

調査設計とデータ特性

若年者のライフスタイルと意識に関する調査の基礎分析

濱本真一

(東北大学 大学院教育学研究科)

はじめに

本論では、東北大学教育学部の2014年度前期開講科目「教育学実習（統計的調査実習）」において実施された調査の設計と、得られたデータの基本特性を整理する。「統計的調査実習（以下本実習と称す）」では調査企画、調査票設計、標本抽出、実査、エディティング、コーディング、データ入力、クリーニング、分析、そして報告書執筆に至る一連の過程を、受講者全員が体得することを目的としている。学生たちは本実習のカリキュラムを通じて、社会調査の実際、および社会調査を用いた論文作成の基本的なスキルを学ぶことになる。

後にも詳しく述べるが、本実習で行われた調査は調査票を用いた標本調査である。標本調査において対象者となる標本が対象とする母集団の適切な縮図でなければならないということは、社会調査法の基礎としてよく知られている。そこでの対象者の選び方は母集団に対する単純無作為抽出が原則とされる。しかし多くの場合、標本調査はその目的や制約によって様々な代替的方法がとられ、その方法に応じて得られた標本は何らかの独特な「クセ」を伴うことが多い。本調査においてもそれは例外ではなく、得られたデータの特徴を理解しておくことは、本調査を用いて何らかの社会科学的な知見・命題を導き出す際の留意点を理解することと同じである。本章では、本調査のサンプリングの方法を整理し、さらに得られた回答者の属性の分布、回答率、回答の傾向に関して、それぞれ基礎的な分析を行う。

1. サンプルデザインと調査のスケジュール

本年度調査の概要は表1に示す。対象者は例年通り20歳以上40歳未満の非学生に設定している。学生は各々の問題関心に沿って質問項目案や分析計画をたてた。調査票（本報告書付録に記載）は各班の関心を最大限生かして作成されたものである¹⁾。対象者の抽出は、協力調査会社のモニター登録者のうち、条件に合致する方々に協力を依頼し、500名の計画サンプルを得た。

抽出に際して性別・学歴・年齢層による割付を行った。昨年までの調査から、特に学歴や年齢層で大きなサンプルの偏りが生じる可能性が予測されており（三輪2012, 石田2013）、基本的な属性周辺分布に極端な偏りを生じさせないための措置である。しかし、これらの割付を行うことは、母集団の代表性を担保することにはならない。想定している母集団（理論母集団）は「全国の若年者」であるが、学生でないという条件のため、多くの20代前半の人々は母集団から除外される。年代の分布が国勢調査と異なることは、むしろ当然とも

表1 調査概要

| | |
|--------|---|
| 調査名称 | 若年者のライフスタイルと意識に関する調査 |
| 調査主体 | 東北大学教育学部・教育学実習／山形大学地域教育文化学部・社会調査演習 |
| 調査目的 | 現代日本の若年層に対し、教育、就業、生活に関する行動や意識について質問紙調査を行い、それらにかかわる社会科学的な命題を検討する |
| 調査対象 | 理論母集団： 調査時点で29~39歳の全国の男女(学生を除く) 調査母集団： 上記の条件を満たすインターネット調査モニター |
| データ数 | モニター調査によって抽出した500名に調査票を郵送し、460名からの回答を得た(回収率92.0%) |
| 調査時期 | 2014年6月17日~28日 |
| 抽出方法 | 上記インターネット調査モニターのうち、学歴・性別・年代ごとに割付を行ったうえで協力を依頼 |
| 調査方法 | 基礎項目等についてウェブ上で尋ねたうえで、調査法を郵送 |
| 調査項目 | 1基本属性項目 2学歴・職業に関する項目 3青少年期の家庭背景に関する項目 4現在の生活習慣・文化活動に関する項目 5現在の幸福感・価値観に関する項目 6ソーシャルメディア等のかかわり方に関する項目 7学生時代の習慣・意識に関する項目 |
| 調査協力機関 | 楽天リサーチ株式会社 |

