

修士論文（2003年度）

インターネット調査における加重修正法の適用可能性

A Study of Weighting Adjustment in Internet Survey

3602B059-8 松田 浩幸

早稲田大学大学院 理工学研究科

機械工学専攻 経営システム工学専門分野

オペレーションズ・リサーチ研究

指導教員

早稲田大学理工学部 逆瀬川 浩孝 教授
文部科学省統計数理研究所 大隅 昇 教授

1 章 研究の概要	1
1.1 研究の背景	1
1.2 研究の目的	3
1.3 研究の概要	4
2 章 インターネット調査	7
2.1 インターネット調査と調査環境の関係	7
2.1.1 国内のインターネット利用状況	7
2.1.2 調査環境の悪化	7
2.1.3 海外の調査環境とインターネット調査の位置づけ	8
2.2 インターネット調査とは	9
2.2.1 定義	9
2.2.2 Web 調査の特徴	11
2.2.3 インターネット調査の諸特性	11
2.3 実験調査	13
2.3.1 調査環境の変化に対応した新たな調査法の研究	13
2.3.2 実験調査の基本方針	13
2.3.3 第 4 次実験調査計画の概要	14
3 章 研究の対象	17
3.1 WEB 調査の形態	17
3.2 母集団と標本の関係	19
3.3 WEB 調査の形態と標本からの推論可能性	20
3.4 第 4 次実験調査で用いた登録者集団とその構築方法	21
3.5 公募系と非公募系	22
3.5.1 公募系の特徴	22
3.5.2 非公募系の特徴	24
3.6 研究の対象と適用可能性	25
4 章 傾向スコアによる加重修正法	28
4.1 実験研究と観察研究	28

4.2 傾向スコア	29
4.3 STRONGLY IGNORABLE の仮定	31
4.4 傾向スコアによる加重修正法	32
4.5 最適化問題	33
4.6 傾向スコアによる加重修正法の適用例	34
4.6.1 Harris Interactive 社による適用例	34
4.6.2 従来型調査の結果と Web 調査の調査結果間の修正	35
4.6.3 適用例の問題点	35
5 章 実験の方法	36
5.1 2つの調査結果に与える要因	36
5.2 適用条件	37
5.3 適用データの特徴	38
5.4 実験の目的	40
5.5 実験の手順	41
5.6 実験計画	44
6 章 結果と考察	48
6.1 傾向スコアの推定	48
6.2 実験 1 の結果	48
6.3 実験 2 の結果	51
6.4 実験 3 の結果	53
6.5 実験 4 の結果	55
6.6 結果のまとめ	57
7 章 考察ならびに今後の課題	59
7.1 考察	59
7.2 今後の課題	61
【参考文献】英文	63
【参考文献】和文	65

1 章 研究の概要

1.1 研究の背景

1990年代前半から一部で普及が始まっていたインターネットはデータの伝送速度の遅さや従量制の料金といった事情もあり、流通する情報はテキスト情報が中心であり、利用方法としても、電子メールや情報検索が主であった。また、利用者は必要なときにインターネット接続サービス事業者にダイヤルアップ接続し、料金を気にしながらインターネットを利用するという状態が最近まで続いてきた。しかし、ブロードバンド時代の到来に伴い、どんな情報でも短時間で送受信が可能となり、ネットワークの常時接続、パソコンだけではなく、携帯電話や情報家電などさまざまな機器を用いて、多様な生活場面でインターネット活用が可能となる状況が現実のものとなりつつあり、インターネットの利用のあり方について考えていく時期にさしかかっている。

「調査」についても再考の時期にあるといえる。ネットワークを利用したインターネット調査は従来型調査方式では不可能であった動画や音声を含めた多様かつ複雑な調査を行えるようになった。しかし、従来にない特性だからこそ、デメリットにもなりうる。そして、その影響についてはあまり意識されず、利点のみが強調されているのが、現状である。加えて、統計の専門家からは「母集団が定義できないから調査でない」、「インターネット調査は駅前調査と同じ」などの否定的な意見も聞こえてくる。インターネット調査が一般的に認知されつつある状況の中、これを科学的でないとして割り切ることに問題がある。むしろ、統計学者や調査法研究者はこの現実に目を向け、インターネット調査が「科学的な調査法または調査方式」になりうるべく研究や議論を行っていかないとはいえないと考える。

こうした現状のもと、インターネット調査の特性を調べるべく、長年に渡り大隅他(2003)は複数の調査機関の協力を得て、実証実験を繰返し行っている。実験結果から、ようやくインターネット調査の特性が臚気ながら見えてきたが、解決すべき問題は沢山ある。

インターネット調査の諸問題の一つに、インターネット調査の回答は属性分布や回答分布の偏りがある。これが原因で信頼性が疑わしいと言われることがある。本研究はこの“偏り”を解決する手段の一つである加重修正法を適用し、偏りを修

正することを考える。

インターネット調査の回答は、セルフセレクション方式(自記式)であるために、その抽出元である登録者集団(リソース/パネル)の構築方法に依存する(大隅,2002a)。構築方法は登録者集団により様々であるが、大別して確率的な抽出を経て構築されたか、否かで分類できる(Couper,2003)。前者の手続きを経て構築された登録者集団を非公募系、後者を公募系と呼ぶ。国内におけるインターネット調査はほとんどが公募系である。

公募系は、Web オープン型で公募し、相当数の標本を確保しやすいという利点があるが、回答者の特徴として、調査サイト登録数が多いこと、調査回答の参加頻度が多いこと、回答率が低い等が指摘されており、インターネット・ユーザーの中のかなり偏った集団 (professional respondents, プロの回答者集団) であることが実証的に示されている (吉村・大隅, 2003)

一方、非公募系は、標本抽出枠が明らかな母集団から確率的な標本抽出を経て作られた標本に基づいて構築されている。しかし、インターネット・ユーザーに限定されること、さらに抽出対象となっても調査への応諾を得た者に限られること等の条件付きの標本となっている。一般に応諾率は高いとは言えないので、公募系と比べて、多数の登録者を確保するための経費面の負担が大きい。公募系と非公募系を比べる限り、非公募系の集団が現実の母集団 (インターネット・ユーザーの集まり) により近いと考えられる。しかし、非公募系はインターネット・ユーザーを代表するわけではない。

調査観点の費用を考えると、公募系の利用頻度を多くし、非公募系のそれを少なくすることが望まれる。しかし、公募系と非公募系は回答者の性質が異なるために、利用目的に合わせて使い分けるのが現状である。そこで、調査実施回数は少ないが構築過程が明らかで、調査結果が信頼できると思われる非公募系の標本構成に基づいて、公募系の結果を修正することを考える。つまり、加重修正法の適用が可能であるとしたとき、利用頻度や登録標本数が多い公募系の結果を、あたかも非公募系の結果と同じ感覚で用いることができるかを考える。具体的には、ある特定の項目の回答、反応変数に対して、回答者の性質のみが異なるとした公募系の回答結果に対して、仮にこれを、非公募系の登録者が回答した場合の反応変数の挙動 (平均や分散の推定) として推論する。

ところで、加重修正法はインターネット調査とは別に長年に渡り研究されてきたという歴史的背景がある(Bethlehem,2002)。加重修正法(weighting adjustment)とは、母集団と標本調査の間に生じる不一致を調整するための補正(修正)、無回答による偏りなどを補正(修正)する一技法のことであり、欠損値処理、補完や決めつけ処理(imputation)、カテゴリカルデータの分析手法(対数線形モデル)、分割表の補正法と関連して発展してきた。最も単純な加重修正法である事後層化法(post stratification)は無回答などによりセル間に生じた偏りを補正するためによく用いられる。例えば、電話世論調査において、高学歴者は一般に回答する傾向が高いため、学歴でブレイクダウンした時に高学歴者層とそれ以外のセル間に偏りが生じ、加重修正が使われることがある。(Gelman & Carlin,2002)。しかし、事後層化法は、抽出率の逆比を重みに用いるために、変数の数が多くなると、ブレイクダウン後にセルが疎になるという問題がある。それらの問題を解決するために、多数の変数を同時に扱うことができる傾向スコア(Rosenbaum & Rubin,1983)を応用した加重修正法(傾向スコアによる加重修正法)が提案された(Rosenbaum,1987)。最近になり、偏りが指摘されるインターネット調査の結果に傾向スコアによる加重修正法を適用するという提案がされ、注目を浴びている。

適用例の一つに米国の Harris Interactive 社 (Taylor 他, 2001) が、2000 年米大統領選挙の予測に用いた例がある。国内では星野・鈴木 (2003) や吉村・松田・大隅他 (2003) の例がある。一方で、インターネット調査における加重修正を行うことについて、目標母集団の不透明さ、登録者集団の作り方、低い回答率などを考えると、従来の方法論をそのまま適用することは難しいという指摘もある(大隅,2003)。

1.2 研究の目的

Taylor 他 (2001) は、電話調査と Web 調査の間で行った傾向スコアによる修正効果の意味があったとしているが、その具体的な手順が書かれていない。しかも、面接や郵送調査などの従来型調査方式とインターネット調査との比較実験を続けてきた大隅らの研究成果に照らし合わせると、異なる調査方式間の修正処理がそう簡単に行えるとは思えない。例えば、吉村・松田・大隅他 (2003) は、傾向スコアによる加重修正法の特性を慎重に吟味すると、異なる調査方式の差異までは修

正が困難であると指摘している。このように加重修正法を適用する状況や前提が異なる中で、修正効果の有無について議論することは難しい。少なくともインターネット調査の諸特性ならびに傾向スコアによる加重修正法の理論の範囲を考慮した上での適用が前提であるが、それを重視した研究も見当たらない。

本研究は、登録者集団の構築過程の異なる公募系、非公募系で得られたインターネット調査における回答間で、傾向スコア加重修正法を用いて修正を行うことを試みる。修正効果の有無を確認し、修正過程で生じる問題を整理しながら、加重修正法の特徴、インターネット調査における加重修正法の適用可能性について考察する。

1.3 研究の概要

公募系、非公募系それぞれに対して、同じ調査環境下で実査を行う。2つの調査結果の違い、反応変数の違いが登録者集団の性質にのみ依存する状況で、公募系の結果に対して傾向スコアによる加重修正法を適用し、仮に非公募系が回答した場合の結果を推定する。非公募系は公募系に比べて、調査結果が信頼できるが、登録者集団を構築するためのコストが大きく、調査疲労を小さくするために、利用回数を少なくしたい。一方、公募系は Web 上での公募により構築された登録者集団(リソース/パネル)であるために、相当数のサンプルを集めやすいという利点があるが、インターネット・ユーザーの中のかかなり偏った集団であり、非公募系の結果の代用にはならない。仮に公募系の結果を用いて、加重修正が適切に行われ、非公募系の結果を推定することができれば、利用価値は十分にあると考える。

公募系、非公募系は標本構成が異なるために、傾向スコアによる加重修正法を用いて共変量調整を行う。具体的には公募系の共変量の分布が非公募系の分布に一致するように、言い換えれば標本構成が同一になるように、ロジスティック回帰により求めた傾向スコアを用いて、加重修正を行う。共変量調整を確認後に、得られた傾向スコアを用いて反応変数の修正を試み、修正効果を測定する。

ここでは、共変量が事前に分かっていないため、幾つかの共変量候補となる変数を拾い、共変量セットを構築する。共変量セットの構築方針として、性別・年齢・職業などの基本属性のみで構成された共変量セット2組(共変量セット1または2)と基本属性にインターネット利用に関する基本属性に相当するインターネッ

ト歴や通信回線などを加えた共変量セットの2組(共変量セット3または4)を用いる。これらの共変量セットが修正効果にどのような影響があるかを確認する。

反応変数には所有している通信機器を取り上げた。ここでは、「PHS」「デジタルカメラ」「携帯電話」「カーナビ」など8種類である。共変量セットと反応変数を用いて、修正効果を測定する。

最初に、傾向スコアによる加重法による修正効果を測定するために、同一の共変量セットを用いて、複数の反応変数に対して修正を試みる。(実験1)。ここでは、修正効果が確認されたことと、反応変数に応じて修正の程度が異なる(修正効果に差がある)ことを確認した。

さらに、異なる共変量セットを用いた場合に、修正効果がどのように異なるかを実験する(実験2)。ここでも、実験1で得られ結果と同様に、反応変数により修正の程度が異なることが観測された。さらに、共変量の数と修正効果は関係なく、反応変数に応じた共変量を探索する必要があることが知見として得られた。また、修正がうまくいかなかった反応変数に対しては、観測されていない共変量が原因であると考えられる。

実験1ならびに2では1変数あたり複数の選択肢(共変量)からなる多値変数を用いて実験を行っていたが、ここで共変量セットの構築方法を変えて、修正を試みる。複数の選択肢を事前に2値化処理をし、それらを組み合わせて共変量セットを構築する。このとき、原則として共変量セット4で使われた共変量を2値化し、それらを組み合わせる。多値変数で構築された共変量セットでは修正がうまくいかなかった反応変数に対して、2値変数の組み合わせた共変量セットでは修正が機能するかを実験する(実験3)。共変量セット4では修正しなかったが、共変量セット5では修正されている様子を確認した。修正されない理由として観測されていない共変量の影響と考えることができたが、共変量セットの構築の仕方を変えれば修正できる可能性があることをこの結果は示唆している。

最後に、実験1また2で得られた修正効果が確認された反応変数と共変量セットの組を用いて(これらの組合せが最適だと仮定して)、2つの反応変数の差の要因が回答者と調査票のみであると考えられるデータセットに加重修正を試みる。加重修正により回答者の影響を除去し、調査票の影響を検出する(実験4)。ここでは、加重修正が適切に機能していると仮定をしているために、修正後の差異が確認さ

れば調査票の影響があったといえる。4つの反応変数ならびに共変量セットの組に対して実験を行ったところ、どれも修正後に反応変数の差異を確認した。つまり、調査票の影響があったことを確認した。

2章では大隅(2003)が行ってきた実験調査から分かったインターネット調査の諸特性について、必要だと思われる事項を取り上げる。3章では登録者集団の特性ならびに本研究が扱う研究の対象について説明をする。4章では、観察研究における共変量調整の必要性、ならびに共変量調整の一手法である傾向スコアによる加重修正法を説明し、5章では扱うデータの特徴、取得環境を整理し、傾向スコアによる加重修正法の修正効果を測定できる実験計画について記す。6章では、5章で受けた実験を行い、その結果をまとめる。7章では、得られた知見をまとめ、考察ならびに今後の課題を述べる。

2章 インターネット調査

2.1 インターネット調査と調査環境の関係

2.1.1 国内のインターネット利用状況

総務省通信白書によると平成11年末における国内の15～69歳までのインターネット・ユーザー数は2,706万人、平成12年末における国内の15歳以上79歳以下の個人におけるインターネット・ユーザー数は4,708万人と推計されている。平成17年(2005年)におけるインターネット・ユーザー数を推計したところ、8,720万人まで増加するものと見込まれる。ここでは、「インターネット・ユーザー」を、「インターネット(ウェブ又は電子メールのどちらかのみの場合も含む。)を自宅・自宅外を問わず、パソコン、携帯電話、携帯情報端末、家庭用ゲーム機、インターネット接続機器を設置したテレビ受像機により利用している人」と定義している。ビデオリサーチネットコム調べでは、インターネット・ユーザーは約5,670万人と推計されている。他の調査会社が算出している数字自体も確定しないが、インターネット・ユーザー数が増加傾向にあることは同じである。今後定額料金、常時接続、高速通信網の整備、携帯端末の普及、高等教育における情報教育の導入などを考えると、今後もインターネット・ユーザー数は増加していくことが想像できる。

2.1.2 調査環境の悪化

大隅(2000)によると国内の調査環境は欧米の調査環境に比べて、恵まれていた。諸外国ではほとんど不可能とされる住民基本台帳、選挙人名簿の閲覧利用に基づくサンプリング操作やそれに基づいて作成されたマスターサンプルを利用した調査が可能であった。このため、統計理論に則った理想的な標本設計が可能な環境下で、実査が行うことができた。しかし、近年では個人情報保護の観点から住民基本台帳の閲覧制限、調査経費の増大、回答拒否・回収不能などによる回収率の低下、質のよい調査員確保の困難、など調査環境の悪化が指摘されている。従来型調査(面接や郵送調査など)の実施が益々困難な方向に向かっている。そのため、調査の現場では調査方法の見直しを余儀なくされている。

一方、国内のインターネット調査の利用が急増している。日本マーケティング・リサーチ協会(JMAR)によると、アドホック調査を100としたときの2001年のイ

インターネット調査の売り上げ比率が 5.5%に対して、2002 年の売り上げ比率が 13.0%と前年比 2 倍以上になっている。

調査環境の悪化に伴い、インターネット利用技術の向上、ネットワークインフラの整備に伴うインターネット利用者の増加などを背景に、インターネット調査に期待が集まっている。

2.1.3 海外の調査環境とインターネット調査の位置づけ

欧米では日本国内のような理想的な標本抽出が困難であり、必然的に別の方法論が発展した。それが電子調査情報収集法(CASIC:Computer-Assisted Survey Information：参考を参照)であり、コンピュータ支援のデータ収集法(CADAC:Computer-Assisted Data Collection)である。これらは電子的にデータや情報をどのように取得するかの方法論の研究であり、インターネット調査もこの中に含まれる。欧米では、長い時間をかけて蓄積された実績や研究の延長線上にインターネット調査が位置づけられている。日本でも RDD 方式(欧米の RDD 方式とは異なる)による電話調査もそのような枠組みの中で議論され、CATI(Computer-Assisted Telephone Interviewing)のような調査方式(調査モード)として定着された。吉村(2003b)は調査方式を①モードの種類②インタビュアーの有無③コンピュータの利用の有無の 3 点の組み合わせから分類している。

表 2.1.3 調査方法の分類

	コンピュータの支援あり			
	なし		CASIC, CADAC	
	インタビュアー		インタビュアー	
調査モード	あり	自記式	あり	自記式
面接	面接	面接留置	CAPI	CASI
郵便		郵送		Disk by Mail
電話	電話	FAX	CATI	
インターネット			OFG	e-mail, Web

CATI:Computer-assisted telephone interviewing

CAPI:Computer-assisted personal interview

CASI:Computer-assisted self-interview

OFG:Online focus group

これらの用語については、大隅(2003)参照してもらいたい。

国内ではすぐれた調査環境を利用することができたために、CASICのような電子的なデータの取得に関する研究が重視されてこなかった背景がある。一般社会に急速に普及したインターネット調査が日本国内で蓄積された調査技術が十分に活かされる事なく普及し、利用のあり方などの十分な研究や議論もなされず、現在に至っている(大隅,2002a).

