

傾向スコアを用いた補正法の有意抽出による標本調査への応用と共変量の選択法の提案

星野 崇宏¹・前田 忠彦²

(受付 2006年1月10日;改訂 2006年3月24日)

要 旨

近年、個人の意識や思考が多様化し、かつ情報化が進んできたために非常に短期間に個人の考え方が変化するという現象が生じている。従って社会科学分野全般において、頻繁に調査研究を行う必要が生じてきたため、コストのかかる無作為抽出標本による既存型調査に代わり、標本の代表性を担保しつつ短期間に費用を抑えて実施できる方法として、インターネット調査などの新しい調査法が利用されるようになってきた。しかし、有意抽出標本からの調査であることから、標本の代表性がなく、信頼性が疑問視されている。本研究では、有意抽出に基づく調査を無作為抽出標本での結果に近似するための傾向スコアを用いた補正法を利用する際に重要となる、共変量の選択法を提案した。さらに日本版一般社会調査のデータと、筆者らが行ったインターネットパネルに対する実験調査データを用いて、その妥当性の検討を行ったところ、補正の再現性が得られた。

キーワード：傾向スコア，社会調査，有意抽出，インターネット調査，共変量調整。

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、正しい評価が必要であるため、以前にも増して情報を早く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。地方政治、国政においても市民の意識を正確につかみとり、ニーズにあった政策を立案し、例えばマニフェストなどといった形で公約し、実行するといった機運が生じつつある。また、個人の意識や思考が多様化し、かつ情報化が極度に進んできたために非常に短期間に個人の考え方が変化するという現象が生じている。従って社会や世論の動向を知るためには、今後益々頻繁に社会調査・世論調査を行っていく必要があると思われる。

そこで、調査を頻繁に行うことが必要になった現在において、時間とコストのかかる確率抽出標本に対する訪問調査(又は訪問留置調査)に代わり、標本の代表性を担保しつつ短期間に費用を抑えて実施できる方法が必要とされてきているが、これまで日本においては選挙人名簿などの利用が比較的容易であったため、学術分野では既存の訪問面接調査の代替方法があまり模索されてこなかった。これに対して民間の調査では、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や、コストの削減、調査スピードの重視のために、RDD法による電

¹ 東京大学 教養学部・総合文化研究科：〒153-8902 東京都目黒区駒場 3-8-1

² 統計数理研究所：〒106-8569 東京都港区南麻布 4-6-7

話調査法、さらには近年のIT技術の進歩とインターネットの普及によってインターネット調査が広く用いられるようになった。さらに現在では市場調査だけでなく、政府関係の調査(例えば内閣府経済諮問委員会「日本21世紀ビジョン調査」「公共料金分野における情報公開の推進調査」など)においても利用されるようになってきた。

本稿で考察するインターネット調査は、大隈(2002)の分類でいう「リソースタイプ」のインターネット調査と呼ばれるもの一般と想定しておく。これは調査主体が何らかの手段で調査協力意志のある個人を集めて登録し(リソースの構築)、登録者集団内で対象者を選定し、Web上に置かれた調査票への回答を求める方法である。調査協力者に対して何らかの形での謝礼が提供されるケースが多い。本稿でデータを扱った調査も具体的にはこの方式によるものである。

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を示す研究は多い。例えば、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; 前田・土屋, 2001; 山岡・林, 1999)。

しかし、現在の調査環境を考えると、このままの形で既存の調査方法を維持していくことは難しい可能性がある。従って、確率抽出標本に基づく既存型調査は、精度を向上させるために慎重に少数回行い、その調査データを用いて、各有意抽出標本を補正するという混合モード調査の考え方による調査手法を研究することは有用であると考えられる。そのためにも、インターネット調査などの有意抽出標本による結果を既存調査の結果に補正する方法論の開発が必要であると思われる。

このような社会的なニーズに統計科学の観点から答えるために、本研究では、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 and Rubin, 1983; 星野, 2003; 星野・繁樹, 2004; 土屋, 2005)を用いた共変量調整法を応用し、関心のある調査項目以外にあえて共変量情報として様々な項目を調査に加え、その共変量情報を利用して、偏りのある抽出による標本を無作為抽出標本に補正をする方法を実際の社会調査データに適用し、その可能性と限界点について検討する。具体的には層化二段無作為抽出を用いた訪問留置調査である日本版一般社会調査(日本版 General Social Surveys: 大阪商業大学比較地域研究所・東京大学社会科学研究所, 2000-2003, 以降 JGSS と略記する〔第7節参照〕)に対して、同様のインターネット調査を行い、インターネット調査の結果からそれぞれの本調査の結果を予測することができるかどうかを検討する。

また、これまでの傾向スコアを用いた調査の補正の応用例(例えば Taylor, 2000; Berrens et al., 2003)では、調整に用いた共変量についての情報が明確でなく、どのような基準で共変量を選択したかについての議論がなされていなかった。しかし、どのような共変量を利用したかということは補正結果に非常に影響を与える。そこで、本研究では、傾向スコアを計算する際に利用する共変量の選択法を提案する。

本論文の構成は以下の通りである。第2節において割り当てが無作為でない準実験での共変量調整の問題としてモデルを定式化し、さらに「異なる2つの調査間での共変量調整」の枠組みとして捉えなおす。第3節では星野(2005)、Hoshino et al.(2005)による、因果効果推定のための傾向スコアによる重み付き対数尤度最大化の考え方による推定手法と検定手法を、第2節の枠組みに従って調査データの補正法として捉えなおす。また、一般に傾向スコアを用いた補正法においては補正に用いる共変量の選択が非常に重要であるが、これまでの研究ではその選択方法についての指針が得られていなかった。そこで、本研究では傾向スコア解析法における簡便な共変量選択法を提案する(第4節)。第5節では提案された傾向スコアによる補正法の再現性を、大阪商業大学及び東京大学社会科学研究所が行った JGSS 調査データと、筆者が行っ