表2 性別×年齢層×学歴のクロス表

| 年齢性別 | 国勢調査 | | | | 計画標本 | | | | 有効回答 | | | |
|-------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|
| | 中学校 | 高校 | 短大・高専 | 大学・大学院 | 中学校 | 高校 | 短大・高専 | 大学・大学院 | 中学校 | 高校 | 短大・高専 | 大学・大学院 |
| 20代男性 | 1.34% | 9.09% | 2.81% | 6.40% | 1.40% | 8.60% | 2.80% | 6.00% | 1.52% | 8.26% | 3.04% | 5.87% |
| 30代男性 | 1.96% | 12.80% | 4.40% | 10.92% | 2.20% | 13.00% | 4.40% | 10.80% | 2.17% | 11.96% | 3.91% | 10.65% |
| 20代女性 | 1.01% | 7.94% | 5.95% | 5.34% | 0.80% | 7.60% | 6.00% | 5.40% | 0.65% | 7.83% | 6.09% | 5.43% |
| 30代女性 | 1.19% | 11.64% | 10.69% | 6.54% | 1.60% | 11.80% | 11.00% | 6.60% | 1.74% | 12.39% | 11.52% | 6.96% |

表3 統計的調査実習のスケジュール

| | |
|-------|--------------------------|
| 2014年 | 4月7日 実習授業開始 |
| | 5月26日 山形大学との合同報告(於:東北大学) |
| | 5月30日 調査会社へ依頼 |
| | 6月2日 プリテスト・質問項目調整 |
| | 6月11日 調査票最終盤完成(付録) |
| | 6月16日 調査票封入・投かん |
| | 6月28日 調査票回収締切 |
| | 6月30日 データ入力開始 |
| | 7月7日 データクリーニング |
| | 7月8日 配布用データ完成 |
| | 8月9日 最終報告会(於:山形大学) |
| | 8月20日 レポート締切 |
| | 9月20日 最終稿提出 |

いえる。また、インターネットモニターという調査方法の性格上、実質的な母集団（調査母集団）は、「インターネットモニターに登録している全国の学生でない若年者」である²⁾。調査母集団が理論母集団から無作為に選ばれていない以上、調査母集団から得られるサンプルには理論母集団からの特性の乖離が生じる。例年同様、留意すべき点である。

割付は性別 2 カテゴリ、学歴 4 カテゴリ、年齢層 2 カテゴリの計 16 カテゴリで分類し、平成 22 年度の国勢調査をもとに行った。国勢調査、計画サンプル、有効サンプルにおける各カテゴリの分布は表 2 に示す通り。計画サンプル、有効サンプルともに国勢調査と比べて大きな偏りは見られない（ χ^2 乗値による適合度検定はいずれも有意でない）。

本実習のスケジュールは表 3 のとおりである。学生は授業内でテーマ設定、仮説構築、質問項目、分析結果に関するプレゼンテーションを行うのと同時に、封入作業、データ入力などの実作作業も行う。表 3 にあるようにテーマ設定から郵送開始の期間、またはデー