参考 CASIC の特徴(大隅,2003)

- サンプルング：CASIC の基本原理は、サンプルング操作の割り付け処理が自動化され、自動サンプル抽出過程と共に統合(integrate)され、かつデータ収集が自動化された
- コスト,モニタリング：コストの査定や管理が的確になること、レポート作成が容易であること
- 進捗過程モニタリング：これが容易であること：例えば、待ち行列リスト、呼び出し記録リスト、現状確認レポート、履歴レポート、面接印実績レポート、所要時間記録、除外記録、診断レポート取得データは個票レベルでホストに転送、管理スタッフやスーパーバイザに提供される
- データ処理：CASIC では、インタビューの段階でオンライン・エディットでデータ・クリーニング等がされることが特徴
- 分析：データ品質がよい、欠測が減る、不適切回答の現象、決め付け処理、などの利点がある
- 調査結果の公開/開示：結果公表が即座にできる。つまり、データ収集後に即、公表可である
- 方法論研究：基本的にデータ収集が自動的であるか、いろんな調査法によるリーチが可能である。よって、調査エキスパートとユーザの両者にとっても利得となる。

2.2 インターネット調査とは

2.2.1 定義

インターネット調査、インターネット・サーベイ(Internet survey),インターネット・リサーチ(Internet research) ,オンライン・リサーチ(Online research), オンライン

ン調査(Online survey),オンライン定性調査(Online qualitative research), 電子調査 (Electronic survey)など, 国内でインターネットを使った調査というと, 様々な名称で呼ばれているようであるが, ここでは一貫して「インターネット調査」と呼ぶこととし, 次のように定義する.

[インターネット調査の定義] (大隅,2003)

- ・ 電子メール調査 (EMS: Electronic mail survey) と,
- ・ Web 調査 (Web-based survey, Web survey)

に分けられ,

- 電子的データ取得法 (CASIC, CADAC) の一つであって, Web 調査, 電子メール調査など, ネットワークや WWW 環境を用いた調査システム下で実施される調査のこと
- “主として”インターネット・ユーザーを調査対象者とする
- コンピュータ支援の自記式調査 (Computer-assisted self-administered) であること
- とくに Web 調査では, ブラウザで閲覧する形式の電子調査票を用いること (HTML, XML などの言語で記述)
- 回答 (取得) がネットワーク上でリアルタイムに転送記録されること

以上から,

「インターネット調査とは, 以上の仕組みを利用するための調査システムの構築を始め, 登録者集団 (リソース, パネル) の構築と登録運用管理, (電子的な) 調査票や設問の設計, サンプリング操作, 調査依頼から回収, データ処理と集計分析, 調査報告など, 調査の全過程にわたって科学的な調査方法論を適用して行う総合的な活動」をいう. ただし, 今後の技術的な進歩改善につれて, 新たな電子的な調査取得方式が登場することが考えられる. 例えば, 高機能の携帯電話, モバイル端末, 双方向デジタル TV, IVR (Interactive Voice Response) 等, 様々な調査方式やそれに関わる機器類が現れることも考えられるので, 現時点における定義とする. 本研究は Web 調査を対象としている.

2.2.2 Web 調査の特徴

大隅(2003), Couper(2003)は Web 調査の特徴を以下の様にまとめている.

- 電子的にデータ取得を調査方式, 電子的な自記式調査である
- 画面設計 (電子調査票の設問レイアウト) の自由度があること
- 双方向性 (interactive) があること
- 面接調査などと違って面接員の誘導, 対応の影響が排除できること
- 画面上に回答に戸惑ったときのガイドや警告を示せること
- 画像 (静止画, 動画), 音声などマルチメディア対応とできること
- 従って, 調査票のデザイン, 修飾が豊富であること
- これらはプログラミングとして回答の制御や抑制が可能であることを意味すること
- 回答者の回答行動を電子的に追跡できること (トラッキング, ロギング) 等々が挙げられる.

こうした諸特性は, 従来の調査法にはない新たな特長であり利点であると同時に, 短所・欠点も併せ持つので利用のあり方について考えていく必要がある.

2.2.3 インターネット調査の諸特性

インターネット調査の利点, 欠点, 特性は様々ある. 詳しくは社会調査ハンドブック(大隅,2002a)を参照されたい. ここでは, 本研究を理解する上で, 必要になると思われる事項のみを社会ハンドブック(大隅,2002a)より抜粋して紹介する.

■ 回答者の捕捉方法

インターネット調査のデータ収集方式はセルフセレクション(自記式)である. これは回答者が自主的に手を挙げて答える方法である. 多くは統計的な意味でのサンプリング操作を伴わない. そのため, 回答は抽出元である登録者集団(リソース/パネル: 後述)に依存し, 登録者集団の構築方法ならびにその特性の評価に依存する.

■ インターネット・ユーザーとインターネット回答者

従来型の調査であれば, 母集団を明確に定義し, そこから適切な標本抽出を行

い、調査対象者を決め、そこから得た回答者の関係を分析することで、母集団の推論をする。しかし、インターネット調査には困難がある。インターネット・ユーザーの特性が分からないことから、統計理論がそのまま適用できず、日本人はもとより、一般生活者を代表しているとはいえない。このことはインターネット調査が調査目的あるいは用途が限定されること、インターネット利用者を対象とした条件付の調査法として利用するものであることを意味する。

■ 母集団をどのように考えるか

従来の統計理論(標本調査理論)に即して考えると、インターネット調査の対象者はインターネット・ユーザーであるので、考えられる母集団にはインターネット・ユーザー全体ということになる。しかし、インターネット・ユーザーの活動方法は流動的であること(例：プロバイダーの乗換えや利用場所、利用コンピュータを変えることは日常的)やインターネット・ユーザー数(各調査会社が示す統計指標が1千万人以上も異なる場合がある)が正確に定まらないことなどから、インターネット・ユーザーを母集団と定義して議論することは難しく、標本誤差などが定義できないなどの問題も抱えている。このような問題を含んでいるため、インターネット調査が統計的な調査ではないと言われるゆえんかと考える。

しかし、母集団が定義できないから、科学的でないと考えるのではなく、新たな統計理論の開発が求められていると考えるべきである。

■ 登録者集団の構築

インターネット調査を行う際には、全インターネット・ユーザーが記載登録されている標本抽出枠がないために、何らかの形で登録者集団を構築しなければならない。それを吉村・大隅(2003)はリソースまたはパネルと呼んでいる。通常は勧誘・公募・自己参加・ボランティア・企業や組織などの何らかの名簿で集め、これを従来調査の標本抽出枠のように考え、調査対象をこの集団からサンプリングを行う。このとき、名簿の集め方が不明あるいは曖昧であることが多い(大隅,2003)。また、大隅らの行っている実験調査(後述)から登録者集団の構築方法がインターネット調査の回答に大きく影響を及ぼすことが実証的に分かったとしている。

2.3 実験調査

2.3.1 調査環境の変化に対応した新たな調査法の研究

国内では調査研究の例が稀であるため、インターネット調査の諸特性を測るためには用意周到な実験計画が必要である。大隅(2000)はこうした調査計画の必要性を感じ「調査環境の変化に対応した新たな調査法の研究」と題して、1997年から複数の調査機関と共同研究を行っている。この共同研究はインターネット調査の特性を理解して、どのように適応可能性を探るのみならず、得られた情報・知見を共同実験調査参加メンバーはもとより、一般にも積極的に公開をしている。本研究もこの実験調査の中で得られた情報を用いている。

2.3.2 実験調査の基本方針

文部科学省統計数理研究所と複数の調査機関による産学協同実験(研究代表者：大隅昇教授)がインターネット調査(特に、Web調査)の特性を実証的に明らかにすることを目的として、1997年から2002年にかけて計4回の実験が行われた。第1次実験調査では、単一のパネルにて約6ヶ月、計12回の調査を実施し、Web調査で起こりうる現象、検証すべき事項の把握や整理に努め、第2次、第3次実験調査では第1次実験調査の結果をふまえ、Web調査の諸特性をより明らかにするような実証研究を進めた。その際に掲げた実験調査の基本方針は以下の通りである。以下、大隅(2003)より引用。

- (1)複数の調査機関で実施する（協同研究を計画）
- (2) 実査の時期をなるべく揃える（調査時期の同時性）
- (3)従来型の複数の調査法を組み合わせ、相互比較を行う（調査法の比較）
 - ・オムニバス（訪問面接、留置自記式）、郵送など
- (4)調査票と設問内容を揃える（調査内容の共通化、整合性）
 - ・Web調査では実施機関間で同じ調査票を用いる
 - ・従来型調査も設問内容を同じとする
- (5)反復調査も場合により組み入れる
 - ・同一リソース、同一回答者に複数回実施、回答変動の観察
- (6)時系列的に比較可能なように、他の調査や過去に用いた設問を利用する（比較可能性）

(7)ただし、調査票へのアクセス方式は、なるべく各機関の現状方式にならう（調査実施環境の現場の保持）

(8)そして、調査を継続的に行うこと（継続性）

2.3.3 第4次実験調査計画の概要

2002年に行われた実験調査を第4次実験調査という。第4次実験調査までにおこなわれた第1次実験～第3次実験調査についての詳細は、大隅(2000),大隅(2001)などを引用していただきたい。

第4次実験調査では2.3.2で述べた基本方針に加え、さらに以下を新たな検証事項として加えている(吉村・大隅,2003)

(9)トラッキング(回答者行動を電子的に追跡し、関連データを取得すること)により Web 調査の特性、特に技術的な問題や事象をより詳細に把握するとともに、Web 調査の質的側面の評価の材料とすることを試みる。

(10)リソース及びパネルの構築方法の類似・差異と調査結果の関連を検証する。特にセルフセクション方式の構築方法と従来調査における調査対象者の抽出方法(例えば、住民基本台帳からの無作為抽出、マスター利用等)に基づいて、作成されたリソース構築方法との比較を焦点にあてる。

また、第4次実験調査では従来型調査と Web 調査の調査方式の比較や調査票のレイアウト実験が行われた。具体的な第4次実験調査の概要は表2.3.3である。

第4次実験調査は以下のように要約できる

- ・3つの異なるサイト Web サイトで
- ・4つの異なる登録者集団(リソース、パネル)を対象に
- ・同じ設問を「ほぼ」同じレイアウトで
- ・ほぼ同時期に
- ・2つの調査テーマでそれぞれ1回ずつ計2回実施
- ・かつ可能な限り同一の設問をほぼ同じ時期に従来型の調査でも実施

表 2.3.3 第4次実験調査の概要

WEB調査	電通リサーチ (DR)		博報堂-東京サーベイ	日本リサーチセンター (NRC)
	第1回：生活意識編			
リソース/パネル名	DENTSU_R-net	Hot Panel	e-HABIT	Cyber Panel
調査方式	リソース内 サンプリング方式	リソース内 サンプリング方式	パネル方式	リソース内 サンプリング方式
調査時期	3/28/2002-4/4/2002	6/13/2002-6/20/2002	3/28/2002-4/8/2002	4/11/2002-4/25/2002
調査票のページ分割	×	△ (半数に実施)	×	○
設問レイアウト実験	○	△ (半数に実施)	○	○
有効回収回答 (%)	939 (81.6)	3,392 (42.4)	931 (88.2)	716 (35.8)
	第2回：インターネット編			
調査時期	5/16/2002_5/23/2002	6/20/2002-6/27/2002	5/16/2002-5/27/2002	5/22/2002-6/7/2002
調査票のページ分割	×	×	×	○
有効回収回答 (%)	894 (59.1)	2,587 (32.4)	896 (84.9)	642 (32.1)
	DRPS	NOS	郵送調査	HABIT2001
オムニバス調査	第1回：生活意識編			生活意識+暮らし向き編
調査期間	3/22/2002-3/31/2002	5/7/2002/-5/15/2002	調査期間	7/4/2002-7/25/2002
方法	面接法	訪問留置自記式	方法	郵送
対象地域	東京30km圏	日本全国	対象地域	首都圏40km・近畿20km
対象者	15~59歳男女	15~79歳男女	対象者	保有パネルHABITの一部 15~69歳男女
サンプリング法	住民基本台帳からの二 段無作為抽出	住民基本台帳からの 層化多段無作為 抽出		
計画サンプル数	1,236	2,200	計画サンプル数	1,000
回収数 (%)	630 (51.0)	1,336 (60.7)	回収数 (%)	946 (94.6) *
不能票数 (%)	606 (49.0)	864 (39.3)	不能票数 (%)	54 (5.4) *
	第2回：暮らし向き編			
調査期間	5/17/2002-5/26/2002	5/7/2002/-5/15/2002		*2002/07/25現在
計画サンプル数	1,001	2,200		
回収数 (%)	630 (62.9)	1,389 (63.1)		
不能票数 (%)	371 (37.1)	811 (36.9)		

(表記法について)

- ・「リソース/パネル名」「調査方式」についての詳細は3章3.1に記載
「調査票のページ分割」「設問レイアウト実験」はレイアウト実験で用いられたときの割り付け方を示しており、詳しくは松田・大隅(2003)を参照のこと。
- ・「有効回収回答」は独自の定義のもと計数したもので、各調査機関がそれぞれ定める有効回答数と異なる場合がある。
- ・調査テーマならびに調査の設問構成は以下の様になっている

第1回 Web 調査：生活意識編

ものの考え方，政治意識（支持政党，内閣支持他），情報技術（IT）への期待感・考え方，科学技術観，日頃の暮らし方・住環境，景気観・雇用意識，情報機器接触度（利用機器，利用頻度），携帯電話の利用・情報授受等を，関連調査から引用した設問（日本人の国民性調査，電通リサーチ定例調査，環境意識調査等），ならびに独自の設問を用意した。

第2回 Web 調査：インターネット編

インターネットへの印象・知識，情報入手経路，インターネット歴，利用時間帯，インターネット活用場面，生活に占める位置づけ，所有アドレス数，情報授受の人数，利用場面と頻度，メールマガジン登録数と内容，情報開示への姿勢（開示したくない情報，提供してよい情報），インターネット上のプライバシーへの意見，各種法整備と規制への意見，インターネット調査への意見・参加頻度，調査協力要因，なりすましへの意見，本音で答えるか否か，モニターとしての登録サイト，ネットワーク社会における人格・人柄，情報技術（IT）への期待感，情報機器接触度（利用機器，利用頻度），情報媒体利用時間，携帯電話の利用・情報授受等についての設問を用意した。これらは，第2次～第3次実験調査で用いてきた設問と同じ設問も含む。