たインターネット調査を用いて例証する。

2. モデル設定

傾向スコアを用いた解析法は、準実験における因果効果の推定のための統計手法として開発された方法であるが、本質的には欠測のある場合の周辺平均を推定することを目的としたものである。従って星野・繁樹(2004)が述べるように、傾向スコアとして共変量情報を一次元化して利用することで、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 y を調査1に回答したときの項目の値、 z を調査1に参加するならば $z=1$ 、調査2に参加するならば $z=0$ 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。従って、標本1の調査協力者で y は観測され、標本2の調査協力者で y は欠測しているとする。また第 i 調査協力者の z の値 z_i が $z_i=1$ となる確率を w_i とおく。このとき、調査への割り当て($=z$)を与えた時の y の分布を

$$(2.1) \quad y|z=1 \sim p(y|\theta_1, z=1), \quad y|z=0 \sim p(y|\theta_0, z=0)$$

と置く。ここで例えば $p(y|\theta_0, z=0)$ は「調査2と同じ抽出を行った場合の項目への反応の分布」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、項目への反応の分布」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、 $p(y|\theta_0, z=0)$ の母数を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「無作為抽出による調査協力者集団での y の分布」を得ることが可能となる。しかし通常推定できるのは母数 θ_1 だけであり、割り当てを所与とした y の周辺分布 $p(y|\theta_0, z=0)$ の母数推定は通常は不可能である。しかしもし θ_0 の推定を行えば、有意抽出による偏り(有意抽出標本からの単純な推定値 θ_1 と、無作為抽出標本からの推定値 θ_0 の差)を補正することが可能になる。

ここで y_i を第 i 調査協力者の従属変数ベクトルとし、 x_i を第 i 調査協力者の共変量ベクトルとする。また、合計の調査協力者数を N とする。一般的に調査状況では y, x の次元が非常に大きいので、そ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 and Rubin(1983)の Strong Ignorability の仮定(割り当ては従属変数の値には直接には依存せず、共変量に依存するという仮定)、つまり

$$(2.2) \quad p(z|y, x) = p(z|x)$$

が成立しているとする。また、この仮定は共変量はすべて観測されているということも示唆している。この仮定の一般的なチェック方法については、星野・繁樹(2004)を参照されたい。また調査の補正においてこの仮定をどのように満たすかについては、本論文の第4節において述べる。また、共変量を所与とした時に y_i が観測される確率を傾向スコアと呼ぶ。本論文では Rosenbaum and Rubin(1983)の定義に従い、傾向スコアは共変量に依存する ($w(x_i, \alpha) = p(z_i=1|x_i, \alpha)$) とし、 α 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

$$(2.3) \quad p(z_1, \dots, z_N | x_1, \dots, x_N, \alpha) = \prod_{i=1}^N w(x_i, \alpha)^{z_i} (1 - w(x_i, \alpha))^{1-z_i}$$

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、プロビットモデルなども利用できる。このとき、当然ながら、 θ_1 は観測された y だけから最尤推定を行っても一致性があるが、 θ_0 についてはそれは保証できない。もし Strong Ignorability の

仮定が成立すれば, Missing At Random (Rubin, 1976) の仮定が成立する. 従って, y の x への回帰モデル $p(y|x)$ のモデル仮定が可能なら, 欠測を無視して回帰モデルの母数を推定することで式 (2.1) の母数も推測できる. しかし, 社会科学における調査では一般に x, y ともに高次元であるので, 回帰モデルの仮定が出来ない場合が多い. 本研究では, この回帰関数の形状も未知であり, 式 (2.1) の y の分布のみ既知であるとする. 従って欠測を無視することはできない.

3. 傾向スコアを重みとする対数尤度最大化推定量について

欠測を考慮した目的関数を

$$(3.1) \quad Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z} | \theta_0, \alpha, z=0) = \frac{1}{N} \sum_{i=1}^N \frac{z_i}{w(\mathbf{x}_i, \alpha)} \frac{1-w(\mathbf{x}_i, \alpha)}{p(z_i=0)} \log p(y_i | \theta_0, z_i=0),$$

とし, α の最尤推定量を $\hat{\alpha}$ とする時

$$(3.2) \quad \frac{\partial}{\partial \theta_0} Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z} | \theta_0, \hat{\alpha}, z=0) = 0$$

を満たす θ_0 を「 θ_0 に対する傾向スコアを用いた重み付き対数尤度最大化推定量」 $\tilde{\theta}_0$ と呼ぶ. 但し $p(z=0) = \int (1-w(\mathbf{x}, \alpha)) p(\mathbf{x}) d\mathbf{x}$ である.

このとき, $\tilde{\theta}_0$ は一貫性と漸近正規性を有する(星野, 2005; Hoshino et al., 2005 参照). ここで, 上記目的関数の $p(z_i=0)$ は個人間で共通の事前確率であるので, これを除いてもかまわない. さらに, この重み付き対数尤度最大化推定量を用いて, 尤度比検定と同様の一般的な検定を行うことが出来る(星野, 2005 参照).

3.1 二項分布への応用と比率の推定値の補正

有意抽出による調査データでの比率の推定値の補正を行うことを考える. $z=0$ のときの y の分布をベルヌーイ分布

$$(3.3) \quad y|z=0 \sim \text{Bernoulli}(p)$$

とおく. このとき $\hat{\alpha}$ を所与とするときの目的関数は

$$(3.4) \quad Q_N^W(\mathbf{y}, \mathbf{x}, \mathbf{z} | p, \hat{\alpha}, z=0) = \frac{1}{N} \sum_{i=1}^N \frac{z_i(1-w(\mathbf{x}_i, \hat{\alpha}))}{w(\mathbf{x}_i, \hat{\alpha})} \{y_i \log p + (1-y_i)(1-\log p)\}$$

となる. 従って, p の推定量は

$$(3.5) \quad \tilde{p}(\hat{\alpha}) = \frac{\sum_i^N \frac{z_i(1-w_i)}{w_i} y_i}{\sum_i^N \frac{z_i(1-w_i)}{w_i}}$$

となる(但し $w_i = w(\mathbf{x}_i, \hat{\alpha})$). これが, 共変量の分布が無作為抽出標本での分布に等しくなるように補正された, 比率の推定値である. ここで, これまでの傾向スコアを用いた調査データの補正の研究の多くでは, 重みを $z_i(1-w_i)/w_i$ でなく, z_i/w_i としてきたが, これは, $p = E(y|z=0)$ の推定値ではなく, $E(y)$ の推定値となることに注意する. 調査の調整の文脈では, 混合集団上での周辺平均 $E(y)$ を求めるのはあまり意味が無い.

