

調査設計とデータの特性に関する基礎分析

東山楓佳・大堀和人・藤井竜哉
(東北大学大学院教育学研究科)

1 はじめに

本稿の目的は、令和4年度(2022年度)の東北大学教育学部開講科目「教育学実習(社会調査の理論と実践)(以下、本実習)」において実施した調査設計の提示と得られたデータの基礎的な特性(学歴や性別などの基本属性により回答の有無に偏りがあるかないか)を分析することである。本実習では、社会調査の手続きにおける、調査企画・調査票設計・標本抽出・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・報告書執筆の過程を、受講者が実際に体得・会得できるように授業が編成されている。近年の教育学実習の報告書では、2020年までは、受講者は2つのグループに分かれていた(藤井ほか 2017; 伊藤・廣谷 2018; 米田・松野 2019; 藤井ほか 2020)。だが、一昨年度からはグループを分けず、より標本規模を大きくして調査を行っている(神山ほか 2021; 藤井ほか 2022)。今年度の調査についても昨年度同様の標本規模となっている。本稿では実施した調査の設計について示すとともに、調査から得られたデータに偏りがあるか、あるとすればどのような偏りがあるかを示す基礎的な分析を行う。

2 調査の設計について

まず調査の設計を示す(概要は表1)。本実習における調査は若年層を対象としてその職業や最終学歴に関するデータ収集を目的としている。そのため対象者は学生を除く20歳以上40歳未満の男女に設定されている。また、一昨年度まで2つのグループに分かれ2種類の調査票を300部ずつ計600部の調査票を送付していたが、一昨年度から1種類の調査票を600部送付しており、今年度も一昨年度、昨年度と同様の方法をとっている。

表1に示した調査の設計の調査項目について述べる。調査項目は年齢・性別といった基本属性項目、職業に関する項目、孤立・孤独状況に関する項目、政治への意識・関心に関する項目、これまでに受けた地域教育に関する項目、災害の被伝承経験に関する項目、「生きる力」に関する項目が盛り込まれている。

有効回収率については表1に示した通り89.3%となっており、昨年度よりも若干低いものとなった。ただし、返送がなかった64票のうち29票がそもそも調査対象者へ不達であり、例年よりも多く不達であった。ちなみに過去の通年での実習における回収率を付記しておく、2016年度が70.3%(藤井ほか 2017)、2017年度が78.1%(伊藤・廣谷 2018)、2018年度が90.7%(米田・松野 2019)、2019年度が91.3%(藤井ほか 2020)2020年度が

89.5%（神山ほか 2021）, 2021 年度が 90.2%（藤井ほか 2022）である。先までに述べた通り 2 種類のグループには一昨年度から分かれていないため, 2016 年度から 2019 年度までの結果とそれ以降の年度では, 単純な比較はできない点は注意が必要である。

次に本実習の調査スケジュールについて述べる。表 2 に示したのが調査スケジュールである。本実習においては調査を実施するにあたって前期までに調査の実施に関する内容, 後期においては実施した調査の分析に関する内容を授業内にて行っている。その中で前期においては調査企画, 調査票設計などの 1 節に示した内容を実際に体験してもらっている。後期において学生たちは分析方法を学ぶとともに, 収集したデータを分析している。

表 1 調査の設計

調査名称	若者のライフスタイルと意識に関する調査
調査対象	1) 母集団：学生を除く日本全国の 20 歳以上 40 歳未満の男女 2) 標本規模：600 3) 標本抽出法：上記の条件を満たすインターネット調査モニターのうち協力依頼に応じた者から年齢, 性別, 学歴による層化無作為抽出。年齢と学歴は最新の国勢調査を用いて層化。
調査期間	2022 年 8 月 10 日～9 月 10 日
調査協力機関	楽天インサイト株式会社
調査項目	・基本属性項目 ・職業に関する項目 ・孤立・孤独状況に関する項目 ・政治への意識・関心に関する項目 ・これまでに受けた地域教育に関する項目 ・災害の被伝承経験に関する項目 ・「生きる力」に関する項目
有効回収票(率)	536 (89.3%)