第4次実験調査の詳細については，吉村・大隅(2003),松田・大隅(2003),横原・武田・細井(2003),中谷・上嶋・渡會他(2003),鈴木・笹田(2003)を参照していただきたい。

3 章 研究の対象

3.1 Web 調査の形態

インターネット調査を行う場合、全てのインターネット・ユーザーが登録または記載されている標本抽出枠がないために、何らかの方法で登録者集団(リソース/パネル)を構築しなければいけない。インターネット調査はセルフセレクション方式で行われるために、抽出元の登録者集団に依存する。そのため、登録者集団の構築方法が異なれば、調査結果に影響がある。

ここでは、現状に行われている Web 調査の形態を整理する。国内で行われている大部分の Web 調査は調査協力者を Web 上で応募することが多く、吉村・大隅(2003),大隅(2002b)は「調査への協力者の確保」ならびに「調査実施方法」の観点から以下の分類ができるという。

① パネルタイプ

WWW 上での広告・告知によって、調査協力の意思を持つものを集め、その全員に対して、複数回の調査を継続的に行う方法。登録者は数千人程度のことが多い。

② リソースタイプ

主として、WWW 上での広告や告知によって、調査協力の意思をもつものを募って登録し、その中から実査の対象を選ぶ。登録者は数万人から十数万人規模に及び、現時点の Web 調査サービスの中心である。リソースタイプはさらに3つに分類することができる。

a) リソース内オープンタイプ

登録者を対象にバナー広告などで調査の協力を呼びかける。特定の個人への調査協力の依頼は行われない

b) 属性絞込みタイプ

調査対象を特定の性、年齢、職業などの属性で絞込み、条件を満たす該当者に調査依頼の電子メールや Web 調査票を送る方法。多くの場合、目標回答数が得られた時点で調査が打ち切られることが多い

c) リソース内サンプリングタイプ

リソースの中から、無作為に調査対象者を選び、調査依頼の電子メールを送り、続けて Web 調査票に回答を行う方法

③オープンタイプ

WWW 上に調査票を公開し、バナー広告などで調査協力を呼びかける。ここでは特定の個人への調査協力依頼は行わない。

しかし、これらの分類は 1999 年当時のものであり、最近ではこれらに該当しない登録者集団の形態が登場している。

一方、Couper(2001)は欧米の Web 調査の形態を標本の構成の仕方が確率的抽出の手続き(Probability-based methods)を経たか否か(Non-Probability methods)の視点で分類している(本章末の参考参照)。Probability-based methods はその標本抽出の手続きから母集団への推論が可能な調査であるが、Non-Probability methods ではそれができない。

大隅(2003)、吉村(2003a)は、上記に挙げた国内の分類形態に、確率的標本抽出の手続きの有無で分類する欧米の視点(Couper,2001)を加えている。つまり、下記のように分類できる。

1) パネルタイプ

Probability-based methods

Non-Probability methods

2) リソースタイプ

Probability-based methods

Non-Probability methods

リソースタイプは更に“リソース内オープンタイプ”，“属性絞込みタイプ”，“リソース内サンプリングタイプ”に分類される

3) オープンタイプ

Probability-based methods

Non-Probability methods

3.2 母集団と標本の関係

Couper(2001)によると、Probability-based methods では得られた回答から、母集団の推論が可能であると言う。そこで、本節では標本と母集団の関係を整理し、回答から何が推論できるかを考える。

3つの母集団の関係(大隅,2003 ; 吉村,2003a)

①目標母集団(target population) : 調査結果を得たいと考えている調査対象集団全体。個人,世帯,組織, などの集まりが挙げられる。例えば, 早稲田大学の工学部生や東京都に居住する 20 歳以上の男女などである

②枠母集団(frame population) : 枠となる台帳やリストのある集団。例えば, 住民基本台帳や選挙人名簿などに記載の人達がこの枠母集団に相当する。そのリストを標本抽出枠あるいはサンプリングフレームと呼ぶ。この標本抽出枠を用いて, 目標母集団の一部を抽出し, 実査の対象にする。対象になった個体を標本と呼び, 抽出時に生じた誤差を標本誤差と呼ぶ。目標母集団の完全なリストが手に入らないことが多いため, 実査の対象になるのは標本抽出枠に記載されている人のみであり, この集団は目標母集団と一致していないと考え, 目標母集団と枠母集団を区別している。また, 目標母集団と枠母集団のずれをカバレッジ誤差(Coverage error)という。

③代表母集団:(representative population): 回収標本が代表すると考えられる集団。調査完了時点で得た最終の回収標本と回収できなかった不能標本とで, 計画標本となる。一般的に回収標本と計画標本が一致してすることはあまりない。そのため, 回収標本が代表すると思われる集団を想定し, 目標母集団と枠母集団と区別する。枠母集団と代表母集団のずれを無回答誤差(Non-response error)と呼ぶ。

3.3 Web 調査の形態と標本からの推論可能性

3.2 で述べた知識を活用して、得られた調査結果からの推論可能性を登録者集団の形態から分類できる(吉村・大隅,2003). ただし, Web 調査における目標母集団の定義や標本抽出枠などの問題は考慮しない. また, オープンタイプについては議論の対象外とする.

1) パネルタイプ(Probability-based methods)

- ・ 枠母集団についての代表性をもつ

回答者→(標本誤差の考慮)→代表母集団→(無回答の考慮)→枠母集団目標母集団

2) パネルタイプ(Non-Probability methods)

- ・ 枠母集団を代表しない

回答者→(推論不可)

3) リソースタイプ(Probability-based methods)

- ・ 枠母集団についての代表性をもつ

a)リソース内オープンタイプ: 回答者→(推論不可)

b)属性絞り込みタイプ: 回答者→(推論不可)

c)リソース内サンプリングタイプ: 標本はリソースについての代表性をもつ

回答者→(標本誤差の考慮)→代表リソース→(標本誤差の考慮)→代表母集団

→(無回答の考慮)→枠母集団・目標母集団

4) リソースタイプ(Non-Probability methods)

- ・ 枠母集団を代表しない

a)リソース内オープンタイプ: 回答者→(推論不可)

b)属性絞り込みタイプ: 回答者→(推論不可)

c)リソース内サンプリングタイプ

- ・ 標本はリソースについての代表性をもつ

回答者→(標本誤差の考慮)→代表リソース→×

リソースを確率的に基づく方法で構築したとしても, リソース内サンプリングタ

イプ以外では、調査結果の適用は回答者に限定されることに注意が必要である。

実験調査では、パネルタイプにはパネル全体に、リソースタイプではリソースから無作為抽出した標本が利用された。リソースタイプの調査結果は、少なくとも、リソース全体を母集団としたときの代表母集団に一般化できるために、リソースの特徴や性質をある程度は把握できる。(吉村・大隅,2003)

登録者集団の作られ方が明示的であることは言うまでもないが、確率的な抽出を経て構築されたか、経ていないかで得られた回答から推論できる可能性が異なるということに注意が必要である。

3.4 第4次実験調査で用いた登録者集団とその構築方法

第4次実験調査では4つの異なるリソース(パネル)を対象に実査が行われた。ここで、第4次実験調査で用いられたリソース(パネル)を紹介する。

表 3.4 各リソース構築方法

リソース名	下位リソース	対象	登録者数内訳	登録者数
R-net	TMS	訪問面接調査回答者（東京30km圏、住民基本台帳）の中の応諾者	3,581	24,053
	DCAMP	郵送調査回答者（関東電波エリアの電話帳）の中の応諾者	4,444	
	TELMS	電話調査回答者（全国電話帳）の中の応諾者	2,827	
	DCP	オープン型Web調査回答者の中の応諾者	13,201	
HotPanel	Hot5	オープン型Web調査回答者の中の応諾者（2001年度）	26,803	64,600
	Hot6	オープン型Web調査回答者の中の応諾者（2002年度）	37,797	
e-HABIT		HABITモニター終了者の中の応諾者 (HABITモニター：首都圏40km、近畿圏20km、地点抽出の後割付、4000人/年)		1,055
CyberPanel		Web上で公募、NRC社HPで登録		81,770

R-net は従来調査回答者の応諾者からなるサブリソース(TMS,DCAMP,TELMS)とオープン型 Web 調査回答者の応諾者からなるサブリソース(DCP)の複数のサブリソースからなっている。Hot Panel も Cyber Panel も Web 上で集められているが、Hot Panel はオープン型 Web 調査への回答者に対象が限定されている点が Cyber Panel とは異なる。e-HABIT は首都圏及び近畿圏における地点抽出及びその後の割り付けに構成されたパネル HABIT の中のインターネット・ユーザーを募集対象としており、Web 上での募集を一切行われていない。R-net ならびに Hot Panel

については横原・武田・細井(2003)を, e-HABIT については中谷・上嶋他(2003)を Cyber Panel については鈴木・笹田(2003)を参照されたい.

以上を踏まえて, 第4次実験調査で用いられたリソースと登録者集団の形態の関係を整理すると以下の様になる.

1) パネルタイプ

Probability-based methods : e-HABIT

2) リソースタイプ

Probability-based methods : R-net(非公募系)

Non-Probability methods : R-net(公募系), Hot Panel, Cyber Panel

ここでは確率的抽出を経て構築された登録者集団を非公募系, 非確率的抽出を経て構築された登録者集団を公募系と呼ぶ. つまり, 第4次実験調査では e-HABIT と R-net の一部が非公募系であり, Hot Panel と Cyber Panel は公募系に相当する.

3.5 公募系と非公募系

3.5.1 公募系の特徴

公募系は Web 上でのセルフセレクション方式で集められた非確率的抽出を経て, 構築された登録者集団のことをいう. 現行のインターネット調査では大部分の登録者集団が公募系であり, 大規模な登録者数を抱えている場合が多い. 中には数十万規模の登録者数を抱えている所もある. 横原・武田・細井(2003)によると公募系の回答者は男性, 30代, IT系技術職の比率の高さが指摘されている. また, インターネット接触時間の長さ, 常時接続化比率の高さ, 調査回答頻度が高いことなどを挙げ, インターネット・ヘビーユース, 調査への関与が高いという. 多少の違いはあるにせよ, 他の公募系に相当するサイトでも同様の傾向が見られる.

本節では吉村・大隅(2003)が第4次実験調査の結果を踏まえて, 公募系の特徴を述べており, 適宜それを引用し, 説明する.

■ 計画標本と回収標本の関係

計画標本と回収標本を比べた場合、男女ともに回収標本の分布が計画標本よりもやや高年齢層側に偏る傾向がある(吉村・大隅,2003)。回答率も約30%~40%程度となっており高くない。このことから、回答者はリソースを代表しているとは言い難い。第2次実験調査と第3次実験調査でも公募系を用いているが、そこでも同様の傾向が見られると指摘している。つまり、低い回答率、回答標本と計画標本の年齢分布の系統的なずれが公募系の1つの特徴である。

■ 特定の設問から見られる特徴

以下、公募系に属する30代の男女を対象とした時の特定の設問に見られる特徴を記す。

(1) インターネット上での調査やアンケートへの参加頻度

第2次、第3次実験調査ではWeb上での調査やアンケートへの参加頻度が高いことが観測された。第4次実験調査でも同様の傾向が観測されている。各リソース(パネル)における週に1回以上の調査参加者の割合、及びリソースの構築法は以下の通りである。

公募系(Hot Panel) : 男性 42.5%, 女性 52.1%

公募系(Cyber Panel) : 男性 52.1%, 女性 62.6%

非公募系(e-HABIT) : 男性 9.7%, 女性 11.1%

* 数値は吉村・大隅(2003)から引用

この結果から見て分かるように、公募系の登録者は約半数近くが週に1回以上、調査に参加しており、調査に積極的な人達の集まりであることが分かる。また、従来型調査終了後に、インターネット調査への参加協力を応諾した人達であり、一般的なインターネット・ユーザーより少しは積極的だと思われる非公募系のe-HABITは1割程度にしかならない。つまり、インターネット・ユーザーではさらにこの割合が小さくなると考えられる。以上を踏まえると公募系の回答者はインターネット・ユーザーの中でもかなり限られた人であることが想像できる。

(2) 登録者サイトと調査参加頻度

上で示した通り、公募系は調査に対して積極的な人が多い集団(プロ回答者； professional Respondents)であることが分かった(吉村・大隅,2003)。第4次実験調査ではインターネット調査の回答に密接に関係があると思われる登録しているサイト状況を調べるために、46のサイト数に対して登録しているか否かの質問を行った。各リソースにおける平均登録サイト数ならびに上位10サイトの登録率は以下のとおりである。

公募系(Hot Panel) : 平均登録サイト数=3.1, 登録率= 約 8.0%~20.0%

公募系(Cyber Panel) : 平均登録サイト数=5.1, 登録率= 約 17.0%~34.7%

非公募系(e-HABIT) : 平均登録サイト数=1.3, 登録率= 約 0.9%~3.5%

* 数値は吉村・大隅(2003)から引用

調査参加常習者の割合が多いリソースほど、登録している他サイトの数が多くなる(平均登録サイト数が多い)。また、「ほぼ毎日」調査やアンケートに参加している者に着目した場合、これは公募系、非公募系に限らず、どのリソースにおいても同じような特徴を示すという(吉村,2003b)。これらのことをあわせると公募系と非公募系でも、プロ回答者が存在し、異なる点は回答者のその占める割合であり、リソースの構築法によって「インターネット・ユーザー」からの対象者の切り取り方が大きく異なることを意味している。

上位登録サイトの相互登録率が高いことも踏まえると、公募系による「インターネット調査」が対象としている集団は想像以上に限定されている可能性がある(吉村・大隅,2003)。

3.5.2 非公募系の特徴

非公募系は標本抽出枠が明らかな母集団から確率的な抽出を経て作られた標本に基づいて構築されている。第4次実験調査では R-net の一部と e-HABIT が該当する。しかし、その標本はインターネット・ユーザーに限定されること、さらに抽出対象となっても調査への応諾を得た者に限られるなどの条件付の標本となっている。一般に応諾率は高いとはいえないので、公募系と比べて多数の登録者を確保するための経費面の負担が大きい。また、調査疲労を少しでも遅らせるために公募系よりも手厚いモニター管理が必要である。

公募系と非公募系は、その構築方法に差異があるが、この両者を比べる限りは非公募系がインターネット・ユーザーの集まりにより近いと考えられるが、しかし非公募系がインターネット・ユーザーの縮図になっているとはいえない。

3.6 研究の対象と適用可能性

調査経費の観点から考えると公募系の利用頻度を多くし、非公募系の利用頻度を少なくすることが望まれる。しかし、公募系と非公募系では回答者の特性が異なるために、代用はできない。ここで、調査実施回数は少ないが、構築過程が明らかで調査結果が信頼できると思われる非公募系の標本構成に基づいて、公募系の結果を修正できるかを検討する。言い換えれば、加重修正法の適用が可能であるとしたときに、利用頻度や登録標本数が多い公募系の結果をあたかも非公募系の結果と同じようにして、用いることができるはずだと考える。もちろん、事前に様々な設定条件が必要であったとしても利用価値は十分にある。

具体的には、ある特定の項目の回答(例 ; 「あなたはデジタルカメラを持っていますか」)に対して、回答者の性質のみが異なるとした公募系の回答結果に対して、加重修正法を適用する。そして、仮に公募系の回答を非公募系が回答した場合の反応変数の挙動(平均や分散)を調べ、推論する。

ここで、具体的に説明をする。

公募系、非公募系に対して同一の調査を行い、その中の特定の項目(例 ; 「あなたはデジタルカメラを持っていますか」)に差異が確認されたとする。これを反応変数 Y (公募系の反応変数を Y_1 , 非公募系の反応変数を Y_0) と呼ぶ。このとき、反応変数 Y に影響を及ぼすと考えられる性別や年齢区分や収入などの共変量 X_1, X_0 (前者が公募系の共変量ベクトル) が観測されており、この分布が偏っているとする。この偏りが反応変数に影響を及ぼしていると考えられるので、調整を行う必要がある。これを共変量調整と呼び、本研究では傾向スコアによる加重修正法を用いることで、調整を図る。再び、この同一の公募系の回答者に対して調査を行ったとする。このとき、非公募系への調査は行われぬ。そのとき、先の調整で用いられた傾向スコアと再び行われた調査で得られた反応変数を用いて、仮に公募系の回答を非公募系の登録者が回答した場合にどのような反応変数の挙動を示すかを推定する。ただし、調査 2 で推定される非公募系は調査 1 の非公募系と同質と