また傾向スコアの推定量が既知の場合の \tilde{p} の漸近分散は

$$(3.6) \quad \frac{1}{N^2} ((p^0)^{-1} + (1-p^0)^{-1})^{-2} \sum_{i=1}^N \left\{ \frac{z_i(1-w_i)}{w_i} [y_i(p^0)^{-1} - (1-y_i)(1-p^0)^{-1}] \right\}^2$$

である(但し p^0 は比率の真値であり, 実際の計算では $\hat{p}(\hat{\alpha})$ で置き換える). 星野(2005)の議論より, これは本来の推定量の分散よりも大きい.

4. 共変量の選択法

星野(2003), 星野・繁樹(2004)が指摘するように, 傾向スコアを用いた調査データの補正において重要なのは共変量の選択である. しかし, これまでの先行研究では, 共変量をどのように選択すると, 補正がうまく働くかについての議論がなされてこなかった. しかし, どの共変量を選ぶかによって, 補正結果が大きく変わってくるのが知られている.

傾向スコアを用いた補正法が機能するためには, Strong Ignorability の仮定が(近似的にでも)成立するかどうかをチェックすることが必要となる. 調査の補正において傾向スコアによる補正法を利用するには, 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.

共変量の選別を, 調査への割り当てと共変量の関係だけ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, 概して補正がうまく行われな可能性が高い(具体的には第5節参照). これは「2つの群の群別(ここでは有意抽出による調査(例えばインターネット調査)と確率抽出に基づく調査(例えば訪問調査)に参加している調査協力者の群別)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. また, このように共変量を選んだとしても, 式(2.2)の基準, つまり Strong Ignorability の仮定が満たされるかはわからない.

共変量選択のもっとも単純かつ明快な基準は, 補正したい項目すべてに対して「 $z=0$ の群での y の期待値と本方法による推定値との二乗誤差」の和を最小にする共変量のセットを見つけることであるが, 一度に共変量をすべて選択しようとする組み合わせが非常に多くなり, 現実的ではない. 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(ここでは以降, 具体的には有意抽出標本による調査としてインターネット調査を, 確率抽出に基づく調査として訪問調査を考える).

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. 補正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)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, さらに二乗誤差の和を減少させる(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4つの基準すべてを満たす共変量を, その候補項目から選び出す. 特に3段階目については, 補正したい項目全体で回帰係数の向きが同じになることが多い変数を選択する.

上記の基準のうち, 1段階目は実施上の理由から, また2段階目は傾向スコア算出のためのロジスティック回帰において有効な共変量を選択するため, そして3段階目は共変量による目的項目への説明力が高いほど補正に有効であると考えられるためである. さらに, 3段階目で回帰係数が二群で同じ方向になる共変量を利用することで, $p(y|x, z=0) = p(y|x, z=1)$ さらにこれをベイズの定理を用いて書き直したものである, 式(2.2)の基準, つまり Strong Ignorability の仮定を満たすことができる.

上記の基準は調査において補正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段階目はそれぞれ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、具体的には例えば2段階目での p 値の基準などは、各社会調査のサンプルサイズや利用可能な共変量の候補の数、目的などに応じて異なるため、事前の実験調査を行い、探索する必要がある。

そのような事前の実験調査で上記の基準を満たす項目群を共変量として利用して傾向スコアを算出し、前節の補正法を利用して補正を行う。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、実際の調査に継続的に利用するという実用性の観点からは20程度がおおよその目安と考えられる。2段階目や3段階目の推定・検定の際の共変量としての採用の基準は、サンプルサイズによって異なることに注意する。また、Strong Ignorability(式(2.2))の条件、つまり3段階目の基準が示唆するところは、補正の目的となる項目によって適切な共変量の組み合わせは変わりうるということである。但し、同じ形式の同様の項目を繰り返し補正する場合は、当然利用する共変量は共通でもあまり補正の効果は変わらないということが予測される。これについては第5節にてJGSS調査データを用いて検討する。

5. JGSS データを用いた傾向スコア調整法の例示

5.1 利用した調査データ

本節では、上記の補正法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。

インターネット調査：2004年7月末から8月初旬に掛けて(株)日経リサーチのネット調査パネルから1万人を無作為抽出してメールで調査依頼を行い、4199の回答が得られた。

JGSS調査：計画標本サイズおよび回収率は2001年調査では4500, 62.4%, 2002年調査では5000, 62.3%, 2003年調査では留置票が2種類あり、今回利用したA票では3500, 55.0%であった。2002年と2003年の調査では、喫煙の項目の選択肢が一部2001年と異なるが、選択肢を統合して2001年と同じ形式に直した。また同様に同居人数に関しても適宜変換を行った。

また、今回インターネット調査で実施された、補正の対象となる項目は

- (1) 一般的対人的信頼感の項目：一般的に、人は信用できると思いますか(はい・いいえ・場合による、の3件法)
- (2) 「夫に十分な収入がある場合には、妻は仕事をもたない方がよい」という意見に、あなたは賛成ですか、反対ですか(賛成・どちらかといえば賛成・どちらかといえば反対・反対、の4件法)
- (3) 「妻にとっては、自分の仕事をもつよりも、夫の仕事の手助けをする方が大切である」という意見に、あなたは賛成ですか、反対ですか(賛成・どちらかといえば賛成・どちらかといえば反対・反対、の4件法)
- (4) 「政府は、裕福な家庭と貧しい家庭の収入の差を縮めるために、対策をとるべきだ」という意見に、あなたは賛成ですか、反対ですか(賛成・どちらかといえば賛成・どちらともいえない・どちらかといえば反対・反対、の5件法)
- (5) あなたの町に外国人が増えることに賛成ですか、反対ですか(賛成・反対、の2件法)
- (6) 不治の病におかされた患者が、痛みを伴わない安楽死を望んでいるとします。その家族

も同意している場合に、医者が安楽死を行える法律をつくるべきだと思いますか（はい・いいえ・わからない、の3件法）

の6項目であった。

5.2 共変量の選択

今回は JGSS2001, JGSS2002, JGSS2003 すべてで採用されている項目の一部についてのみインターネット調査を行ったので、利用可能な共変量は少ない(32項目：付録参照)。そこで今回は、前節で取り上げられた共変量選択法の妥当性を検証するという目的から、以下の3通りの方法で共変量のセットを構成し、それぞれを用いて算出された傾向スコアによって補正を行った。