表4 回答者の分布

| | カテゴリ | 実数 | 比率 | 国勢調査 H22 | χ^2 乗値 | p値 |
|------------|-----------|-----|-------|-------------|-------------|-------|
| 性別 | 1 男性 | 218 | 47.4% | 48.7% | 0.307 | 0.580 |
| | 2 女性 | 242 | 52.6% | 51.3% | | |
| 年齢層 | 20-24歳 | 66 | 14.7% | 14.8% | 0.802 | 0.849 |
| | 25-29歳 | 101 | 22.5% | 24.2% | | |
| | 30-34歳 | 128 | 28.5% | 28.0% | | |
| | 35-39歳 | 154 | 34.3% | 33.0% | | |
| 婚姻状況 | 1 未婚 | 230 | 50.4% | 53.5% | 2.141 | 0.343 |
| | 2 既婚 | 207 | 45.4% | 43.2% | | |
| | 3 死別・離別 | 19 | 4.2% | 3.4% | | |
| 従業上の 地位 | 1 経営者・役員 | 4 | 1.2% | 1.7% | 80.423 | 0.000 |
| | 2 常時雇用 | 192 | 55.3% | 66.5% | | |
| | 3 臨時雇用 | 81 | 23.3% | 22.6% | | |
| | 4 派遣社員・嘱託 | 34 | 9.8% | 4.0% | | |
| | 5 自営業主 | 26 | 7.5% | 3.1% | | |
| | 6 家族従業者 | 6 | 1.7% | 2.0% | | |
| | 7 内職 | 4 | 1.2% | 0.2% | | |
| 職業 | 1 専門・技術 | 81 | 23.5% | 18.4% | 58.409 | 0.000 |
| | 2 管理 | 6 | 1.7% | 0.5% | | |
| | 3 事務 | 107 | 31.0% | 20.8% | | |
| | 4 販売 | 47 | 13.6% | 15.5% | | |
| | 5 サービス | 33 | 9.6% | 12.7% | | |
| | 6 運輸・保安 | 21 | 6.1% | 4.7% | | |
| | 7 生産現場・技能 | 48 | 13.9% | 26.0% | | |
| | 13 農業 | 2 | 0.6% | 1.3% | | |
| 学歴 | 中学校 | 28 | 6.1% | 5.4% | 0.905 | 0.824 |
| | 高等学校 | 186 | 40.4% | 39.6% | | |
| | 短大・専門 | 113 | 24.6% | 24.4% | | |
| | 大学・大学院 | 133 | 28.9% | 30.7% | | |

タ完成から報告書執筆に至る期間は非常に短い。そのため、授業外の時間に各班が自主的に作業を行うことが求められている。500人の調査対象者のうち300名を東北大の学生が、200名を山形大学の学生がそれぞれ封入・投かんし、返送された調査票もそれぞれ入力した。

2. 回収サンプルの特徴

回収されたサンプルの属性の分布を表4に示す。これらの分布が母集団を代表しているものなのかを判断するには、外的な基準が必要である。本論では、やや年度が異なるが、平成22年度の国勢調査を用いる。国勢調査における分布を真の分布と仮定し、本調査で得られたデータの属性との整合性を検証する。割り当てを行った3つの変数（年齢、性別、学歴³⁾）に関しては、およそ国勢調査と変わらない分布が得られた。例年、中卒学歴者の比率が低い傾向にあるが、本年度はそのような傾向は確認できない。

従業上の地位は国勢調査と異なる結果となっている。特に国勢調査とのかい離が大きいのは派遣社員・嘱託社員カテゴリである。職業⁴⁾にも有意なゆがみが生じている。とりわけ管理職および事務職は過剰に、生産現場・労務職は過小にサンプリングしていることがわかる。本実習では例年職業に関する分布のゆがみが大きい傾向にある。インターネット調査の特性としてある程度は仕様のないこととはいえ、職業に関する分析を行う際には、留意が必要である。この傾向は本実習で調査を開始した2011年度から共通の傾向である。

3. 欠票の分析

本調査の回答率は 92%であり、例年と同様高い値となっているが、事前のインターネット調査において尋ねた質問群を用いて回答と基本属性の関係について検証しておく。まず、性別・学歴・年齢層を用いた割付と回収率の関係は表 5 のとおりである。最終学歴が中学校の 20 代女性の回収率がやや小さいように見えるが、それほど大きな違いではない。属性ごとにみると、男性の回収率 (88.6%) は女性 (95.3%) よりやや低く、20 代の回収率 (92.2%) は 30 代の回収率 (1.9%) よりもやや高い。社会調査の回収率は一般に女性のほうが男性よ

表5 割り当てごとの回収率

| 年齢・性別 | 最終学歴 | | | |
|-------|------|------|-------|--------|
| | 中学校 | 高等学校 | 短大/高専 | 大学/大学院 |
| 20代男性 | 100% | 88% | 100% | 90% |
| 30代男性 | 91% | 85% | 82% | 91% |
| 20代女性 | 75% | 95% | 93% | 93% |
| 30代女性 | 100% | 97% | 96% | 97% |