みなせることが条件である。

ここで扱う研究の対象は調査1について実験を行い、修正効果を測ることに努める。

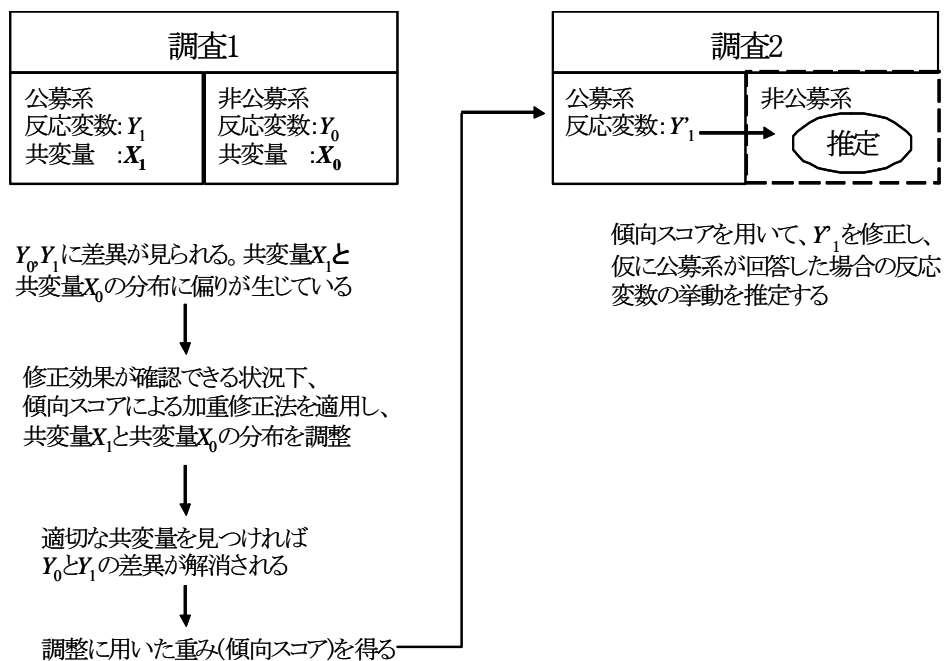


図 3.6 適用可能性

参考

欧米での Web 調査の形態の分類(Couper,2001;吉村,2003;大隅,2003)

* 確率的抽出(Probability-based methods)

① Intercept Surveys :

ターゲット母集団をそのサイトへの訪問者とし、 N 番目ごとに調査への参加を依頼する。

② List-based Samples :

E-mail のリストが作成できるような集団 (会社や学校) をターゲット母集団とし、リストから標本を無作為抽出する。

③ Web option in mixed-mode surveys :

例えば電話調査で Web 上での回答も可能であるという選択肢を用意する。

④Pre-recruited panels of Internet users :

インターネット・ユーザーを母集団とし、標本を確率抽出によるパネルを構成する。

⑤Pre-recruited panels of full population :

より一般的な母集団から標本を確率抽出しパネルを構成する。

*非確率的抽出(Non-probability methods)

⑥Polls as entertainment :

インターネット上の投票など、娯楽を目的とするもの。

⑦Unrestricted self-selected surveys :

ポータルサイトなどで行われるオープン型の Web 調査。

⑧Volunteer opt-in panels :

典型的な Web 調査。公募型パネルやオプトインパネルによる Web 調査。現行の Web 調査の大部分を占める。

4章 傾向スコアによる加重修正法

3章までは Web 調査全体を俯瞰し、Web 調査の回答者はインターネット・ユーザーの一部であり、登録者集団の構築方法により属性分布や回答分布の偏りが生じるという問題を指摘した。本章ではこの偏りへの対処法の一つである加重修正法をとりあげる。特に、米国の Harris Interactive 社による 2000 年米国大統領選挙の予測に Web 調査の結果に傾向スコア加重修正法を用いた例が紹介され、注目されている。

4.1 実験研究と観察研究

処置を与えた群と処置を与えない群の 2 群を比較して、処置の要因を測る観測値に差異があったとする。このとき、処置のことを処遇(treatment)と呼び、観測値を反応変数と呼ぶ。また、2 群の反応変数の差を処遇効果といい、特に 2 群の間の反応変数の平均の差を平均処遇効果という。例として、交通事故発生時に「シートベルトの着用有無」が「事故の被害度」と関係するかという問題を考えた時に、「シートベルトの着用有」が処遇有、「シートベルトの着用無し」が処遇なし、事故の被害の大きさ(程度)が反応変数となる(Rubin,1973)

ここで処遇の有無によって、反応変数に差異が生じるか、つまり処遇効果があるかに関心があり、反応変数に関する共変量 (または処遇前に得られた測定値) が観測されているとする。

用意した調査対象となる 2 群に対して、事前に処遇を無作為に割り付けることができれば、反応変数に影響を与えうる共変量(観測の有無に関らず)の分布は両群で確率的に等しいと考えることができるので、共変量の影響を無視することができる。群間で反応変数の差異を確認することができれば、処遇効果があったといえる。このように、研究者が事前に処遇の割付を行う研究を実験研究という。

ところが、研究者が事前に処遇の割付を行うことができない場合がある。何らかの処遇を伴うが、研究者自身で無作為に割り付けが行えない研究を実験研究と対比させて、観察研究と呼んでいる(Cochran,1965)。観察研究では反応変数に差異が生じたとしても、それは共変量に影響していることも考えられるために、即座に処遇効果があるとは言えない。

先のシートベルトの着用有無を例で説明すると、

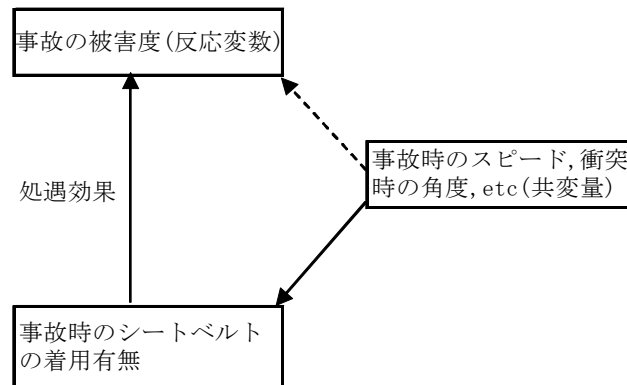


図 4.1 事故当時のシートベルトの着用効果

「事故当時の被害度」が大きいときは、「シートベルトを着用無し」だけでなく、「衝突時のスピード」や「衝突時の角度」などの他の要因も影響していると考えられる。そのため、これらの要因を無視して、「シートベルトの着用無し」が「事故当時の被害を大きくする」と結論づけることはできない。

上記の例では事前にシートベルトの着用を割り付けることができない、つまり観察研究では、処遇の異なる群間で、共変量の分布が等しいとは言えないので、共変量の影響を無視できず、何らかの方法で共変量の影響を除去しないと行けない。よく用いられる共変量調整の方法にマッチングがある。これは共変量の値が同じ個体を各群から選び、それらの反応変数を比較する方法である。例えば、共変量が性別のみであれば、男性同士、女子同士のペアを作って、反応変数を処遇間で比較して、処遇効果があるかを確認する。しかし、多数の共変量を用いると該当するサンプルが存在しないなどの問題がある。

4.2 傾向スコア(propensity score)

傾向スコアは Rosenbaum & Rubin(1983)によって観察研究において平均処遇効果を推定するために提案されたバランシングスコアの1つである。バランシングスコア $b(\mathbf{X})$ とは共変量ベクトル \mathbf{X} の関数で、これを用いることで比較したい群間の共変量ベクトル \mathbf{X} の分布のバランスをとることができる。つまり、バランシングスコアを所与として、共変量の分布

が等しくなる性質がある.

傾向スコアはバランシングスコアの1つであり, 観測された共変量を所与としたときの処遇が割り付けられる確率と定義される.

ここでは公募系と非公募系からなる2つの登録者集団を考え, 公募系と非公募系に対して, 同一質問の調査(例:「あなたはデジタルカメラを持っていますか」)を行う. そのとき, 公募系,非公募系からそれぞれ n_1, n_0 の回答を得たものとする ($N=n_0+n_1$). 各 $i=1,2,\dots,N$ について以下の変数を定義する. (表 4.2 参照)

Y_{1i} :公募系に属する回答者 i の質問に対する回答を数量化したもの

(例:「はい」と答えれば1,「いいえ」と答えれば0)

Y_{0i} :非公募系に属する回答者 i の質問に対する回答を数量化したもの

(例:「はい」ならば1, 他は0)

Y を「反応変数」と呼ぶ.

\mathbf{X}_i :回答者 i の共変量ベクトル(性別, 年齢, 年収...)

Z_i :回答者 i が公募系に属するならば1,非公募系に属すれば0

Z を「処遇変数」と呼ぶ

$e(\mathbf{X}_i)$:傾向スコア

ただし, $z_i, y_{1i}, y_{0i}, \mathbf{x}_i$ はそれぞれ対応する大文字の観測データを表すことにする.

$$e(\mathbf{x}_i) = \Pr(Z_i = 1 | \mathbf{X}_i = \mathbf{x}_i) \quad (i=1,2,\dots,N) \quad (1)$$

回答者 i の共変量ベクトル \mathbf{x}_i を所与として, その回答者 i が公募系と判断される条件付確率が傾向スコアである.

傾向スコアの推定値は, 処遇変数を予測変数, 観測された共変量ベクトルを説明変数として, ロジスティック回帰を用いられることが多い (Rosenbaum & Rubin, 1983; D'Agostino, 1998).

表 4.2 : データの構造

	Y		Z	X			
	Y_1	Y_0		X_1	X_2	\dots	X_M
1	Y_{11}	0	1	X_{11}	X_{21}	\dots	X_{M1}
2	Y_{12}		1	X_{12}	X_{22}	\dots	X_{M2}
\vdots	\vdots		\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots		\vdots	\vdots	\vdots	\vdots	\vdots
i	Y_{1i}		1	X_{1i}	X_{2i}	\dots	X_{Mi}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	
k	0	Y_{0k}	0	X_{1k}	X_{2k}	\dots	X_{Mk}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
N		Y_{0N}	0	X_{1N}	X_{2N}	\dots	X_{MN}

4.3 strongly ignorable の仮定

\mathbf{x} を観測された Y の共変量ベクトルとして, 以下が成り立つとき, 処遇は strongly ignorable という.

$$Y \perp\!\!\!\perp Z \mid \mathbf{x}, 0 < pr(Z = 1 \mid \mathbf{x}) < 1, \text{ for all } \mathbf{x} \quad (2)$$

注 1) $A \perp\!\!\!\perp B \mid C$ は, C を所与としたときに A と B が独立であることを示す.

strongly ignorable の仮定とは, \mathbf{x} を所与として Y と Z が独立であることの仮定である.

事前に処遇の割付ができない観察研究においても, この仮定をおいた下では, マッチング, 層別, 共分散調整などを使って, 平均処遇効果の推定値が求められることができる (Rosenbaum & Rubin, 1983). Y に関する共変量がすべて観測されることが必要であり, もし観測されていない共変量の存在があれば, strongly ignorable の仮定は成り立たない.

Rosenbaum (1984), 吉村 (2003a) によると strongly ignorable の仮定が成り立っているか確認するためには, 以下の手順を踏めばよい.

- ・処遇の影響を受けない変数を用意する.
- ・傾向スコアでブレイクダウンをする(または共変量でバランスをとる)
- ・比較する. 差が小さければ **strongly ignorable** の仮定は成り立っており, もし差が大きければ観測されなかった共変量の存在を検討する. つまり, 傾向スコアの再推定が必要である.

4.4 傾向スコアによる加重修正法

2つの登録者集団の共変量の分布(例えば, 性別や年齢の分布など)が異なっているために, その違いが回答に影響を及ぼし, 推定値が偏る場合がある. その問題を解決するために, 傾向スコアを用いて, 公募系の共変量の分布を非公募系のそれにあわせるような共変量調整を行い, その調整に基づいて, 推定値の修正を行う.

修正前の推定量

$$\text{非公募系の反応変数の平均} \quad \bar{Y}_0 = \frac{1}{n_0} \sum_{i=1}^N (1 - Z_i) Y_{0i} \quad (3)$$

$$\text{公募系の反応変数の平均} \quad \bar{Y}_1 = \frac{1}{n_1} \sum_{i=1}^N Z_i Y_{1i} \quad (4)$$

加重修正法 1(Rosenbaum,1987)

$$\bar{Y}'_1 = \frac{1}{N} \sum_{i=1}^N \frac{Z_i Y_{1i}}{e(\mathbf{x}_i)} \quad (5)$$

ignorable selection(Rosenbaum,1987)の仮定のもとで, 公募系と非公募系をあわせた集団に相当する Y の値の推定量を(5)式で求めた. (5)式を傾向スコアによる加重修正法という.

ところで, (5)式の \bar{Y}'_1 は公募系と非公募系をあわせた集団における推定量であり, 本研究に必要なのは非公募系の結果を基に公募系の結果を修正した推定量である.

Hahn(1998)は Y_i を Y_{1i} と Y_{0i} を用いて

$$Y_i = Z_i Y_{1i} + (1 - Z_i) Y_{0i} \quad (6)$$

と表現している. ここで, 非公募系の結果を基に, 公募系の結果を修正した推定

量 \bar{Y}_1'' を考える．これは非公募系における n_0 個の Y_0 の標本平均として得られる．

$$\begin{aligned}\bar{Y}_1'' &= \frac{1}{n_0} \sum_{i=1}^{n_0} Y_{0i} \\ &= \frac{1}{n_0} \left\{ \sum_{i=1}^N \frac{Z_i Y_{1i}}{e(\mathbf{x}_i)} - \sum_{i=1}^N Z_i Y_{1i} \right\} \\ &= \frac{1}{n_0} \sum_{i=1}^N \frac{Z_i Y_{1i} (1 - e(\mathbf{x}_i))}{e(\mathbf{x}_i)}\end{aligned}\quad (7)$$

(7)式を用いて以後議論を行う．また， Y_0 の分散の推定量 \hat{v} は以下の様に表現される．

$$\begin{aligned}\hat{v} &= \frac{1}{n_0 - 1} \sum_{i=1}^{n_0} (Y_{0i} - \bar{Y}_1'')^2 \\ &= \frac{1}{n_0 - 1} \left[\left\{ \sum_{i=1}^N \frac{Z_i Y_{1i}^2 (1 - e(\mathbf{x}_i))}{e(\mathbf{x}_i)} \right\} - n_0 (\bar{Y}_1'')^2 \right]\end{aligned}\quad (8)$$

4.5 最適化問題

共変量ベクトル X の部分集合をうまく取り出すと，修正の共変量の選び方によって，修正の程度が異なる．そのような中で共変量ベクトルの部分集合をうまく取り出すことが必要である． $\mathbf{x} = (X_1, X_2, \dots, X_M)$ ， $I = \{1, 2, \dots, M\}$ としたとき， I の部分集合 $J \subset I$ を使い，加重修正法を適用する．ただし，集合 I の中に反応変数に関する共変量が全て観測されているものとする． $\mathbf{X}(J) = (X_j)_{j \in J}$ を共変量ベクトルとしたときの(7)式に対応する推定値を以下のように書く．

$$\bar{y}_1''(J) = \frac{1}{n_0} \sum_{i=1}^N \frac{z_i y_{1i}}{\hat{e}(\mathbf{x}_i(J))} (1 - \hat{e}(\mathbf{x}_i(J)))\quad (9)$$

ただし， z_i, y_{1i}, y_{0i} ， $\mathbf{x}_i(J)$ はそれぞれ対応する大文字の観測データを表し， $\hat{e}(\mathbf{x}_i(J))$ はロジスティック回帰を使って，(10)式で求められる

$$\hat{e}(x_i(J)) = \frac{1}{1 + \exp(-\hat{\beta}_0 - \hat{\beta}_1 x_i(J))} \quad (10)$$

また、 \bar{Y}_0 に対応する観測値を \bar{y}_0 としたときに、

$$\bar{y}_0 = \frac{1}{n_0} \sum_{i=1}^N (1 - z_i) y_{0i} \quad (11)$$

部分集合 J をうまく選ぶと、 $\bar{y}_1''(J)$ と \bar{y}_0 の差を小さくすることが期待される。

つまり、

$$\text{目的関数：} \min_{J \subset I} |\bar{y}_1''(J) - \bar{y}_0| \quad (12)$$

となる J の組み合わせを考えることが目的である。

4.6 傾向スコアによる加重修正法の適用例

4.6.1 Harris Interactive 社による適用例

アメリカの米国大手市場調査会社である Harris Interactive 社は 2000 年 11 月に行われた大統領選挙において、Web 調査で得られた結果に対して傾向スコアによる加重修正法を適用した調整(修正)を行い、選挙前に G.ブッシュと A.ゴアの得票率が同じであることを予測した。他にも同様の予測を行い、成功を収めている。しかし、具体的な方法が記載されておらず、追試ができないという問題を抱えている。星野(2003)は、Harris Interactive 社の傾向スコアによる調整法を次のように推測している。