方法1. インターネット調査を用いて2001年の訪問留置調査データの結果をうまく予測できるように、前節で取り上げた基準に合わないものを外すという形で共変量を除去し(具体的には32の候補のうち、共変量選択法の2段階目で25に、3段階目で18に絞った)、以下の18項目を最終的な共変量として取り上げた。性・年齢・地域ブロック(北海道・東北/関東など6地域)・都市規模・学歴・父学歴・母学歴・職業分類・世帯収入・住居形態・新聞を読む頻度・喫煙・飲酒・先週の状況・先週仕事をしていただどうか・所属階層の自己評定・同居人数・仕事でメールをするか？

方法2. インターネット調査に協力したか、2001年の訪問留置調査に協力したかという二値変数を基準変数としたロジスティック回帰分析における変数減少法から共変量を選択し(p値が0.01を閾値とし、得られた20変数からさらに寄与の小さい2変数を除去し)、上記と同じ数の18項目になるようにした。即ち性・地域ブロック・都市規模・学歴・母学歴・配偶者学歴・職業分類・世帯収入・住居形態・本を読む頻度・先週仕事をしていただどうか・15歳のころの居住地域・15歳のころの父親の職業・15歳のころの母親の職業・15歳のころの世帯収入・仕事でメールをするか?・私用でメールするか?・インターネットで株取引をするか?

方法3. 基本属性(性・年齢・地域ブロック・学歴・職業分類・収入)を単純に利用した。

また方法2は、前節の共変量選択法の2段階目だけを、各変数ごとでなく、一度に行ったものとして考えることができ、もっとも単純に考えられる共変量選択法であるためにあえて取り上げた。この方法では前節の3段階目の基準を利用していないために、Strong Ignorabilityの仮定のチェックが行われていない。これに対して方法1では3段階目の基準を用いているために、近似的にでも仮定が成立するように共変量を選択しているという点が大きく異なる。

これまでも層別抽出の際には層の重みをかけたり、事後層別として国勢調査の際の属性変数の分布に周辺比率を合わせる Raking 法(Deming and Stephan, 1940)などが利用されてきた。これに対して本研究で利用された方法は、一方の参照標本に共変量分布を合わせる手法であり、その特徴として(1)共変量の周辺分布でなく、同時分布をあわせる(2)傾向スコアという値に一元化することで、Raking 法などと異なり共変量の数が多くても利用できるという点がこれまでの補正法と大きく異なる。本研究では、既存の調整手法の中でよく利用される方法である Raking 法も比較の対象として利用した。

5.3 傾向スコアの計算

本研究での第*i*調査協力者に対する傾向スコア $w(x_i, \alpha)$ の計算には、上限と下限のあるロジスティック回帰モデルを最尤法を用いて母数推定し、その母数 α と共変量ベクトル x_i の値を

表 1. 各方法の正判別率.

年度	方法1	方法2	方法3
2001	90.6%	97.6%	86.5%
2002	89.9%	96.9%	86.5%
2003	89.7%	96.8%	86.5%

用いて下記のように計算する：

$$(5.1) \quad w(x_i, \alpha) = c + \frac{(d - c)}{1 + \exp[-(a^t x_i + b)]}$$

ここで c と d はそれぞれ「インターネット調査に参加する」確率の下限と上限であり、母数として推定する。従ってここでの母数 α は a, b, c 及び d である。一般的によく利用される上限と下限のない ($c=0, d=1$ の) ロジスティック回帰モデルでなく、このモデルを利用したのは、調査への割り当ては、共変量によっては説明されない要素が存在すると考えられるからである。また、このモデルを利用して傾向スコアを計算することで、あまりに小さい傾向スコアはその調査協力者の重みを非常に大きくさせ、解析を不安定にさせる危険性を避けることができる。

それぞれの共変量のセットを用いてインターネット調査か訪問留置調査かという二値変数を基準変数としたロジスティック回帰分析を行い、正判別率を計算した結果を表 1 に記載した。正判別率が高いということは説明変数によって調査の群分けの予測がうまく行われていることを示しているため、方法 1, 3 より方法 2 の方が正判別率が高いのは当然である。

5.4 補正の効果

ここでは 2001 年, 2002 年, 2003 年の 3 回の JGSS 調査に対する 5 つの結果を比較する。具体的には (1) 各 JGSS 本調査の訪問留置調査の回答分布 (2) インターネット調査での回答分布 (3) 提案手法による補正值 (方法 1) (4) 方法 2 によって計算された傾向スコアを用いた補正值 (5) 方法 3 によって計算された傾向スコアを用いた補正值, 及び (6) 方法 3 での共変量である「性・年齢・地域ブロック・学歴・職業分類・収入」に関して Raking 法を行った補正值 (これを方法 4 とする。Raking 法では利用できる共変量数がせいぜい 5, 6 程度であることに注意する) である。それぞれの方法での結果を表 2 から表 4 に示した (それぞれ項目の選択肢順に 1, 2, 3 とつけている。また欠測を表すカテゴリーには 0 をつけている)。また、各 JGSS 本調査の訪問留置調査の回答分布を真値とした、項目単位での各方法の二乗誤差を表 5 に示した。また、インターネット調査での二乗誤差和を 100% とした時の、各方法の二乗誤差和を表 5 に記載した。

更に、各回答カテゴリーごとの訪問留置調査での結果とインターネット調査での結果のプロット (調整前: 左), 及び訪問留置調査での結果と方法 1 による補正值のプロット (方法 1: 右) を図 1 に記載した。また、曲線の中は被験者数を 1000 とした場合の単純無作為抽出による二項分布の比率の推定値の 95% 信頼区間である。

提案手法による補正值 (方法 1) はすべての年, およびすべての項目で、インターネット調査での値に比べて本調査である訪問留置調査の結果に近づいていることがわかる。また基本属性のみを利用して傾向スコアを算出し、これを用いて補正を行った場合 (方法 3) も訪問留置調査に近づくが、提案された共変量選択法を用いて様々な共変量を加えた方がより訪問留置調査に近づくことが分かる (但し方法 3 で利用した属性変数は、前節での 3 段階目の基準に適合している)。これに対して、方法 2 はかえって補正前のインターネット調査の結果よりも二乗誤差が大きくなっている場合がある。このことから、提案された共変量調整法の 3 段階目である、

表 2. 2001 年調査での結果 .