表 2 調査スケジュール

年月	授業・調査実習の内容
2022 年	4 月 14 日 実習授業開始
	6 月 15 日 調査会社へ依頼
	7 月 15 日 質問項目締切
	7 月 21 日 事前インターネット調査開始
	7 月 26 日 事前インターネット調査終了
	7 月 29 日 調査協力者リスト納品
	7 月 27 日 調査票最終確認
	8 月 2, 9 日 調査票封入・投函
	9 月 10 日 調査票回答締切
	9 月 20 日 データ入力・クリーニング開始
2023 年	1 月 12 日 最終報告会
	2 月 15 日 最終レポート締切

3 欠票の分析

ここでは、本調査において、調査対象者 600 名のうちの返答者の属性の特徴を分析する。なお、調査対象者の属性については、事前のインターネット調査によって集計されたデータを使用する。

まず、今回の調査で対象となったサンプル全体の記述統計量と、そのうち質問紙に対し返送があったものの記述統計量は以下の通りとなった（表 3）。

表 3 記述統計

	全体				返送あり			
	mean	min	max	sd	mean	min	max	Sd
年齢	31.4	20	39	5.1	31.5	20	39	5.0
	%				%			
性別								
男性	50.0				49.8			
女性	50.0				50.2			
学歴								
中学校	4.7				4.3			
高等学校	22.8				23.5			
専修・高専・短大	29.8				29.1			
大学・大学院	42.7				43.1			
雇用形態								
正規雇用	60.3				60.3			
非正規雇用	20.0				19.8			
自営業・家族従業者・内職	7.0				6.9			
無職・その他	12.7				13.1			

次に、調査対象者の属性の違いによって返送の有無に偏りがあるのかを検討する。そのためにまず返送の有無と「年齢」、「性別」、「学歴」、「雇用形態」のそれぞれの関連を確認する。

まず返送の有無と「年齢」の関連を確認するため、相関係数を確認すると、相関係数は 0.057、有意確率は 0.16 という結果になった。結果より、返送の有無と年齢には関連がないということが明らかとなった。

次に返送の有無と「性別」、「学歴」、「雇用形態」について、カイ二乗分析によって関連を確認する。結果を先に述べれば、「性別」、「学歴」、「雇用形態」のいずれにおいても返送の有無との有意な関連を示したものはなかった。以下、詳述する。

返送の有無と「性別」については、以下の結果を得た（表 4）。男性の場合、返送なしは 11.0%、返送ありは 89.0% となっており、女性の場合には返送なしが 10.3%、返送ありが 89.7% となっていた。カイ二乗検定の結果は先にも述べた通り、10% 水準でも有意な結果となっておらず、返送の有無と「性別」との間に関連は見られなかった。

次に返送の有無と「学歴」についての結果を確認する（表 5）。最終学歴が「中学校」の場合、返送なしは 17.9%、返送ありが 82.1%、「高等学校」の場合、返送なしが 8.0%、返送あり 92.0%、「専修・高専・短大」の場合、返送なしが 12.8%、返送ありが 87.2%、「大学・

大学院」の場合、返送なしが 9.8%，返送ありが 90.2%となっていた。カイ二乗検定の結果は有意な結果となっておらず、返送の有無と「学歴」との間に関連は見られなかった。

次に、返送の有無と「雇用形態」の関連についての結果を確認する（表 6）。雇用形態が「正規雇用」の場合、返送なしは 10.8%，返送ありが 89.2%，「非正規雇用」の場合、返送なしが 11.7%，返送あり 88.3%，「自営業・家族従業者・内職」の場合、返送なしが 11.9%，返送ありが 88.1%，「無職・その他」の場合、返送なしが 7.9%，返送ありが 92.1%となっていた。カイ二乗検定の結果は有意な結果となっておらず、返送の有無と「雇用形態」との間に関連は見られなかった。

以上の分析から、年齢、性別、学歴、雇用形態のいずれについても、返送の有無とは関係が見られなかった。しかし、以上の分析は単独の変数のみに着目しており、その他の変数を考慮していない。そこで、その他の変数を考慮した場合に同様の結果になるかどうかを検証するため、従属変数を返送の有無（返送なし=0，返送あり=1）とした二項ロジスティック回帰分析を行った。結果は、表 7 の通りである。

表 4 性別と返送の有無の関連

	返送の有無	
	なし	あり
男性	33 (11.0)	267 (89.0)
女性	31 (10.3)	269 (89.7)