表6 回答有に関する2項ロジット

| | Coef. | s.e. | オッズ比 | Coef. | s.e. | オッズ比 |
|-------------|----------|----------|--------|----------|----------|--------|
| WEB男性 | -0.915 * | (0.39) | 0.401 | -0.972 * | (0.39) | 0.378 |
| WEB年齢 | 0.012 | (0.03) | 1.012 | 0.014 | (0.03) | 1.014 |
| WEB職業(正規雇用) | | | | | | |
| WEB非正規 | -0.159 | (0.43) | 0.853 | -0.193 | (0.44) | 0.825 |
| WEB自営・内職 | -0.440 | (0.51) | 0.644 | -0.434 | (0.51) | 0.648 |
| WEB無職・その他 | 0.508 | (0.59) | 1.662 | 0.452 | (0.59) | 1.571 |
| WEB学歴(高卒) | | | | | | |
| WEB中卒 | 0.436 | (0.78) | 1.546 | 0.421 | (0.78) | 1.524 |
| WEB短大専門 | 0.135 | (0.45) | 1.144 | 0.156 | (0.45) | 1.169 |
| WEB大学 | 0.249 | (0.41) | 1.283 | 0.280 | (0.41) | 1.323 |
| WEB幸福度 | | | | | | |
| 定数 | 2.515 * | (1.11) | 12.367 | 2.889 * | (1.21) | 17.974 |
| N | | 500 | | | 500 | |
| L.L. | | -134.048 | | | -133.714 | |
| BIC | | 324.028 | | | 329.574 | |

*: $p < 0.05$

り回収率がよくなり、また年齢が上がるにつれ回収率がよくなる傾向を持つことが知られている (三隅・三輪 2008)。本調査では、性別に関してはそれと同様の傾向が、年齢に関しては逆の傾向がみられた。

本調査に対して「返送有=1」「返送なし=0」とする二項ロジスティック回帰を行った結果を表 6 に示した⁵⁾。2012 年度の調査では、中卒者が返送しにくく (石田 2013)、2013 年度では「無職・その他」カテゴリが返送しにくいという結果が得られた (濱本 2014) が、本調査ではそのいずれの傾向はない。学歴・職業による返送傾向の違いは認められない一方で性別の効果のみが検出され、男性が返送しにくいという結果になった。同じサンプリング計画によって調査対象者を選定しているにもかかわらず、返送確率オッズ比の傾向が年度ごとに異なっているという事実は興味深い。ただしここではそのメカニズムを突き止めることはできないため、今後の課題とする。なお、回収率と特に関係があることがこれ

表7 WEBと質問紙における従業上の地位のクロス表

| WEB働き方 | 質問紙働き方 | | | | | | | | | | | Total |
|---------|--------|-----|------|------|---------|---------|------|----|----|-----|-----|-------|
| | 経営者・役員 | 正社員 | 臨時雇用 | 派遣社員 | 契約社員・嘱託 | 自営業・自由業 | 家族従業 | 内職 | 無職 | その他 | 無回答 | |
| 経営者・役員 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 |
| 正社員 | 0 | 187 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 7 | 199 |
| 臨時雇用 | 0 | 1 | 72 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 77 |
| 派遣社員 | 0 | 0 | 1 | 7 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 10 |
| 契約社員・嘱託 | 0 | 1 | 2 | 0 | 18 | 1 | 0 | 0 | 0 | 1 | 1 | 24 |
| 自営業・自由業 | 0 | 0 | 1 | 0 | 0 | 20 | 0 | 0 | 0 | 0 | 1 | 22 |
| 家族従業 | 0 | 0 | 0 | 0 | 0 | 2 | 4 | 0 | 13 | 0 | 0 | 19 |
| 内職 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 3 | 2 | 0 | 0 | 7 |
| 無職：求職あり | 0 | 0 | 4 | 2 | 1 | 0 | 0 | 0 | 27 | 1 | 0 | 35 |
| 無職：求職なし | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 52 | 1 | 1 | 57 |
| その他 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 6 |
| Total | 4 | 192 | 81 | 12 | 22 | 26 | 6 | 4 | 98 | 5 | 10 | 460 |