	訪問	ネット	方法1	方法2	方法3	方法4
信頼0	1.15%	0.24%	0.34%	0.15%	0.32%	0.33%
信頼1	23.23%	29.10%	25.88%	27.38%	26.13%	26.92%
信頼2	14.01%	12.46%	12.80%	11.79%	13.06%	12.84%
信頼3	61.61%	58.20%	60.99%	60.68%	60.48%	59.90%
妻の仕事0	2.04%	0.21%	0.14%	0.08%	0.17%	0.18%
妻の仕事1	17.46%	7.72%	9.98%	10.63%	10.42%	9.85%
妻の仕事2	33.91%	31.39%	33.82%	30.10%	33.26%	33.12%
妻の仕事3	31.94%	44.03%	40.93%	45.33%	41.91%	42.21%
妻の仕事4	14.66%	16.65%	15.13%	13.86%	14.24%	14.64%
妻は夫の0	2.69%	0.17%	0.16%	0.10%	0.12%	0.13%
妻は夫の1	11.86%	6.03%	7.12%	6.53%	7.18%	7.40%
妻は夫の2	34.30%	30.94%	33.65%	29.11%	33.99%	33.32%
妻は夫の3	39.39%	44.61%	44.50%	51.74%	44.57%	44.62%
妻は夫の4	11.76%	18.27%	14.56%	12.51%	14.14%	14.53%
貧富解消0	1.25%	0.33%	0.07%	0.07%	0.12%	0.13%
貧富解消1	29.21%	19.84%	22.64%	18.47%	23.28%	22.54%
貧富解消2	24.77%	27.96%	29.87%	27.83%	29.51%	29.76%
貧富解消3	34.12%	36.08%	37.29%	44.25%	37.12%	37.64%
貧富解消4	6.56%	10.91%	5.91%	5.89%	6.38%	6.41%
貧富解消5	4.09%	4.88%	4.21%	3.49%	3.57%	3.51%
外国人増加0	6.38%	0.38%	0.19%	0.27%	0.28%	0.31%
外国人増加1	35.02%	46.92%	42.88%	43.60%	44.39%	46.26%
外国人増加2	58.60%	52.70%	56.94%	56.14%	55.33%	53.44%
安楽死是非0	0.82%	0.21%	0.05%	0.07%	0.17%	0.18%
安楽死是非1	68.64%	72.37%	69.47%	65.36%	71.57%	71.83%
安楽死是非2	5.88%	4.45%	3.69%	5.51%	4.63%	4.45%
安楽死是非3	24.66%	22.96%	26.79%	29.07%	23.63%	23.53%

表 3. 2002 年調査での結果 .

	訪問	ネット	方法1	方法2	方法3	方法4
信頼0	0.51%	0.24%	0.28%	0.10%	0.39%	0.30%
信頼1	20.05%	29.10%	25.66%	26.10%	25.43%	25.32%
信頼2	9.62%	12.46%	13.05%	11.56%	13.06%	13.66%
信頼3	69.83%	58.20%	61.01%	62.24%	61.12%	60.72%
妻の仕事0	1.32%	0.21%	0.16%	0.06%	0.26%	0.24%
妻の仕事1	18.29%	7.72%	10.10%	9.86%	10.45%	10.14%
妻の仕事2	31.66%	31.39%	34.71%	30.75%	33.74%	33.57%
妻の仕事3	34.64%	44.03%	40.70%	44.60%	41.87%	41.96%
妻の仕事4	14.09%	16.65%	14.33%	14.72%	13.68%	14.08%
妻は夫の0	1.59%	0.17%	0.15%	0.12%	0.13%	0.13%
妻は夫の1	13.07%	6.03%	7.33%	6.98%	7.51%	7.25%
妻は夫の2	33.73%	30.94%	34.67%	28.49%	34.27%	33.94%
妻は夫の3	36.91%	44.61%	43.48%	51.05%	43.68%	44.73%
妻は夫の4	14.70%	18.27%	14.37%	13.36%	14.41%	13.95%
貧富解消0	0.75%	0.33%	0.11%	0.06%	0.13%	0.13%
貧富解消1	22.38%	19.84%	23.69%	19.37%	24.33%	24.55%
貧富解消2	25.60%	27.96%	29.28%	28.01%	29.68%	31.31%
貧富解消3	39.99%	36.08%	36.89%	42.39%	36.40%	34.64%
貧富解消4	8.23%	10.91%	6.08%	6.04%	6.05%	6.23%
貧富解消5	3.05%	4.88%	3.96%	4.13%	3.41%	3.13%
外国人増加0	6.33%	0.38%	0.16%	0.57%	0.34%	0.44%
外国人増加1	37.08%	46.92%	42.90%	43.31%	44.44%	45.24%
外国人増加2	56.59%	52.70%	56.94%	56.13%	55.22%	54.31%
安楽死是非0	0.24%	0.21%	0.10%	0.07%	0.26%	0.22%
安楽死是非1	59.36%	72.37%	69.56%	64.68%	71.46%	72.25%
安楽死是非2	7.11%	4.45%	3.97%	4.98%	4.69%	4.32%
安楽死是非3	33.29%	22.96%	26.37%	30.28%	23.60%	23.20%

表 4. 2003 年調査での結果 .