■ Harris Interactive 社の傾向スコアによる調整法の手続き

- (1) Web 調査のための調査協力者パネル(前述のパネルより登録者集団という意味合いで使われている)を大量に用意する。2002 年現在で Harris Interactive 社の場合、700 万人 (Harris Poll Online Panel)である。
- (2) 傾向スコア算出を算出するための無作為抽出による調査(電話調査)を行う。無作為抽出による調査では、傾向スコアを算出するための、Web 調査と共通の調査項目を測定する。この項目群は属性変数や態度・行動を測定する心理・社会・医療などの質問項目があるとうまくいく場合が多い。調査の目的に応じて変化させる必要があり、数種類は容易している。

- (3) 本調査としての Web 調査を行う。調査協力者は、パネルの中から無作為抽出または何らかの属性変数上の比率に従うようなサンプリング(標本抽出)がされる。本調査の中には調査目的に直接関係する項目だけではなく、傾向スコアを計算するための共通の項目も含まれる。
- (4) 無作為抽出のデータと Web 調査のデータを用いて傾向スコアを算出する
- (5) 傾向スコアを用いた加重修正法により、Web 調査の結果を調整する。

4.6.2 従来型調査の結果と Web 調査の調査結果間の修正

星野(2003)は Web 調査の調査結果を訪問留置調査のそれに近づけるために、つまり Web 調査の回答者が従来調査方式で回答した場合を推定するために、傾向スコア加重修正法を適用した。

4.6.3 適用例の問題点

Taylor 他 (2001) は、電話調査と Web 調査の間で行った傾向スコアによる修正効果の意味があったとしている。星野(2003)も訪問留置調査と Web 調査の間で修正を行い、修正効果を確認したとしている。しかし、面接や郵送調査などの従来型調査方式とインターネット調査との比較実験を続けてきた大隅らの研究成果に照らし合わせると、異なる調査方式間の修正処理がそう簡単に行えるとは思えない。例えば、吉村・松田・大隅他 (2003) が、傾向スコアによる加重修正法の特性を慎重に吟味したところ、異なる調査方式の差異までは修正が困難であると指摘している。このように加重修正法を適用する状況や前提が異なる中で、修正効果を一括りにして議論することは難しい。本研究はインターネット調査の諸特性を考慮しながら、傾向スコアによる加重修正法による修正効果の測定を試みる。

5 章 実験の方法

5.1 2つの調査結果に与える要因

一般に、2つの調査結果、反応変数 Y の間に差異が見られるときは複数の要因が相互に影響を及ぼしていることが予想される。(図 5.1)

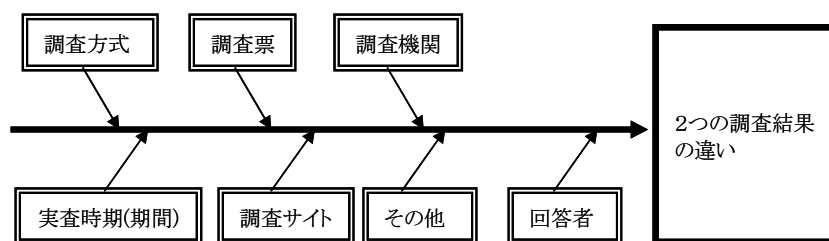


図 5.1 特性要因図

- ・ 調査方式
- ・ 回答者の偏り
- ・ 調査時期
- ・ 調査サイト(調査機関)
- ・ 調査票のデザイン
- ・ 回答形式
- ・ その他

例えば、調査票のデザインが回答に及ぼす影響がある(松田・大隅,2003;)。Web 調査の特徴の一つは、設問形式の多様性にあるとされている。すなわち、調査票の作成時に、様々な様式の設問形式を取り入れることが可能であるとされてきた。しかし、調査票のデザインや回答形式の影響は未知であり、無視することはできない。調査票設計を例に挙げると、松田・大隅(2003)は第4次実験調査で行われたレイアウト実験の結果を用いて、同一設問に対して、初期状態で選択肢が全て閲覧できるラジオボタン形式と一部が閲覧できないコンボボックス形式と一切閲覧できないプルダウン・メニュー形式の3つの異なる設問形式の回答パターンの比較を行った時に、ラジオボタン形式と他2つの形式で回答パターンに差異があることを確認し、デザイン効果が回答に及ぼす影響が無視できないことを実証的に示した。

5.2 適用条件

先の適用例で示したように、それぞれが適用する状況や前提が異なっており、それらを一括りにして修正効果の有無について議論することは難しい。Web 調査の特性ならびに調査結果に与える要因を踏まえた上で傾向スコアによる加重修正法を適用し、その修正の効果を測るためには、どのようなデータを用いるべきか整理する必要がある。

傾向スコアによる加重修正法の適用範囲は観測された共変量だけであり、観測されない共変量や処遇効果(ここでは、調査方式の差異による影響、調査時期、調査サイトなど)は修正することができない。そのため、2つの調査結果の差異が回答者の性質のみに起因すると考えられる場合以外の要因が制御された調査環境下で修正を試みるべきである。

以上より、Web 調査の結果に対して加重修正法による修正効果を測定するためには少なくとも次の要件を満たすことが必要だと思われる。

- 登録者集団（リソース/パネル）が計画的にまた明示的に設計されている
- 調査方式が同一の調査モードである(ここでは Web 調査で統一)
- 同一調査票の利用、調査時期が同じ（ほぼ同時期に実施）など、ほとんど同じ調査環境下で実施
- 2つの調査結果の差異が回答者の特性や性質だけに起因すると考えられること

実は加重修正を行うか否かに関わりなく、Web 調査の結果を科学的に考察するためには、データ取得環境がいかなる条件で行われたかを知っておくことは肝要である。これらの条件が満たされた環境下で取得したデータを用いるなら、傾向スコアによる加重修正法を適用した結果から、標本構成からのアンバランスから生じたものであるかどうかを類推できると考えられる

5.3 適用データの特徴

■ 登録者集団の性質

データの構成は表 5.3.1 にあるように電通リサーチ社の用いている「DENTSU_R-net」サイトで取得されたデータである。電通リサーチ社では、2種類の登録者の公募方法を使い分けており、これに対応して大別して2種の登録者集団（R ネット・モニター、HOT Panel モニター）がある。ここでは、R ネット・モニターを用いるが、これの構成は表 5.3.1 にあるように、さらに「公募系」と「非公募系」（注：電通リサーチ社では前者をインターネット公募系、後者をサンプリング系と呼称）がある。この違いは表 5.3.1 にある通りであるが、大きな違いは、公募系は完全なオープン型 Web 調査で公募・登録した標本のうち、初期の頃に集めた集団（DCP 系）に含まれる標本、非公募型は従来型調査（表にある3種；TMS, DCAMP, TelMS）の標本抽出枠に基づいて構成された標本である。また、いずれも、対象者の調査参加への「応諾」を取ったうえで登録者とするという合意形成を行っていることが特徴である。尚、R-net についての詳細は横原・武田・細井(2003)に記されている。

表 5.3.1 登録者集団の構築方法(本研究対象)

運用サイト名	類型区分	名称	総登録者数(名)	標本の出处(標本抽出枠)	調査方式	登録者居住地域
R ネット・モニター	非公募系	TMS	3,581	住民基本台帳	訪問面接法	首都圏
		DCAMP	4,444	電話帳	郵送 ⁶ 札法	
		TelMS	2,827	電話帳	電話調査法	全国
HOT Panel モニター	公募系	DCP('98-99)	13,201	オープン方式 (約200のウェブサイトを対象としたオープン型調査から収集)		全国
		HOT5('00)	26,803			
		HOT6('01)	37,797			

■ 解析対象

第4次実験調査で取得された「インターネット編」と「生活意識編」のデータを用いる。詳細は吉村・大隅(2003)を参考。調査票の詳細については、2章を参照のこと。

分析に用いた有効解析対象者数は以下である。

表 5.3.2 有効解析対象者数(生活意識編)

	計画標本数	有効回収回答数	有効解析対象者数
公募系	840	536	484
非公募系	702	403	361

表 5.3.3 有効解析対象者数(インターネット編)

	計画標本数	有効回収回答数	有効解析対象者数
公募系	849	502	464
非公募系	693	392	353

インターネット調査では、画面上でみる調査票の形式が、従来の質問紙形式と似ていても、その裏側では CGI などを利用して電子的に回答の取得を制御する仕組みがあることが多い。第 4 次実験調査ではこれらの影響が回答に与えることを懸念して回答の制御や誘導を行っていない。そのため、加重修正を行う上で不都合になる一部のサンプルを解析対象から外した。

Rosenbaum & Rubin(1987)によると、共変量に無回答がある場合、傾向スコアを算出することができない。そのため、共変量に相当する変数に無回答があるサンプルを解析対象から外した。また、回答に論理矛盾があるようなサンプルも解析対象から外すことにした。具体的な基準は以下の通りである。

～サンプルの除外基準～

- ① 共変量として用いた属性設問に無回答があるサンプル
- ② 登録時と回収時で論理的な矛盾がある
 - (ア)登録時と回収時で性別が異なる場合
 - (イ)登録時の年齢と回収時の年齢に隔たりが大きい場合
 - (ウ)登録時に既婚、回収時に未婚(論理矛盾)と回答する場合
 - (エ)(インターネット編)登録時と回収時のインターネット歴に隔たりが大きい場合
 - (オ)世帯年収・家族構成・回答時間など、総合的に判断して解析対象にふさわしくない場合

③ 設問間の論理的な矛盾がある

例 1) 未既婚「未婚」かつ家族構成「夫婦のみ」

例 2) (インターネット編のみ)インターネットの利用可能場所と調査回答時の場所が一致しない など

■ 本実験で用いる適用データの特徴

取得環境としては「5.2 適用条件」に記した要件を満たすこととする。このため、本実験の公募系と非公募系の反応変数を比較したときに、その差異は回答者の性質によるものだと考えられる(図 5.3.2)。回答者の性質の違いによる影響は傾向スコアによる加重修正法の適用範囲であるので、仮に共変量の影響を的確に除去できたならば、反応変数の差異が解消されることが期待される。言い換えれば、公募系の結果を修正した結果と、修正を行わない非公募系の結果が等しくなることが期待される。これは 4 章の(12)式が 0 に近づけることを意味する。

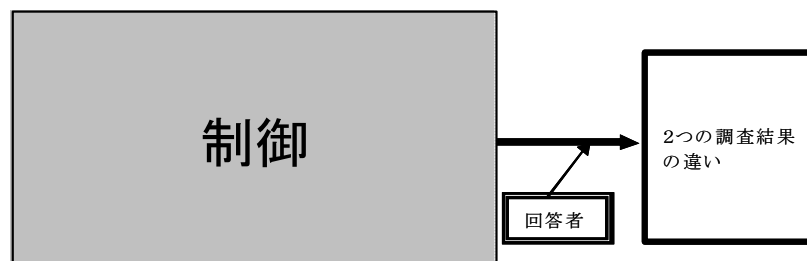


図 5.3.2 回答者以外の要因を制御した上での特性要因図

5.4 実験の目的

修正効果が測定できるデータを用いて、修正効果に関する以下の仮説を立て、それらを検証する。

- ① 修正の効果が本当にあるのか
- ② 同一の共変量を用いた時に反応変数に対する影響があるのか
- ③ 全ての反応変数に対して、修正効果がある共変量があるのか
- ④ 共変量の組み合わせ(共変量セット)による影響があるのか

- ⑤ 修正ができなければ、どのようにすれば修正をすることが可能であるか
- ⑥ 異なるデータセットを用いた場合、何を意味するのか(注 1)

注1) 加重修正法の対象となる設問・選択肢(反応変数,共変量)は同一であり,他の設問の調査内容や調査テーマが異なるデータセット,ここでは生活意識編のデータセットを用いる.

5.5 実験の手順

傾向スコアによる加重修正法は回答者の偏りを解消できる方法であるが,前提条件や使用状況が不明な中で修正効果を論ずることに問題があった.本節では修正効果を測定できるように考慮した修正過程を詳細に説明する.

手順 1: 共変量の形式の決定

まず共変量の候補とするいくつかの説明変数を選ぶ.設問が多項選択肢型の場合(例:年齢区分),選択肢をそのまま用いる場合と,事前に2値化してから用いる場合がある.

前者は,選択肢別の頻度分布をみて,必要に応じて選択肢の併合(プーリング)を行い,新たな変数を作り共変量として用いる.プーリングする目安をつけるために数量化Ⅲ類(対応分析)などを用いるとよい.本実験でも数量化Ⅲ類を用いて,一部の変数をプーリングした.1変数に対して複数の共変量からなる変数を多値変数と呼ぶことにする.

2値化処理については,例えば「性別(男性)」は「男性」であれば1,「男性以外」であれば0を対応させて共変量として用いる.年齢区分を例にした処理過程が図 5.5 である.このように1変数あたり0または1で構成される変数を多値変数と区別して2値変数と呼ぶことにする.

手順 2: 共変量セット(組み合わせ)の構築

ここでは事前に適切な共変量の組み合わせが分かっていないので,選択候補の共変量を適宜組み合わせた共変量の集合,共変量セットを用意し,これを用いて実験を行う.例えば

- ・ 共変量セット 1 (3変数): 性別, 年齢区分, 未既婚

- 共変量セット 2 (3 変数) : 性別, 年齢区分, 職業
- 共変量セット 3 (7 変数) : 性別, 年齢区分, 職業, 家族構成, 住居形態, 世帯年収, インターネット歴
- 共変量セット 4 (9 変数) : 共変量セット 3 に回答場所, 通信回線の 2 変数を追加

* 各変数は 1 変数あたり複数の選択肢から構成されている(先の例でいうところの年齢区分). の内訳については, 表 5.5 を参照.

共変量セット構築の方針として, 共変量セット 1,2 は, 基本属性のみで構成されており, 共変量セット 3 と 4 は基本属性の変数に加え, インターネット利用形態等に関わる変数を一部追加した. つまり, 共変量を基本属性だけとした場合と, インターネット歴や利用環境などを加えた場合とで, 反応変数の修正の程度の異なり具合を確認する.