不一致率= 14.1%

まで言及されたことのない「幸福度項目」は、回収率予測に含んでもあまりモデルは改善されなかった。

4. 回答の信頼性に関する考察

前年度同様、調査票にて得られた回答の信頼性（正確さ）について若干の考察をしていく。前節までに述べたように、本調査では一部の基礎項目をインターネット上で尋ね、そのうえで調査票を郵送している。郵送した調査票の中にも基礎項目は含まれているため、一部の変数については、同一人物に異なる形で同じ質問をしている。性別、年齢、従業上の地位、最終学歴の項目⁶⁾がそれにあたる（インターネットと質問紙で一部選択肢カテゴリの違いがある）。これらの項目は、インターネット上の回答と調査票の回答を比較することができる。このような試みは社会調査法の研究の中でも多くはなく（数少ない例として、林 2004 など）、考察する意義があると考えられる。

表 7～10 は、インターネット質問における質問と、質問紙における質問のクロス表である（学歴は 4 カテゴリに統合）。ウェブ回答と質問紙回答が一致しているときには全度数はクロス表の対角成分にあるはずである。表 7～10 を見ると、従業上の地位、学歴、年齢において不一致（非対角セルに正の度数）が生じている。

これらの回答のずれが生じる原因としては、以下のものが考えられる。

- 1) 入力ミス：入力の打ち間違い、わかりづらい回答に対する誤判断等
- 2) 回答者の変化：ウェブ上での回答時から調査票回答時（約 1 か月）の間に状況が変化した等
- 3) 質問項目の違い：Web 質問と質問紙で質問順序やワーディングが異なることによる回答のズレ
- 4) 回答者のミス：回答者の誤答、質問意図の誤認等
- 5) 意図的な誤答：見栄や社会的望ましさによる回答者の虚偽回答等
- 6) 代理回答：代理回答者による事実誤認や質問意図の誤認等

このうち、1) の打ち間違いの可能性は低いと考えられる。なぜなら本実習では、入力には細心の注意を払い、1 つの調査票につき 2 名が独立に入力を行い、それらを照合すること

表8 WEBと質問紙における学歴のクロス表

| WEB学歴 | 質問紙学歴 | | | | | | Total |
|-------|-------|------|-----------|-----|----|-----|-------|
| | 中学校 | 高等学校 | 短大・ 専門 | 大学 | 不明 | 無回答 | |
| 中学校 | 10 | 16 | 0 | 2 | 0 | 0 | 28 |
| 高等学校 | 0 | 166 | 3 | 14 | 0 | 3 | 186 |
| 短大・専門 | 0 | 69 | 43 | 1 | 0 | 0 | 113 |
| 大学 | 0 | 0 | 1 | 131 | 1 | 0 | 133 |
| Total | 10 | 251 | 47 | 148 | 1 | 3 | 460 |

不一致率= 23.9%

表9 WEBと質問紙における性別のクロス表

| WEB性別 | 質問紙性別 | | |
|-------|-------|-----|-------|
| | 男性 | 女性 | Total |
| 女性 | 218 | 0 | 218 |
| 男性 | 0 | 242 | 242 |
| Total | 218 | 242 | 460 |

不一致率= 0.0%

表10 WEB年齢と質問紙年齢の組み合わせ

| | |
|-----------|-----|
| 一致 | 437 |
| 質問紙=WEB+1 | 23 |
| 質問紙欠損 | 0 |
| 合計 | 460 |

不一致率= 5.0%

で入力ミスを発見し訂正しているからである⁷⁾。また、わかりづらい回答（たとえば選択肢の1とも2ともとれる○のつけ方をしているなど）に対しては、入力者が作業監督者に判断を仰ぐような体制になっているが、これらの基礎項目においては、そのような例は報告されていない。