	訪問	ネット	方法1	方法2	方法3	方法4
信頼0	0.46%	0.24%	0.23%	0.13%	0.30%	0.28%
信頼1	20.95%	29.10%	24.26%	24.81%	24.85%	25.21%
信頼2	11.19%	12.46%	12.64%	10.04%	13.48%	12.94%
信頼3	67.40%	58.20%	62.88%	65.02%	61.37%	61.57%
妻の仕事0	1.23%	0.21%	0.10%	0.04%	0.19%	0.20%
妻の仕事1	19.88%	7.72%	10.51%	11.21%	10.73%	10.53%
妻の仕事2	31.99%	31.39%	33.65%	28.98%	32.93%	33.22%
妻の仕事3	31.89%	44.03%	40.13%	44.67%	42.17%	43.10%
妻の仕事4	15.02%	16.65%	15.61%	15.10%	13.98%	12.94%
妻は夫の0	2.10%	0.17%	0.12%	0.10%	0.12%	0.12%
妻は夫の1	13.85%	6.03%	7.77%	7.18%	7.67%	7.45%
妻は夫の2	33.88%	30.94%	34.01%	27.98%	33.29%	33.33%
妻は夫の3	35.10%	44.61%	42.86%	50.59%	44.29%	44.56%
妻は夫の4	15.07%	18.27%	15.24%	14.15%	14.62%	14.54%
貧富解消0	0.51%	0.33%	0.09%	0.08%	0.13%	0.13%
貧富解消1	26.47%	19.84%	24.61%	19.70%	25.17%	24.67%
貧富解消2	28.41%	27.96%	30.20%	28.14%	29.55%	29.10%
貧富解消3	34.75%	36.08%	36.35%	43.71%	36.13%	36.23%
貧富解消4	6.59%	10.91%	5.55%	5.68%	5.74%	6.13%
貧富解消5	3.27%	4.88%	3.20%	2.70%	3.28%	3.73%
外国人増加0	6.85%	0.38%	0.23%	0.35%	0.32%	0.34%
外国人増加1	38.17%	46.92%	42.32%	44.72%	44.45%	45.22%
外国人増加2	54.98%	52.70%	57.45%	54.93%	55.23%	54.43%
安楽死是非0	0.15%	0.21%	0.06%	0.09%	0.23%	0.24%
安楽死是非1	62.08%	72.37%	68.91%	65.14%	71.64%	73.43%
安楽死是非2	7.61%	4.45%	4.29%	4.40%	4.79%	4.54%
安楽死是非3	30.15%	22.96%	26.74%	30.37%	23.34%	21.78%

表 5. 2001-2003 の項目ごとの二乗誤差 .

	ネット	方法1	方法2	方法3	方法4	
2001年	信頼	0.0049	0.0010	0.0024	0.0011	0.0019
	妻の仕事	0.0255	0.0141	0.0245	0.0153	0.0167
	妻は夫の	0.0121	0.0063	0.0215	0.0061	0.0062
	貧富解消	0.0122	0.0081	0.0230	0.0068	0.0083
	外国人増加	0.0212	0.0103	0.0117	0.0136	0.0190
	安楽死是非	0.0019	0.0011	0.0031	0.0012	0.0014
	対ネット比		52.36%	110.57%	56.60%	68.71%
2002年	信頼	0.0225	0.0121	0.0098	0.0117	0.0130
	妻の仕事	0.0208	0.0115	0.0173	0.0119	0.0121
	妻は夫の	0.0131	0.0079	0.0269	0.0079	0.0090
	貧富解消	0.0038	0.0031	0.0027	0.0039	0.0071
	外国人増加	0.0147	0.0072	0.0072	0.0092	0.0069
	安楽死是非	0.0283	0.0162	0.0042	0.0246	0.0277
	対ネット比		56.07%	65.93%	66.98%	73.74%
2003年	信頼	0.0153	0.0034	0.0022	0.0057	0.0055
	妻の仕事	0.0300	0.0160	0.0249	0.0193	0.0220
	妻は夫の	0.0174	0.0101	0.0324	0.0127	0.0135
	貧富解消	0.0067	0.0011	0.0128	0.0006	0.0006
	外国人増加	0.0123	0.0067	0.0085	0.0082	0.0092
	安楽死是非	0.0168	0.0069	0.0020	0.0146	0.0208
	対ネット比		44.86%	84.06%	61.97%	72.84%

Strong Ignorability の仮定が成立するように共変量を選択することは非常に重要であることがわかる。また方法 4 は同じ共変量を利用している方法 3 よりも調整がうまく行っていない。

また、表 2 から表 5 は欠測も 1 つのカテゴリーとした場合の結果であるが、欠測値の割合は調査方法によって大きく違うことが様々な研究により指摘されており(前田・土屋, 2001; Couper, 2000), 実際今回の場合も極端に異なっていることがわかる。これは調査対象の集団の違いと

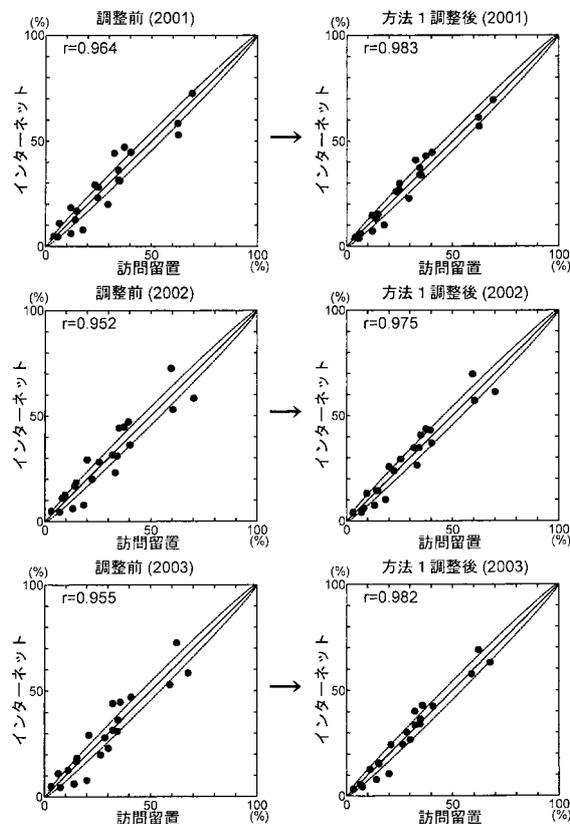


図 1. 2001 年・2002 年・2003 年各調査での訪問留置調査とネット調査(調整前・調整後)の
 カテゴリごとの比率のプロット。

いうより、調査方法の違い、つまり「調査モード」の効果によるものと考えられる。

そこで、各推定値からそれぞれ欠測カテゴリーを除去して、あらためて各項目で回答の合計が 100% になるように基準化した場合の結果を表 6 に記載した。

欠測カテゴリーを除去した方が方法 1 による補正は効果があることがわかる。また、図 2 は欠測カテゴリーをそのままにした場合の、項目単位での訪問留置調査と提案推定値の二乗誤差をプロットしたものである。

図 2 中の直線は $y = x$ の直線であり、そこから下にずれるほど、「インターネット調査の平均」よりも提案された補正法による推定値の方が「訪問留置調査の平均」に近いことが分かる。上記の結果から、項目の内容とはあまり関係なく、インターネット調査の結果が訪問留置調査の結果と大きく異なる項目ほど、提案された手法による補正がうまく働いていることが分かる。

このように、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、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 2004 年に実施されており、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

表 6. 2001-2003 の項目ごとの二乗誤差(無回答除去).