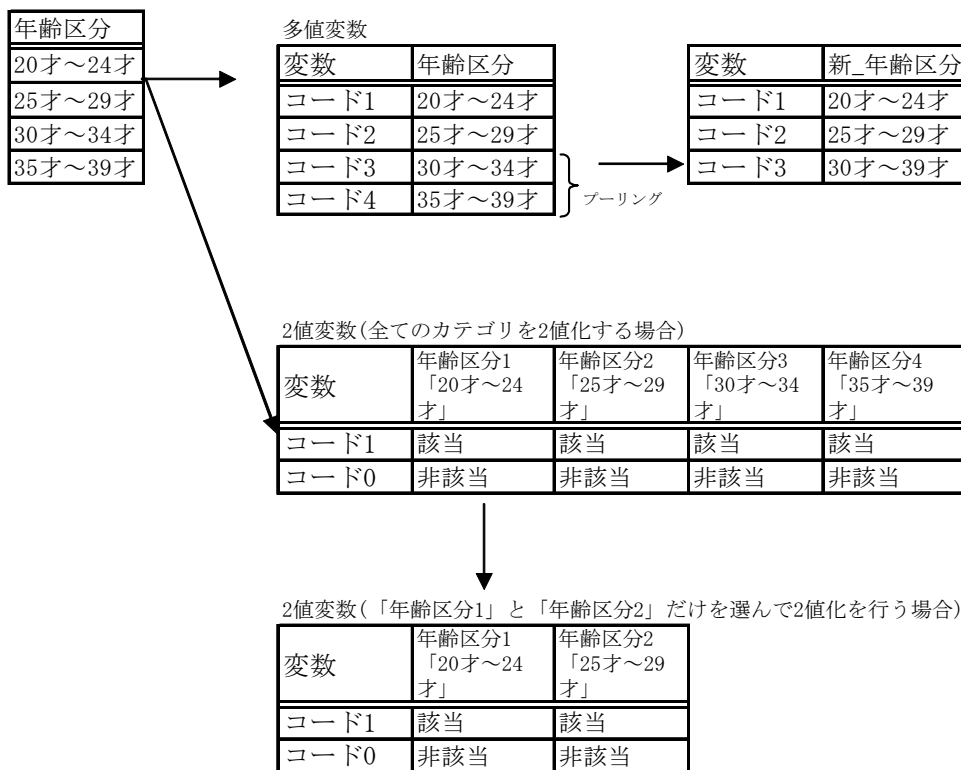


図 5.5 共変量の形式(多値変数と 2 値変数)

表 5.5 共変量セット

共変量セットと共変量の組合せ

	性別	年齢区分	未既婚	職業	家族構成	住居形態	世帯年収	回答場所	通信回線	インターネット歴
選択肢数	2	12	2	21	6	8	15	4	9	12
共変量セット1	○	○	○							
共変量セット2	○	○		○						
共変量セット3	○	○		○	○	○	○			○
共変量セット4	○	○		○	○	○	○	○	○	○

各共変量の内訳

1	男性	20代後半未満	既婚	労務職	単身世帯	持ち家	200万円未満	自宅	常時接続 (ケーブル含む)	2年未満
2	女性	30代前半	未婚	技能職	夫婦のみ	非持ち家	200~300万円未満	自宅外	非常時接続	2年以上3年未満
3		30代後半		販売・保安・サービス職	夫婦他		300~400万円未満		専用回線	3年以上4年未満
4		40代前半		営業職	その他		400~500万円未満		その他	4年以上5年未満
5		40代後半・50代前半		研究開発職	親と子供		500~600万円未満		分からない	5年以上6年未満
6		50代後半		技術職			600~700万円未満			6年以上7年未満
7		60代以上		技術系以外の専門職			700~800万円未満			7年以上
8				管理職			800~900万円未満			
9				その他一般事務職			900~1000万円未満			
10				自由業(開業医、開業弁護士、大学教授など)			1000~1200万円未満			
11				経営者(会社役員など)			1200万円以上			
12				自営者およびその家族			わからない 答えたくない			
13				農林漁業						
14				パート						
15				学生						
16				主婦専業						
17				無職						
18				その他:具体的に						

注)選択肢数とは、調査票に記載された選択肢の数を表す

手順 3 : 反応変数の用意

ここでは反応変数として、情報通信機器の所有の有無を取りあげる。

- ・ PHS
- ・ 携帯電話
- ・ カーナビゲーションシステム
- ・ レーザ以外のプリンター
- ・ PHS
- ・ PDA
- ・ デスクトップ型 PC
- ・ デジタルカメラ

手順 4：傾向スコアの算出(4 章(10)式参照)

傾向スコアは処遇変数を目的変数,手順 3 で準備した共変量セットを説明変数としたロジスティック回帰により求める。オプションとして,ステップワイズ法の使用も考えられるが,ここでは用いない。

手順 5：共変量調整

手順 4 で求めた傾向スコアならびに(7)式を用いて共変量調整を行う。それが適切に行われたかを確認するために,各共変量の調整前後の平均の差を用いて,比較をする。もし,この差が大きければ共変量の組み合わせを再度変えて,傾向スコアの推定を行う必要がある。

手順 6：反応変数の修正

共変量調整の後,(7)式を用いて反応変数の修正を試み,以下の量を求める。

- 修正前の非公募系の平均(4 章(3)式に相当)
- 修正前の公募系の平均(4 章(4)式に相当)
- (各共変量セット毎)修正後の公募系の平均(4 章(7)式に相当)

評価尺度

- 低減率(percent reduction)

$$100 \times \left(1 - \frac{|\bar{y}_1''(J) - \bar{y}_0|}{|\bar{y}_1 - \bar{y}_0|} \right)$$

ただし, $\bar{y}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} z_i y_{1i}$ とする。(各記号については 4 章を参照)

ここでは,低減率が 100 に近いほど,修正効果があるということにする

5.6 実験計画

本研究は「非公募系の結果を基に公募系の結果を修正したときに反応変数がどのような挙動を示すか」という想定の中,修正を行う。さらに,5.4 で掲げた仮説を検証するために次のような実験計画を用意した。

実験 1『修正効果の有無の検証』

実験 1 では、傾向スコアによる加重修正の効果を確認するために、同一の共変量セットを用いて、傾向スコアの推定、推定された傾向スコアを用いて共変量調整・反応変数の修正を順次行う。実験 1 の条件は以下の通りである。

適用データ : 「インターネット編」のデータの一部

共変量セット : 共変量セット 1

反応変数 : デスクトップ型 PC, デジタルカメラ, PHS, レーザ以外のプリンタ

実験 2『異なる共変量セットを用いたときの修正効果の検証』

実験 2 では同一の反応変数に対して、複数の共変量セットを用いたときに、反応変数の修正の程度を表す低減率がどのように変化するか、またどのような特徴が観測されるかなど、探索的に調べる。共変量セットの組合せについては 5 章で示した通りであるが、再び説明をすると共変量セット 1,2 は基本属性を中心とし、共変量セット 3,4 はインターネット調査における基本属性を加えた多値変数で共変量セットが構築されている。

適用データ : 「インターネット編」のデータの一部

共変量セット : 共変量セット 1~4 を使用

反応変数 : PHS, 携帯電話, カーナビゲーションシステム, レーザ以外のプリンター, PDA, デスクトップ型 PC

実験 3『2 値変数により構築された共変量セットの影響』

実験 3 では共変量の形式を変えて共変量セットを構築し、その影響を調べる。共変量の形式は 1 変数あたり複数の共変量からなる多値変数として用いる場合と複数の選択肢を事前に 2 値化処理をし、それぞれを変数として扱う 2 値変数があった。2 値変数を全て選択すれば、多値変数と同じであるが、2 値変数を用いる時は、全て選択せず幾つかを拾い、それらを組合せて用いることとする。前者は、設問の選択肢をそのまま利用できる(頻度バランスに応じてプーリングを行うことが

望ましい場合もある)ことから、解釈や利用がしやすいという点が挙げられる。一方、後者は2値化された共変量を試行錯誤的に組み合わせることで、推定の精度が多値変数に比べて高くなることが予想される。どの程度、修正効果に違いが生じるかを観察する。

尚、原則として共変量セット5は共変量セット4で用いた共変量の中から、試行錯誤的に2値変数を組み合わせる。

適用データ : 「インターネット編」のデータの一部

共変量セット : 共変量セット4と共変量セット5

反応変数 : PDA, デジカメ, デスクトップ PC

- 共変量セット5(全て2値変数で構成)

例) 年齢(30-34才)は年齢が30才から34才にあてはまれば1, そうでなければ0であることを示す。

年齢 ... 年齢(30-34才), 年齢(70才以上),

性年齢区分... 性年齢区分(男性 35-39才), 性年齢区分(25-29才), 性年齢(男性 35-39才), 性年齢(男性 65-69才), 性年齢(女性 20-24才),

職業 ... 職業(労務職), 職業(保安・サービス), 職業(研究開発職), 職業(経営者), 職業(一般事務職), 職業(大学生),

家族構成 ... 家族構成(親と夫婦と子供)

住居形態 ... 住居形態(賃貸一戸建て), 住居形態(民間アパート), 住居形態(社宅・官舎), 住居形態(その他),

世帯収入 ... 世帯収入(500-600万), 世帯収入(500-700万), 世帯収入(1500万以上),

回答場所 ... 回答場所(職場), 回答場所(学校),

通信回線 ... 通信回線(ダイヤルアップ), 通信回線(その他),

インターネット歴... インターネット歴(1.5-2年), インターネット歴(2-2.5年), インターネット歴(3-4年), インターネット歴(4-5年), インターネット歴(5-6年), インターネット歴(6-7年), インターネット歴(7年以上)

実験 4『加重修正法を用いた処遇効果の検出』

実験 4 では、実験 1～3 で用いた「インターネット編」のデータセットでなく、「生活意識編」のデータセットを用いる。そこで、同一の反応変数、ほぼ同一の共変量セットを用い、公募系と非公募系の反応変数の差異に与える要因を回答者と調査票である状況下で実験をする。この場合、反応変数に与える影響は、調査票と回答者である。実験 1 ならびに 2 で得られた反応変数の修正効果が高かった反応変数とその共変量セットが最適な組合せであると仮定し、その組合せを用いる。「生活意識編」における反応変数の差異を修正することができる組合せであれば、回答者の違いによる影響は除去されるので、反応変数に与える要因は調査票の違いと考えることができる。そこで、調査票の影響を処遇効果として考え、調査票の影響の有無があるかを確認する。

適用データ : 「生活意識編」のデータの一部

反応変数と共変量セット

- 共変量セット 1 に対して、「PHS」「ノートパソコン」
- 共変量セット 2 に対して、「レーザープリンタ」
- 共変量セット 3 に対して、「レーザー以外のプリンタ」

ただし、インターネット編の共変量セットの構成と生活意識編の共変量セットの構成は若干異なるが、ほぼ同じと仮定して用いることにする。

6章 結果と考察

6.1 傾向スコアの推定

手順に従って、公募系または非公募系を表すダミー変数(処遇変数) Z を目的変数、共変量セットを説明変数としたロジスティック回帰分析を用いて、傾向スコアの推定値を求めた。共変量セット1を用いた時の公募系と非公募系のそれぞれの傾向スコアの推定値の分布を図6.1.1ならびに図6.1.2に示す。

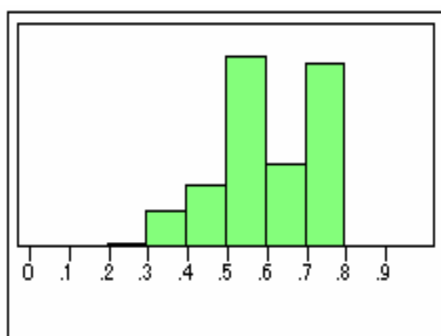


図 6.1.1 傾向スコアの推定値の分布
(公募系)

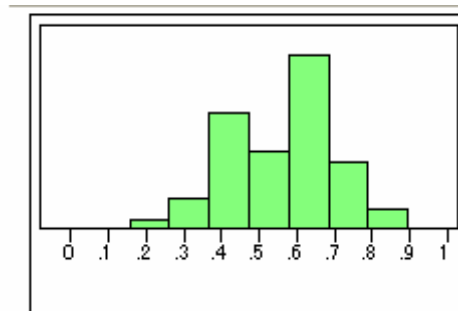


図 6.1.2 傾向スコアの推定値の分布
(非公募系)

図6.1.1, 図6.1.2から分かるように、公募系は全体的に傾向スコアが高いところに集中しており、非公募系は裾が長い分布になっている。特に、傾向スコアの小さい値では公募系の割合が小さく、非公募系の割合が大きくなっている。非公募系の回答分布を推定するには、公募系のこうした少ない層に大きな重みを与えるような加重修正を行う。これが傾向スコアによる加重修正法の本質である。

6.2 実験1の結果

図6.2.1は修正前後の共変量の分布の様子である。この図は公募系、非公募系それぞれの共変量の値の平均を示している。例えば、加重修正前の性別(男性)であれば、公募系は65%ぐらいが男性であり、非公募系は50%強が男性であることを示している。右側は加重修正後の様子を表しており、意味するところは加重修正前と同じである。気をつけたいのは、非公募系の値は修正されない。言い換えれば、修正前と修正後では同じ値になる。

図6.2.2は各共変量について公募系と非公募系の平均の差の変化を表している。

図 6.2.1, 図 6.2.2 共に, 各共変量の公募系と非公募系の平均の差が小さくなっていることが確認できる. この状態は共変量の分布が揃っていることを意味し, 共変量調整が行えたと言える.

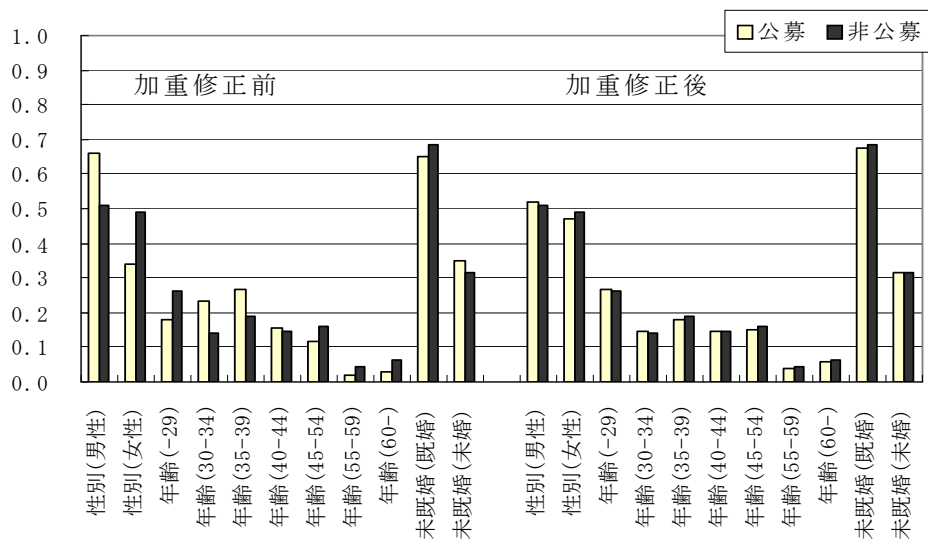


図 6.2.1 共変量調整 1(共変量セット 1)

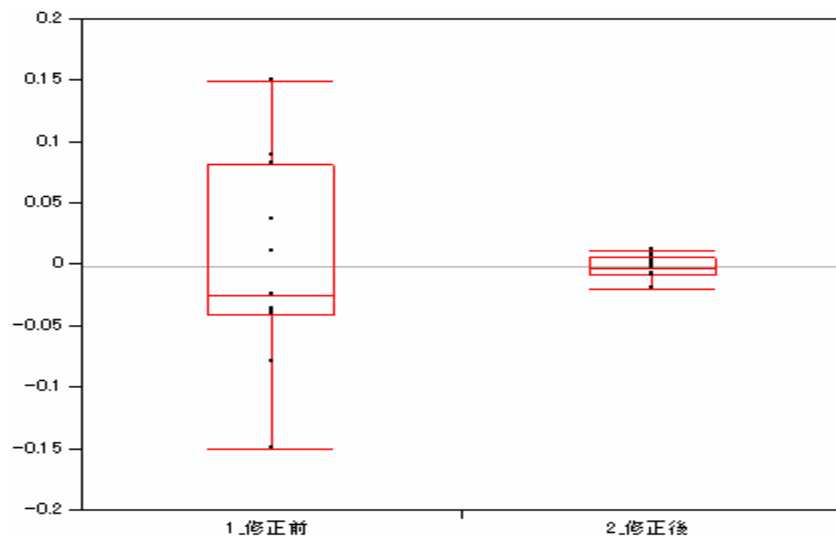


図 6.2.2 共変量調整 2(共変量セット 1)

共変量調整が行われたので, 共変量の影響が除去されたと考えてよい. 次に共変量セット 1 で求めた傾向スコアを用いて, 反応変数の修正を行う. その様子が

図 6.2.3 である。

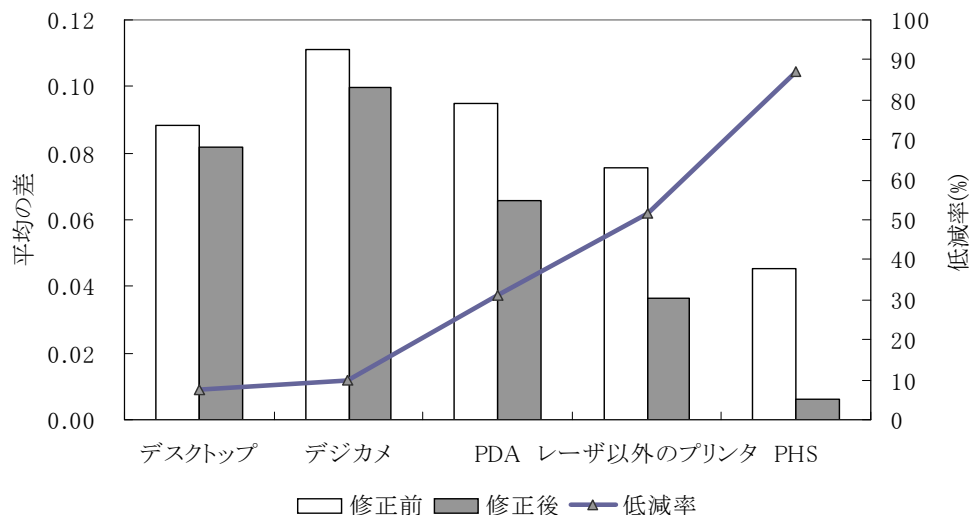


図 6.2.3 低減率 (共変量セット 1)