また2)の回答者の変化は限定的には起こりうることである。特に表11に示した年齢は、ウェブ回答時から一つ年齢が上がっているケースが一定数いるのはむしろ自然なことである。また、この期間に転職をしたり、職場で従業上の地位が変化したりすることも考えられる。しかしそれらがどの程度の頻度で起こり、表7, 8の不一致率が自然なものなのかの判断は困難である。2013年度の調査で従業上の地位の不一致率は13.6%であり、(メカニズムはともかくとして)不一致の生じる傾向に大きな差は生じていない。

表8の学歴の不一致は昨年度(11.5%)と比較して大幅に増加しているが、これは3)質問項目の違いによるものと考えられる。具体的に、Web質問と質問紙ではそれぞれ以下のように尋ねている。

WEB：あなたが最後に卒業した学校を、以下の中から1つ選んでください。

質問紙：あなたが最後に通った学校(中退も含む)はどれですか。

この2つ(下線強調は筆者)を見て分かるように、質問紙では中退も卒業と同じ扱いで含んでいる。日本の学校中退率が20%よりもはるかに低く、表8の下三角セルにも度数が発生していることからすべての原因をこれに期すことはできないが、理由の一つとして中退者による回答のズレは考えられる。

残りの4) 5) 6)は回答者に起因するものである、ただしこれらは特定することまたは調査者がコントロールすることは困難である⁸⁾。

さいごに

本調査の基本的な特性を分析したところおおむね調査計画段階での割り当ては成功したといえる結果となった。しかし、本実習に関して重要な変数である職業に関していえば、依然偏りを持った状態であった。本実習における調査は今年度で5回目となり、調査計画

や調査票設計にも過去の教訓から学んだ改善を施している。回収に関しては著しい返送傾向の違いは見られなくなったものの、いくつかの属性に無視できない分布の偏りが存在していた。この傾向は例年同様であり、サンプリングをインターネットモニターに頼っている調査の宿命なのかもしれない。本調査の第一義的な目的は調査主体である本実習の受講者に対し、社会調査の実査に触れさせることにある。そのためある程度のサンプル偏りやエラーは、「必要な失敗」として学生の教育効果を持ちうる。しかしながら、社会調査そのものが現在置かれている状況を省みると、授業における調査といえいたずらに質の悪い調査または調査データを作ることにも慎重にならなければいけない。

さいごに本年度実習を通じての課題をまとめる。まず第1に、標本の設計と抽出である。これは2011年度から共通の課題であり（三輪2012）サンプリングは基本的に教員と調査会社の調整のもとで行われ、学生は「20~39歳の非学生500名」という計画サンプルを所与としたうえで研究を企画する。学生らが標本を設計、または抽出する作業はなく、それゆえ彼らが本調査におけるこれらの「失敗」を意識する機会が乏しい。さらに、調査対象者を主体的に選んでおらず、自ら調査対象者のもとに赴いて調査依頼をしているわけではないので「回収率をどう上げるか」という現在の社会調査における大きな課題を意識する機会にも恵まれない。第2に、コーディングがあげられる。本調査では調査票を見て分かるように、自由記述欄がいくつかあるものの、それらは選択肢カテゴリの中に入らない人の補足的な情報として用いているものがほとんどである。授業時数の都合上、短時間でデータ入力、クリーニングを行うことが求められ、膨大な作業量を要する自由記述に対するアフターコーディングはなるべく行わないという方針によるものである。これによって学生たちが調査データを整える作業として非常に重要な位置を占める部分を意識せずに終わってしまうことが危惧される。

