. P. (2000). Web surveys: are View of issues and approaches. *Public*

- Opinion Quarterly*, 64, 464-494.
- Deming, W. E., & Stephan, F. F. (1940). On a Least Square Adjustment of a Sample Frequency When the Expected Marginal Totals Are Known. *The Annals of Mathematical Statistics*, 11, 427-444.
- Duncan, G. M. (1987). A simplified approach to M-estimation with application to two-stage estimators. *Journal of Econometrics*, 34, 373-389.
- 星野崇宏 (2003). 調査データに対する傾向スコアの適用 品質, 33 (3), 44-51.
- 星野崇宏・繁桝算男 (2004). 傾向スコア解析法による因果効果の推定と調査データの調整について, 行動計量学, 31, 43-61.
- Hoshino, T., Kurata, H., & Shigemasu, K. (in press). A Propensity Score Adjustment for Multiple Group Structural Equation Modeling. *Psychometrika*.
- Huber, P. J. (1967). The Behavior of the Maximum Likelihood Estimates Under Nonstandard Conditions. *Proceedings the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 221-233.
- Huber, P. J. (1981). *Robust Statistics*. New York: John Wiley.
- 古川一郎・守口 剛・阿部誠 (2003). マーケティング・サイエンス入門 有斐閣.
- 狩野裕 (2002). 構造方程式モデリング, 因果推論、そして非正規性. 多変量解析の展開 (甘利俊一・狩野裕・佐藤俊哉・松山裕・竹内啓・石黒真木夫編) 64-130 岩波書店.
- Little, R. J. A., & Rubin, D. B. (1987) *Statistical Analysis with Missing Data*. Wiley, New York.
- 前田忠彦・土屋隆裕 (2001). 日本人の国民性 2000 年度吟味調査報告 統計数理研究所研究レポート, 87.
- 大隅昇 (2002). インターネット調査. 社会調査ハンドブック (林知己夫編) 200-240 朝倉書店.
- Rosenbaum, P. R. (1987). Model-Based Direct Adjustment. *Journal of the American Statistical Association*, 82, 387-394.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70,

41-55.

Rubin, D.B. (1976). Inference and, missing data. *Biometrika*, 63, 581-590.

Rubin, D.B. (1985). The Use of Propensity Scores in Applied Bayesian Inference. In J・M. Bernardo, M.H.DeGroot, D.V.Lindley, &, A.F.M.Smith (Eds.), *Bayesian. Statistics 2* (pp.463-472). North-Holland :Eisevier Science Publisher B.V.

鈴木達三 (2003). 電話調査法についての一考察 行動計量学, 30, 73-92.

Thylor, H. (2000). Does internet research work? *International Journal of Market Research*, 42, 51-63.

Taylor, H., Bremer, J., Overmeyer, C., Siegel, J.W. & Terhanian, G. (2001). The Record of Internet-based Opinion Polls in Predicting the Results of 72 Races in the November 2000 U.S. Elections, *International Journal of Market Research*, 43, 127-136.

土屋隆裕 (2005). 調査不能者の特性に関する一研究:「日本人の国民性第11次全国調査」への協力理由に関する事後調査から, 統計数理, 53, 35-56.

山岡和枝・林知己夫 (1999). 電話帳記載・非記載者をめぐる諸問題:首都圏調査から 行動計量学, 26, 114-124.