棒グラフは修正前ならびに修正後の公募系と非公募系の平均の差を表し、折れ線グラフは低減率である。本研究のデータの特徴は、各反応変数に対して適切な共変量調整を行うことができれば、公募系と非公募系の反応変数の差が解消される。言い換えれば、修正の程度を表す低減率は 100%に近づく。

ここで、反応変数をみると「デスクトップ」「デジカメ(注：デジタルカメラの略称)」はほとんど修正されていないことが見て取れる。一方、「PHS」は低減率の値が大きく、共変量セット 1 でほぼ的確に機能したといえる。このことから、「PHS」は共変量セット 1 で構成される「性別」、「年齢区分」、「未既婚」が共変量になっていると思われる。一方、「PHS」以外の反応変数は共変量セット 1 の共変量だけでは不足であり、ここで用いた以外の「観測されていない共変量」の存在を考慮しないとイケないといえる。このように、同一の共変量セットを用いたとしても反応変数の修正の程度が異なること、つまり、それは反応変数に応じた共変量セットを探索する必要があること意味している。また、ここでは「PHS」のように適切だと思われる共変量セットを見つけたが、反復的な検証を行わないと確かだとはいえない。

6.3 実験2の結果

最初に、共変量セットに用いる変数の数が増えると、判別率(実際に公募系の回答者を公募系と判定する確率)ならびに共変量調整にどのような影響を与えるかを確認する。

表 6.3 判別率

	セット1	セット2	セット3	セット4
判別率(%)	61.6	65.0	74.2	74.1

共変量の数が増えるに従って、判別率が上がる様子が確認できる。

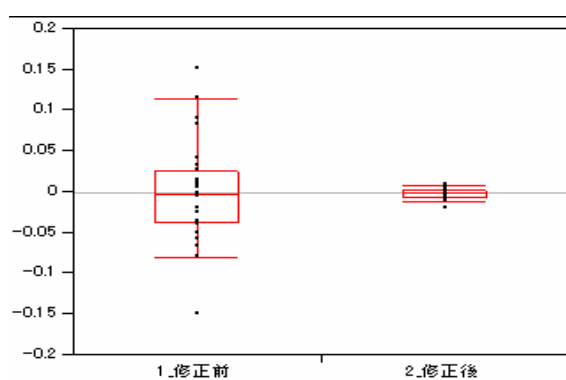


図 6.3.1 共変量調整(共変量セット2)

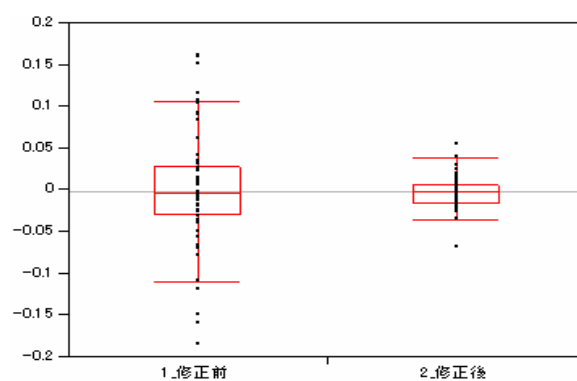


図 6.3.2 共変量調整(共変量セット3)

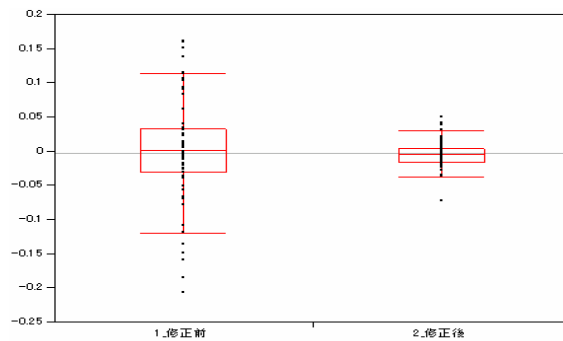


図 6.3.3 共変量調整(共変量セット 4)

修正前の公募系と非公募系の各共変量における平均の差は共変量セット 2 が小さく、共変量の数が多くなる共変量セット 3,4 では、その差が大きくなっている様子が見て取れる。共変量セット 2 では共変量セット 1 のように的確に調整がなされたが、共変量セット 3 ならびに共変量セット 4 では調整しきれなかった共変量の存在が幾つか見られる。調整が機能しなかった共変量の存在がどのように影響するかは分からないが、相対的に修正前後で差が小さくなっているため調整が機能していると考え、共変量セット 2,3,4 で求められた傾向スコアを用いて、以下反応変数の修正を試みる。

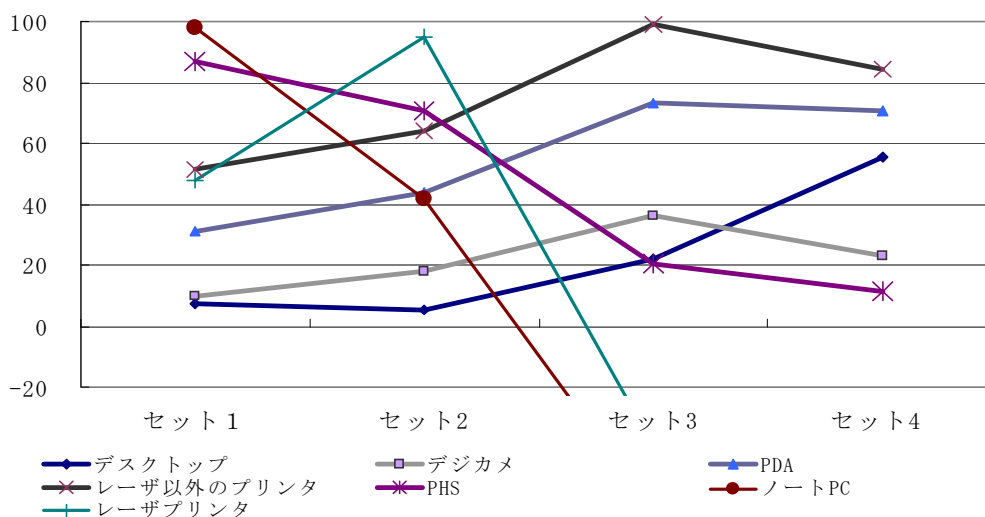


図 6.3.4 低減率(共変量セット 1~4)

図 6.3.4 は低減率の様子を示している。共変量セット 1 は実験 1 に示した「PHS」と「ノート PC」が修正されている。また、共変量セット 2 では「レーザプリンタ」、共変量セット 3 では「レーザ以外のプリンタ」、共変量セット 4 でも「レーザ以外のプリンタ」の修正効果があったと言える。

反応変数に着目してこの結果の解釈を行う。共変量セット 1 から共変量セット 4 になるにつれ、共変量の数が増えているが反応変数が修正されるとは限らない。「ノート PC」や「レーザプリンタ」や「PHS」ではむしろ、共変量の数が増加するにつれ、修正の程度が悪化している。このことから、共変量の数と修正の程度は関係がなく、実験 1 で得た結果と同じになるが反応変数に応じた共変量セットを探索する必要がある。一方、修正の程度が小さいが、「PDA」「デスクトップ PC」は基本属性で構成された共変量セット 1 よりも基本属性にインターネット上の基本属性を加えた共変量セット 4 の方が修正される反応変数もある。これは、共変量の数ではなく共変量セット 3, 4 に含まれる「インターネット歴や利用形態などに関する変数」が影響していると考えられる。

また、「デジタルカメラ(図内：デジカメ)」はここで取りあげた共変量セットでは十分な修正を行えなかった。

6.4 実験 3 の結果

修正効果に大きく影響を与えるのは傾向スコアである。そのため、多値変数で構築された共変量セット 4 と 2 値変数の組合せで構築された共変量セット 5 から求められる傾向スコアの分布が同一であれば、修正効果が変わらないと考えられる。そこで最初に共変量セット 4 と共変量セット 5 で傾向スコアの分布がどのように異なるかを観察する。

その様子が図 6.4.1 である。各点は回答者毎の共変量セット 4 と共変量セット 5 で求められた傾向スコアに対応する。この図は 45 度の直線上に点が集まれば、共変量セット 4 で求めた傾向スコアと共変量セット 5 で求めた傾向スコアが一致することを意味する。この図をみると、傾向スコアの分布は一致していないが、完全に異なるとも言えない。

ここで求めた傾向スコアを使って、修正効果がどの程度異なるかを調べるために、「PDA」「デスクトップ PC」「デジタルカメラ」を取りあげる。その結果が図

6.4.2 である.

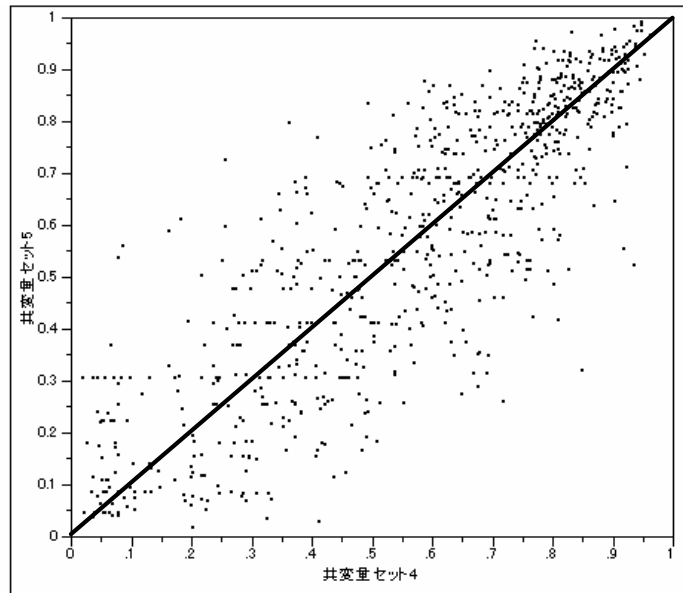


図 6.4.1 傾向スコアの分布

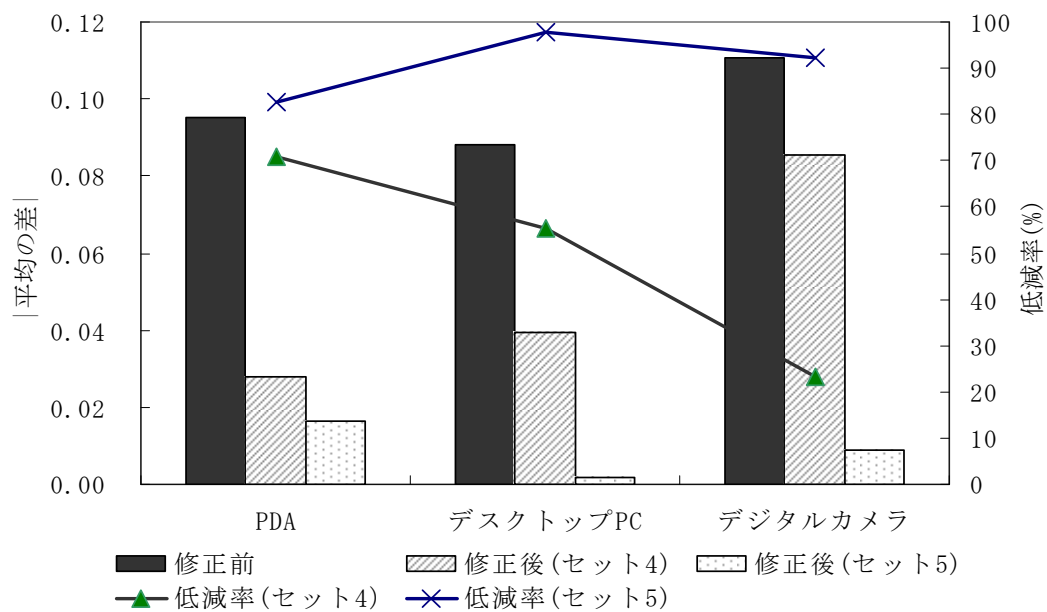


図 6.4.2 低減率比較(共変量セット 4 と共変量セット 5)

共変量セット4を用いて、「デジタルカメラ」や「デスクトップPC」への修正を試みた時は、ほとんど修正されなかったが、共変量セット5では修正が適切に行われている。特に「デジタルカメラ」は、共変量セット4で修正が適切に行われなかったのは“観測されていない共変量の存在”だと考えられたが、共変量セット5内の変数を2値化し、適当に組み合わせると、修正されることが分かった。また、傾向スコアの分布が類似であったとしても、修正効果には大きな影響を与える可能性がある。

ここでは確認しないが、実験2にて修正が適切に行われなかった反応変数についても同様のことが言える可能性がある。

6.5 実験4の結果

実験4の環境下では、反応変数の差が見られる要因としては調査票と回答者が主な要因として考えられる。そこで、実験2において、修正効果のあった反応変数と共変量セットの組合せが最適であると仮定して、「生活意識編」の反応変数に対して修正を試みれば、回答者の偏りによる影響は除去されることが期待され、反応変数に与える要因は調査票の影響だけだと考えてよいであろう。そこで修正後の反応変数を観測した時に

- ・ 反応変数の差異がある→調査票の影響がある
- ・ 反応変数の差異がない→調査票の影響がない

と言える。つまり、実験4では調査票の影響を処遇効果と考え、傾向スコアによる加重修正法を用いて処遇効果を検出する。

用いる反応変数と共変量セットは以下の通りである。

- ① 共変量セット1:「ノートPC」「PHS」
- ② 共変量セット2:「レーザープリンタ」
- ③ 共変量セット3:「レーザー以外のプリンタ(図ではノンレーザープリンタと記載)」

* 「インターネット編」と「生活意識編」では本研究が扱っている属性設問、反応変数に対しては同一の設問内容、同一の選択肢からなる。共変量セットを構築する際に選択肢の一部をプーリングしているために、若干共変量の構成が異なっている。

図 6.5 が結果である。インターネット編で得られたデータを用いた時には修正効果の確認された反応変数と共変量の組合せであるが、図 6.5 では全ての反応変数において修正されていない、つまり修正後に反応変数の差異が確認されるので、調査票の影響があったと考えてよい。ここでは調査票の影響を検出したが、調査時期や調査サイトなどの影響の有無を調べるためにも使うことができる。

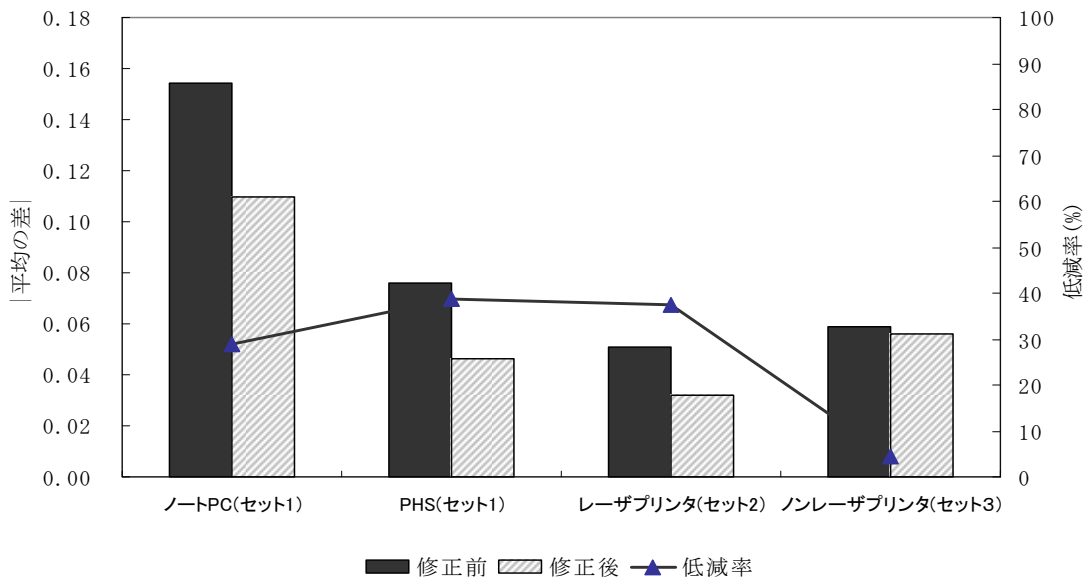


図 6.5 低減率(生活意識編のデータを用いた場合)