以上のような課題を含みつつではあるが、本実習では、学生に対して問いの設定からレポート作成、プレゼンテーションに至るまで幅広いスキルを涵養する機会となり、また本調査は全国の若年層に対して青少年期の生活環境、自身・家族の学歴や職業、教育・健康・職業に対する意識などをとらえた有意義な調査であったといえる。石田（2013）が指摘するように、この手の調査データは授業の終了とともにほとんど「封印」され、二度と使われなくなるというのは非常に惜しまれる。これらのデータは、社会科学的な命題を導出するにはあまりにも心許ないデータであることは言うまでもない。しかし、過去のデータの蓄積を観れば、教育現場における社会調査法の教授過程でどのような問題が認識され、どのように克服しようとしてきたか（してこなかったか）または克服できた（できなかったのか）を発見する良い材料となる。これまでの蓄積も踏まえて、本調査データを有効に活用する方法を考えていくという課題は、本年度も共通である。

【注】

- 1) 本年度は諸事情により、実習非受講者が作成した質問項目を調査票に組み込んだため、昨年度と比較して質問項目の量は多くなっている。ただし、それらの項目の入力は学生に負担を強いていない。
- 2) さらに言えば、モニター登録者のうち、素早く対象者としての条件を理解し、応募するような特性を持った人である。
- 3) 学歴の分布は、質問紙における回答の分布ではなく、有効回答者の、ウェブ質問における学

- 歴を使用した。その理由は、4 節で詳しく述べるように、ウェブ質問は卒業したという意味での学歴を用いているため、マクロ統計の定義と整合的だからである
- 4) 事前の質問には含まれていなかったが、後から「労務職」を新しいカテゴリとして追加したが、本節のまとめにおいては「生産現場・技能職」にまとめている。
 - 5) 割り当てインデックスを用いれば、性別×年代×学歴の交互作用を検討したことにひとしくなるが、表 5 で確認したように、いくつかの割り当てにおいては完全な識別となってしまうため、それぞれの特性を独立であると仮定して投入した。また、働き方の項目においても無職は回収率が 100%となっている。無職者をケースから除外するとサンプルサイズが大幅に減ってしまうため、その他を同一カテゴリに含めて推定した。
 - 6) 本年度のウェブ調査はこの 4 項目に加えて「全体として、あなたは普段どの程度幸福だと感じていますか」という 11 点尺度の幸福度を尋ねている。この項目は、幸福度の個人内変動を考慮して調査票に先立って Web で質問したものであり、Web 回答と調査票回答が一致することを志向していないので、分析には含んでいない。
 - 7) ただし、質問紙の作り方や入力の手法に体系的なミスを生じさせるような構造的な問題がある場合はその限りではない。例えば、本実習では多重回答項目の前後に空白セルとなるようなデータの配列において、空白とすべきセルに 1 や 0 が入力されているケースがクリーニングの際に複数発見されている。
 - 8) とはいえ、質の良い調査データを作成するためには取り組まなければならない問題でもある。特に職業に関する質問などは、回答者が容易にこたえられるように質問文や選択肢に工夫をしなければならない。本調査では国勢調査を始め多くの社会調査が用いている職業分類を用いたが、回答者によってはこの分類自体が回答困難であるという可能性も否定できない。

【文献】

- 石田賢示, 2013, 「調査データの特性に関する基礎分析」東北大学教育学部・教育政策科学研究室編『平成 24 年度東北大学教育学部・教育学実習「統計的調査実習」報告書』: 1-7.
- 濱本真一, 2014, 「調査設計とデータ特性に関する基礎分析」東北大学教育学部・教育政策科学研究室編『平成 24 年度東北大学教育学部・教育学実習「統計的調査実習」報告書』: 1-8.
- 林英夫, 2004, 『郵送調査法』関西大学出版部.
- 三隅一人・三輪哲, 2008, 「2005 年 SSM 日本調査の欠票・回収状況の分析」三輪哲・小林大祐編『2005 年 SSM 日本調査の基礎分析——構造・趨勢・方法』2005 年 SSM 調査シリーズ 1, 2005 年 SSM 調査研究会: 17-29.
- 三輪哲, 2012, 「調査設計とデータの特性」東北大学教育学部・教育政策科学研究室編『平成 23 年度東北大学教育学部・教育学実習「統計的調査実習」報告書』: 1-7.