		ネット	方法1	方法2	方法3	方法4
2001年	信頼	0.0051	0.0009	0.0023	0.0011	0.0019
	妻の仕事	0.0248	0.0132	0.0236	0.0145	0.0159
	妻は夫の	0.0112	0.0051	0.0198	0.0048	0.0050
	貧富解消	0.0124	0.0079	0.0226	0.0067	0.0082
	外国人増加	0.0188	0.0062	0.0080	0.0101	0.0162
	安楽死是非	0.0017	0.0009	0.0032	0.0009	0.0011
	対ネット比		45.97%	105.66%	51.53%	65.39%
2002年	信頼	0.0230	0.0124	0.0102	0.0119	0.0130
	妻の仕事	0.0204	0.0110	0.0168	0.0115	0.0121
	妻は夫の	0.0126	0.0072	0.0260	0.0072	0.0090
	貧富解消	0.0039	0.0031	0.0026	0.0039	0.0071
	外国人増加	0.0113	0.0023	0.0031	0.0050	0.0069
	安楽死是非	0.0284	0.0161	0.0041	0.0247	0.0277
	対ネット比		53.14%	69.60%	61.70%	73.74%
2003年	信頼	0.0155	0.0035	0.0023	0.0058	0.0057
	妻の仕事	0.0297	0.0156	0.0245	0.0189	0.0216
	妻は夫の	0.0165	0.0091	0.0312	0.0116	0.0124
	貧富解消	0.0068	0.0010	0.0127	0.0005	0.0006
	外国人増加	0.0075	0.0004	0.0030	0.0026	0.0039
	安楽死是非	0.0169	0.0069	0.0020	0.0147	0.0210
	対ネット比		40.21%	86.19%	56.49%	66.30%

調整後の二乗誤差

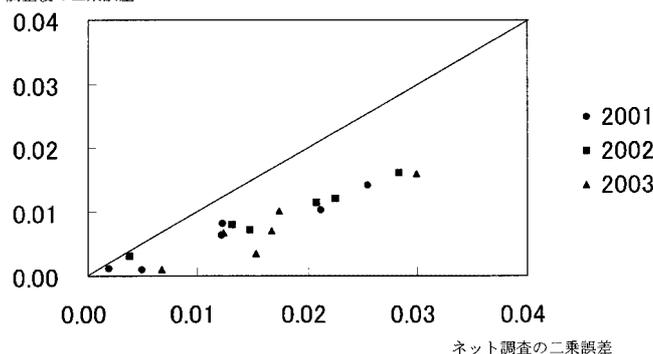


図 2. 二乗誤差のプロット.

あることを示している.

5.5 リサンプリングによる調整法の妥当性の検討

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人ずつを無作為抽出し、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を行い、補正を行うという処理を1000回行った.

各質問でのネット平均及び提案手法(方法1)について、1000回のリサンプリングごとに推定値を得た. また、6つの質問項目の回答分布を多項分布とみなして、訪問平均と有意な差があるかどうかを最尤法による尤度比検定及び星野(2005)によって提案された重み付き尤度比検定を用いて、有意水準5%で棄却されるかどうかを調べた. これを2001年から2003年度の3年度分を行うことで、 $3 \times 6 = 18$ の質問のうち、訪問調査での回答と有意に差があるとされた質問数の平均は、ネット平均では8.21、方法1では4.34であった. 棄却された質問数の分布を図3に記載した. また、インターネット調査での二乗誤差和を100%とした時の、各方法の二乗誤

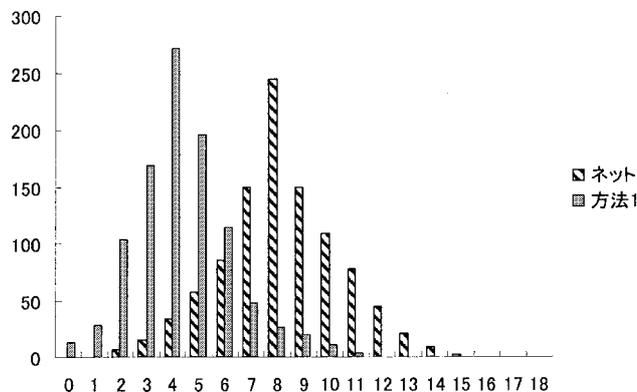


図 3. 1000 のリサンプリングデータ中における訪問調査と差が有意であると判定された質問数の分布。

表 7. 二乗誤差和の平均と標準偏差(下段は無回答除去後の結果)。

	方法1		方法2		方法3		方法4	
	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差
2001年	53.12%	4.31%	112.21%	5.33%	56.72%	4.68%	69.03%	4.55%
	46.03%	3.99%	107.73%	4.90%	51.83%	4.33%	65.77%	4.11%
2002年	56.85%	4.41%	66.31%	4.79%	67.13%	4.44%	74.00%	3.98%
	53.67%	4.03%	70.31%	4.78%	61.73%	4.20%	74.13%	4.03%
2003年	45.03%	3.89%	84.20%	4.80%	62.33%	4.11%	73.16%	3.89%
	40.54%	3.98%	86.30%	4.82%	56.79%	4.32%	66.65%	3.79%

差和の平均と標準偏差を表 7 に記載した。

図 3 と表 7 からは、明らかに提案手法はネット平均をそのまま利用するよりも訪問調査の結果に近い調整を行っていることがわかる。

6. 議論

本論文では傾向スコアを用いた調査の補正の際の共変量の選択基準を提案し、実際の社会調査に対して、有意抽出標本データから無作為抽出標本データへの補正を行った。

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。これ以外の複数の市場調査においても、同様に提案された共変量選択法を用いた傾向スコアによる補正の再現性が確認されているが、このことについては今後別論文にて発表していく予定である。