6.6 結果のまとめ

(1) 傾向スコアの推定について

傾向スコアはロジスティック回帰により簡便な方法で求められることができた。また、求めた傾向スコアを用いて共変量調整や反応変数の修正が容易に行えることを確認した。また、傾向スコアの推定に用いた変数の数が多くなるにつれ、公募系である回答者を公募系(非公募系である回答者を非公募系)と判断する判別率が高くなることが分かった。

(2) 共変量調整について

傾向スコアを用いると概ね共変量調整ができることも確認した。しかし、共変量の数が増えると調整が難しくなる共変量も存在した。この存在が後に行う反応変数の修正にどのように効いてくるのか、未知であり今後の課題になる。

(3) 反応変数への修正について

実験 1 ならびに実験 2 において、同一の共変量セットを用いて修正を行っても修正が適切に行われた反応変数と修正されない反応変数があることが分かった。このことは、全ての反応変数に効果が期待される汎用的な共変量(セット)はなく、反応変数に応じた適切な共変量を模索する必要があることを示唆している。また、適切な共変量セットを見つけたとしてもさらなる反復検証が必要である。このことから、加重修正は容易でない。

(4) 共変量セット 1,2 と共変量セット 3,4 による修正効果の差異

「PHS」や「ノート PC」や「レーザープリンタ」は基本属性のみで構成された共変量セットで修正効果があることを確認できたが、インターネットの利用形態等を加えた共変量セット 3 または 4 では修正効果が確認できなかった。一方、「レーザー以外のプリンタ」、「PDA」などは共変量セット 1,2 では修正効果がなかったが、共変量セット 3,4 で修正効果があることが見て取れる。また、「デジタルカメラ」はここで取りあげた共変量セットに対してほとんど修正効果が見られなかった。

(5) 2 値変数により構築された共変量セットの影響

多値変数で構成された共変量セット 4 では修正できなかった反応変数に対して、2 値変数の組合せによる共変量セット 5 で修正を試みたところ、修正の程度が共変量セット 4 よりも大幅に改善された。

(5) 加重修正法を用いた処遇効果の検出について

傾向スコアによる加重修正法の適用可能性の一つとして、処遇効果(調査票の影響)の検出に用いた。実験 2 にて、修正効果の確認された反応変数と共変量の組み合わせが最適であるとしたならば、修正後の反応変数の差異は調査票の影響と考えることができる。修正後の反応変数に差異が確認されたので、調査票の影響があったといえる。

7章 考察ならびに今後の課題

7.1 考察

傾向スコアの推定はロジスティック回帰を用いて、簡単に求めることができたが、最適な傾向スコアの推定には困難が伴う。実験2の結果から分かるように、傾向スコアを算出するための変数(共変量)の数を増やせば、判別率は上がる。しかし、それは共変量調整や反応変数の修正に有効であるとは限らない。むしろ、共変量調整が適切に行われなかった共変量の存在による影響が懸念される。

実験1または実験2より、同一の共変量セットを用いて反応変数の修正を行ったところ、修正される反応変数と修正されない反応変数があった(図6.3.4)。これは、汎用的な共変量(セット)はなく、反応変数に応じた適切な共変量(セット)を探索する必要があることを示唆している。本研究では、「PHS」「ノートPC」「レーザープリンタ」「レーザー以外のプリンタ」において、適切だと思われる共変量セットを確認した。これが確かであるかを確認するためには、反復的な検証が必要である。検証をする際には、5章5.1で述べたような修正効果を測定できるような実験デザインが必要である。つまり、同一の調査環境を再現する必要がある。また、調査時期、社会環境が異なれば、せっかく見つけた共変量セットも変化する可能性がある。これらのことを考えると加重修正は容易ではない。

ここでは「基本属性」を表す変数のみで構築した共変量セットと「基本属性」に「インターネット利用に関する基本属性」を加えた共変量セットを用いて実験を行った。反応変数の修正の程度により、①基本属性のみで修正可能②基本属性＋インターネット利用に関する基本属性の組合せで修正可能③ここで取りあげた共変量では修正不可能の3分類することができる。①、②の場合は先に述べたように、反復的な検証が必要であるが、特に③の場合、十分に多くの共変量を測定し、専門知識や背景情報などを考慮しながら、探索する必要がある。

共変量セット1から4までは1変数あたり複数の共変量からなる多値変数を共変量の形式として用いた。共変量セット4の共変量を2値化し、それらを試行錯誤的に組み合わせた共変量セット5を用いて修正を試みたところ、共変量セット4では修正効果がなかった「デジタルカメラ」などが大幅に修正された。共変量セット4を用いた時の修正の失敗は、観測されていない共変量の存在かもしれないが、共変量の探索の失敗とも考えられる。このように、共変量の形式を変えて

共変量セットを構築すると修正できる可能性がある。ここで用いた共変量セット 5 は相当な試行錯誤の上で、見つけられた。最適な共変量セットの探索と効率的なアルゴリズムの開発が必要である。

傾向スコア加重修正法による適用例の一つとして調査票の影響を検出するために用いた。他にも調査機関や調査時期などの検出にも利用できる。例えば、1 年前の登録者集団(リソース/パネル)と現在の登録者集団が同質であるかを調べるために用いることができるであろう。本研究では公募系と非公募系を対象に実験を行ったが、従来型調査による「インターネット・ユーザー」と「非インターネット・ユーザー」でも同様の実験が行うことができ、これらの知見を総合すれば、インターネット・ユーザーの特性が分かるかもしれない。いずれにせよ、修正効果を測定できる環境が整っていることが重要である。修正効果を測定できない環境下で実験を行ったとしても、修正効果か他の要因の影響かを判断することができないためである。例えば、従来型調査と Web 調査の間で加重修正の適用を考えた場合、処遇効果に相当する調査方式の影響が入ってくる。そのため、修正後の反応変数の差異が確認された場合、共変量調整の失敗か、調査方式の影響は区別できない。加重修正法が適用するためには、加重修正法が適切に機能するためには、適切な調査計画や調査目的にあった設問形式のあり方に依存すると考える。

7.2 今後の課題

(1)修正効果の測定

ここでは、修正効果を測定するために低減率を用いた。低減率は修正の程度を表すために、一律に比べることができるが、修正前の公募系と非公募系の反応変数の平均の差が小さい場合、修正率は敏感に反応する。低減率とあわせた指標が求められる。

(2)ステップワイズ法の使用

傾向スコアを求めるときに本研究では全変数を選択したが、モデル選択としてステップワイズ法の使用が考えられる。吉村・松田・大隅他(2003)ではステップワイズ法を用いたが、重要な共変量を選択されない可能性があること、その影響が未知なので検討を要すべき事項である。

(3)交互作用項の検討

多値変数で構成された共変量セット 1～共変量セット 4 は交互作用を検討しなかった。実際には、交互作用の影響も考えられる。全ての交互作用項を用いることは現実的に無理であるので、交互作用項をどのように検出し、選択するかは今後の課題になる。べき乗項の検討についても同じことが言える。

(4)共変量調整について

共変量調整は、各共変量の値の修正前後の差を確認することで、成否を決めた。共変量調整が客観的に判断できる指標が必要である。また、調整ができなかった共変量の影響があることは多分に考えられ、これらの影響を測定する必要がある。

最後に

本研究ではインターネット調査における回答の“偏り”を解消する一手段である傾向スコアによる加重修正法を用いて、その修正効果を確認することができた。しかし、この方法は偏りが生じた後の処置である。偏りが生じないように、登録者集団のあり方や実査の行い方などを詳細に検討し、少しでも偏りが小さくするような取り組みが大事であることを忘れてはいけない。

謝辞

本論文を書くにあたり，熱心にご指導をいただいた統計数理研究所大隅昇教授，大学入試センター吉村宰助教授，(株)電通リサーチ横原東首席部長，武田正樹氏，ならびに第4次実験調査参加メンバーの方々に，この場を借りて厚く謝意を表します．また，4年間に渡り，親身なご指導してくれた当研究室の逆瀬川浩孝教授にも心より感謝の意を表します．

また，当研究室の大学院生ならびに学部生にはお世話になった．ここにあわせて，感謝の意を表す．

【参考文献】 英文

- [1] Bethlehem, J. G.(2002) , Weighting Nonresponse Adjustments Based on Auxiliary Information, pp275 – 287.
- [2] Berrens, P., Bohara, A.K., Jenkins-Smith, H., Silva, C. and Weimer, D.L. (2001), The Advent of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples, *Political Analysis*, **11**, 1-22.
- [3] Cochran, W.G.(1965), The Planning of the Observational Studies of Human populations(with discussion). *Journal of the Royal Statistical Society*, **A,128**.234-55
- [4] Couper, M. P. (2000), Web Surveys: A Review of Issues and Approaches, *Public Opinion Quarterly*, Winter 2000, 464 – 494.
- [5] Couper M. P. (2001), The Promises and Perils of Web Surveys, A. Westlake and others. (eds), *The Challenge of the Internet*. London: Association for Survey Computing, 35-56.
- [6] Couper, M.P. (2003), The Internet and Other Survey Opportunities, JMRA 第 33 回トピックスセミナー, 2003 年 10 月 23 日開催,
- [7] D'Agostino RB Jr(1998), Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group, *Statistics in Medicine*, **17**, 2265-2281
- [8] Gelman, A. and Carlin, J. B.(2002), Poststratification and Weighting Adjustments, pp289 – 302.
- [9] Groves, R. M. (1989), *Survey Errors and Survey Costs*, John Wiley, New York.
- [10] Hahn, J. (1998) On the role of the propensity score in efficient semiparametric estimation of average treatment effects, *Econometrica*, **66**, 315-332
- [11] Imbens, G. W. (2000), The role of the propensity score in estimating dose-response functions, *Biometrika*, **87** (3), 706-710.
- [12] Miller, T. W. and Panjikaran, K. J. (2001), Studies in Comparability: The Propensity Scoring Approach, A.C. Nielsen Center for Marketing Research, University of Wisconsin-Madison, Madison, WI.
- [13] Ohsumi, N. and Yoshimura, O. (1999), The online survey in Japan: An evaluation of emerging methodologies, Bulletin of the International Statistical Institute 52nd Session, Book2, 171-174.

- [14] Rosenbaum, P. R. (1984) ,From Association to Causation in Observational Studies: The Role of Test of Strongly Ignorable Treatment Assignment. *Journal of American Statistical Association*, **79**, 41-48.
- [15] Rosenbaum, P. R. (1987), Model-Based Direct Adjustment, *Journal of American Statistical Association*, **82**, 387-394.
- [16] Rosenbaum, P. R. (1995), *Observational Studies*, Springer Series in Statistics, Springer.
- [17] Rosenbaum, P. R. and Rubin, D. B. (1983), Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome, *Journal of Royal Statistical Society, Series B* **45** (2), 212-218.
- [18] Rosenbaum, P. R. and Rubin.D. B. (1984), Reducing Bias in Observational Studies Using Subclassification on the Propensity Score, *Journal of American Statistical Association*, **79**, 516-524.
- [19] Rosenbaum, Paul R. and Rubin, Donald B. (1985), Constructing a Control Group Using Multivariate attached Sampling Methods that Incorporate the Propensity Score, *The American Statistician*, **39**, 1, 33-38.
- [20] Rosenbaum,P. R. and Rubin.D. B. (1983), The central role of the propensity score in observational studies for causal effects., *Biometrika*, **70** (1), 41-55.
- [21] Rubin.D.B.(1973),Matching to remove bias in observational studies,*Biomtrics*,**29**, 159-183
- [22] Saris, W. E. (1991), *Computer-Assisted Interviewing*, A Sage University Paper, Sage Publications.
- [23] Taylor, H. (2000) Does internet research work?, *International Journal of Market Research*, **42**(1) 51-53.
- [24] Taylor, H., Bremer, J., Overmeyer, C., Siegel, J.W. and 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
- [25] Terhanian, G. and Taylor, H (1999),Heady Days Are Here Again-online polling is rapidly coming of age,*Public Perspective*,June/July,20-23

【参考文献】 和文

- [26] 大隅昇(2000), 『「調査環境の変化に対応した新たな調査法の研究」報告書』, 文部省科学研究費, 特定領域研究「統計情報活用のフロンティアの拡大」(略称: ミクロ統計データ), 研究計画 A02 班(公募研究)「ミクロデータ利用の社会的制度の問題点」(課題番号: 09206117), 東京.
- [27] 大隅昇(2001a), 調査環境の変化と新しい調査法の抱える問題, 統計数理研究所公開講演会抄録, 統計数理, **49**, 201-213.
- [28] 大隅昇(2001b), 電子調査: その周辺の話題-電子的数据取得法の現状と問題点-, 統計数理,**49** (1), 201-213
- [29] 大隅昇(2002a), インターネット調査, 「社会調査ハンドブック」, 朝倉書店
- [30] 大隅昇(2002b), インターネット調査の適用可能性と限界-データ科学の視点からの考察-, 行動計量学, **29**, 1, 20-44.
- [31] 大隅昇(2002c), インターネット調査の適用可能性と限界-実験調査から見えること-, 輿論科学協会創立 56 周年記念講演, 市場調査, 250 号
- [32] 大隅昇(2003), インターネット調査を検証する-質の評価と標準化に向けて-, JMRA 研修セミナー, 1-57
- [33] 佐野紳也(2003), Web 調査の現状と可能性, 品質, **33**(2), 289-295
- [34] 島村一俊(1999), メールマガジンを使用したインターネットセルフセレクション調査に関する考察, 早稲田大学理工学部卒業論文
- [35] 鈴木文雄, 笹田幸典 (2003), インターネット・サーベイと従来型調査の比較検証-Cyber Panel と NOS (オムニバス調査) を用いた比較実験調査から-, ISM シンポジウム「インターネット調査の現状を検証する-調査法としての評価方法と標準化を考えるか-, 予稿集, 95-110
- [36] 住本隆(2002), インターネット調査に要求されるもの-ハリスインタラクティブのデータ・ウェイトイング方法-, エストレーラ, **95** 号, 2 月号, 11-19.
- [37] 中谷吉孝, 上嶋幸則, 渡會隆, 瀧中勢子, 蓑原勝史 (2003), e-HABIT の特徴と今後のインターネット調査パネル構築-各実験サイトのデータ比較から-, ISM シンポジウム「インターネット調査の現状を検証する-調査法としての評価方法と標準化を考えるか-, 予稿集, 75-93.
- [38] 能見正 (2000), 双方向性ネットワークを利用した調査手法とその影響, 郵政研

究月報，9月号，72-97.

- [39] 星野崇弘(2003)，調査データに対する傾向スコアの適用，品質，33(3)，312-319
- [40] 星野崇宏，鈴木督久(2003)，傾向スコアを用いた Web 調査の無作為抽出への近似，日本行動計量学会第31回大会発表抄録集，184-185
- [41] 細井勉(2002)，マーケティング・リサーチ領域におけるインターネット調査概論－現状と展望－，エストレーラ，95号，2月号，2-10.
- [42] 横原東(2001)，マーケティングにおけるインターネット調査の実状と課題，統計数理研究所公開講演会抄録，統計数理，49，215-222.
- [43] 横原東，武田正樹，細井勉(2003)，DENTSU_R-net に基づくインターネット調査の検証－とくに第4次実験調査結果を中心として－，ISM シンポジウム「インターネット調査の現状を検証する－調査法としての評価方法と標準化を考えるか－」，予稿集，55－74.
- [44] 吉村宰(2001)，インターネット調査にみられる回答者像，その特性，統計数理研究所公開講演会抄録，統計数理，49，223-229.
- [45] 吉村宰(2003a)，インターネット調査を検証する－質の評価と標準化に向けて－，JMRA 研修セミナー，1-32
- [46] 吉村宰(2003b)，Web 調査の現状と課題－調査誤差の分類と対処の観点から－，日本行動計量学会第31回チュートリアルセミナー
- [47] 吉村宰，大隅昇(2003)，インターネット調査の質の評価を考える，ISM シンポジウム「インターネット調査の現状を検証する－調査法としての評価方法と標準化をどう考えるか－」，予稿集，15-32
- [48] 吉村宰，大隅昇，清水信夫(2002)，インターネット調査の諸特性と今後の展開のあり方－第4次実験調査から見えてきたもの－，第30回日本行動計量学会大会特別セッション
- [49] 吉村宰，松田浩幸，大隅昇，横原東，武田正樹(2003)，インターネット調査における加重修正法に関する一考察，行動計量学会予稿集，120-123

参考 Web サイト

総務省 HP : <http://www.soumu.go.jp/index.html>

ビデオリサーチネットコム : <http://www.vrnetcom.co.jp/>