新しい技術の性質を精査し、研究の一般化可能性の範囲を検討しておくために、このような研究事例をより多く積むことが重要である。例えば、本稿の結論が (i) おおむねどのような有意標本に対しても成り立つのか (ii) 特定の集め方をした登録者集団でより良く成立するのか、等の点については今後の検討を待たねばならない。

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、第 4 節で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。

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。

さらに、近年調査環境が著しく悪化しており、今後とも台帳等からの無作為抽出による訪問調査を行っても拒否が多くなると予想される。従って、わざわざ回答してくれた協力者が益々特殊な集団になっていくことが今後予想され、たとえ住民基本台帳等からの無作為抽出による調査であっても、その結果が本来の結果から大きく異なる可能性は大きくなっていくであろう。従って、調査に回答するかどうかを共変量情報を用いて説明し、補正するという本方法は、今回のようなインターネット調査の補正だけでなく、既存型調査を国勢調査などのより大規模な調査に向けての補正を行う方法としても有効かもしれない。これについての応用可能性は今後の研究課題である。

本稿の補正法は、拒否による調査不能に対する直接の解決はもたらさないが、調査協力者から得られる情報のみにもとづき、拒否による回答をある程度推察する際の加重補正法に関する示唆的な研究が土屋(2005)でなされている。

本研究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。今後の様々な実験研究が望まれる。

7. 文末注

日本版 General Social Surveys(JGSS)は、大阪商業大学比較地域研究所が、文部科学省から学術フロンティア推進拠点としての指定を受けて(1999-2003 年度)、東京大学社会科学研究所と共同で実施している研究プロジェクトである(研究代表:谷岡一郎・仁田道夫,代表幹事:佐藤博樹・岩井紀子,事務局長:大澤美苗)。東京大学社会科学研究所附属日本社会研究情報センター SSJ データアーカイブがデータの作成に協力している。

謝 辞

(株)日経リサーチの鈴木督久氏、坂巻秀一氏、長野孝彦氏には、本実験調査にあたってご協力を頂きましたことをここに深謝いたします。匿名の査読者の方には本論文の改稿に際して参考になるコメントを頂いたことを感謝いたします。また、本研究は文部科学省科学研究費補助金基盤研究(B)(2)課題番号 16330114(研究代表者:前田忠彦)、および平成 16 年度財団法人大川情報通信基金研究助成(研究代表者:星野崇宏)、2005 年度財団法人吉田秀雄記念事業財団研究助成(研究代表者:星野崇宏)による財政的支援を受けました。

付録: JGSS で利用した共変量の全候補

今回、共変量の候補として、補正の目的変数以外にインターネット調査を行ったのは以下の 32 項目である。それぞれ括弧内には JGSS での変数名を記載する。

性(sex)年齢(age)地域ブロック(block)都市規模(f04)学歴(xxlstsch)父学歴(pplstsch)母学歴(mmlstsch)配偶者学歴(sslstsch)職業分類(xxwpl)世帯収入(szhsincm)住居形態(tp8dwel)本を読む頻度(fq5read)新聞を読む頻度(fq5newsp)仕事以外の旅行(fq5trip)飲酒(dodrink)

喫煙(dosmoke+dosmokex) 先週の状況(tp5unemp) 先週仕事をしていたかどうか(xjob1wk) 15歳のころの居住地(tp6loc15) 15歳のころの父親の職業(ppjbto15) 15歳のころの母親の職業(mmjbtp15) 15歳のころの世帯収入(opffix15) 世帯収入水準の自己評定(op5ffinx) 同居人数(szff+szfft1) 仕事でメールをするか(?(doemailj) 私用でメールするか(?(doemailp) インターネットで株取引をするか(?(donetstk) 携帯電話またはPHSを利用するか(?(dophs) ファックスを利用するか(?(dpfax) 所属階層の自己評定(op5levk) 実子の数(ccnumttl) テレビ視聴時間(hrtv).

参 考 文 献

- Berrens, P., Bohara, A. K., Jenkins-Smith, H., Silva, C. and Weimer, D. L. (2003). The advent of internet surveys for political research: A comparison of telephone and internet samples, *Political Analysis*, **11**, 1–22.
- Couper, M. P. (2000). Web surveys: A review of issues and approaches, *Public Opinion Quarterly*, **64**, 464–494.
- Deming, W. E. and 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.
- 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(2005). 欠測群の周辺分布の母数に対する傾向スコアを用いた重み付き M 推定量の提案と介入効果研究への応用, 行動計量学, **32**, 121–132.
- 星野崇宏, 繁樹算男(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. and Shigemasu, K. (2005). A propensity score adjustment for multiple group structural equation modeling, *Psychometrika* (in press).
- 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, No. 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, 東京.
- 大阪商業大学 比較地域研究所, 東京大学 社会科学研究所(2000–2003). 『日本版 General Social Surveys 基礎集計表・コードブック』, 東京大学 社会科学研究所.
- Rosenbaum, P. R. and 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.
- Taylor, H. (2000). Does internet research work?, *International Journal of Market Research*, **42**, 51–63.
- 土屋隆裕(2005). 調査不能者の特性に関する一考察——「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.

Applying Propensity-score Adjustment to Social Surveys with Non-random Sampling and a Selection Criterion for Covariates

Takahiro Hoshino¹ and Tadahiko Maeda²

¹Department of Cognitive and Behavioral Science, The University of Tokyo

²The Institute of Statistical Mathematics

The advent of the “information age” has led to a great diversity of individual views and has created an environment in which individuals’ beliefs can change rapidly. Given these rapid changes, it is necessary to conduct frequent surveys in every area of the social sciences. However, survey response rates have gradually decreased due to privacy concerns and lifestyle changes. There are therefore increased calls for new survey methods that can cost-effectively produce representative samples.

Web-based surveys have been devised in various areas. However, the conventional Web-survey method of purposive sampling has drawn criticism for its lack of representativeness. In this study, we applied a propensity-score adjustment method to Web-based surveys to predict Japanese general social surveys (nation-wide random-sample surveys). We propose a new method for selecting covariates and show that the propensity-score adjustment by this method is more effective than that using conventional demographic variables as covariates for estimating propensity scores.