$\chi^2=0.07$ $p=0.791$

表 5 学歴と返送の有無の関連

	返送の有無	
	なし	あり
中学校	5 (17.9)	29 (82.1)
高等学校	11 (8.0)	126 (92.0)
専修・高専・短大	23 (12.8)	156 (87.2)
大学・大学院	25 (9.8)	231 (90.2)

$\chi^2=3.63$ $p=0.304$

表 6 雇用形態と返送の有無の関連

	返送の有無	
	なし	あり
正規雇用	39 (10.8)	323 (89.2)
非正規雇用	14 (11.7)	106 (88.3)
自営業・家族従業者・内職	5 (11.9)	37 (88.1)
無職・その他	6 (7.9)	70 (92.1)
$\chi^2=0.81$ $p=0.847$		

表 8 返送の有無を従属変数とした二項ロジスティック回帰分析の結果

	Coef.	S. E	Exp
切片	0.01	1.00	2.49
年齢	0.04	0.03	1.04
性別 (ref. 男性)			
女性	0.14	0.29	1.15
学歴 (ref. 大学・大学院)			
中学校	-0.90	0.57	0.41
高等学校	2.65	0.39	1.30
専修学校・高専・短大	-0.31	0.31	0.74
雇用形態 (ref. 正規雇用)			
非正規雇用	-0.20	0.36	0.82
自営業・家族従業者・内職	-0.18	0.52	0.84
無職・その他	0.42	0.51	1.52
-2LL	399.90		
N	600		

結果を確認すると、どの項目においても統計的有意差は認められなかった。つまり本調査においては返送の有無に年齢、性別、学歴、雇用形態による偏りは生じていないということになる。

4 さいごに

本実習の欠票の分析を行ったところ返送の有無に年齢、性別、学歴、雇用形態による偏りは生じていないことが分かった。昨年度の欠票の分析では、返送の有無に年齢による偏りがあったことが報告されている。先行研究でも年齢が未提出に関する重要な要素であることが明らかにされている（埴淵・山内 2019 など）が、今年度の分析では年齢による返送の偏

りは見られなかった。

最後に本実習の課題を挙げる。本実習の課題としては一昨年度、昨年度の報告書（神山ほか 2021；藤井ほか 2022）で述べられた調査方法の再考，過去データの活用，回収率に関する次年度以降へのフィードバックといった 3 点は今年度も挙げられる。今年度の基礎分析の結果から考えられる課題に絞り述べると，居住地や住居形態などの他の要因も考慮し標本を抽出することが挙げられる。先行研究による欠票分析では，個人が居住している地域や住居形態も回収率と関連していることが明らかにされている（松岡ほか 2015 など）。今回の分析で使用できたデータは年齢，性別，学歴，雇用形態に限られており，その他の変数が返送の有無に影響を与えていた可能性も十分考えられる。また，学生たちの実施した質問の中に都市出身か地方出身かを聞く質問があったことも含めて考えれば都市と地方の返送の有無が本調査の分析に影響を与えていることもあり得る。そのため，今後の課題としては，事前調査で居住地域や居住形態などの項目を含め標本抽出を行う必要があると考える。

〔文献〕

- 藤井奈々子・中島日向子・廣谷貴明，2017，「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 28 年度東北大学教育学部・教育学実習「社会調査の理論と実践」報告書』：1-7.
- 藤井竜哉・遊佐賢・伊藤愛莉，2020，「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『令和元年度東北大学教育学部・教育学実習「社会調査の理論と実践」報告書』：1-7.
- 埴淵知哉・山内昌和，2019，「国勢調査『不詳』発生に関連要因——インターネット調査を用いた未提出者の分析」『E-journal GEO』14(1)：14-29.
- 伊藤愛莉・廣谷貴明，2018，「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 29 年度東北大学教育学部・教育学実習「社会調査の理論と実践」報告書』：1-8.
- 神山真由・菅原純一・藤井竜哉，2021，「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『令和 2 年度東北大学教育学部・教育学実習「社会調査の理論と実践」報告書』：4-13.
- 米田佑・松野広，2019，「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 30 年度東北大学教育学部・教育学実習「社会調査の理論と実践」報告書』：1-6.
- 藤井竜哉・上野春実・熊夢詩，2022，「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『令和 3 年度東北大学教育学部・教育学実習「社会調査の理論と実践」報告書』：1-6.
- 松岡亮二・前田忠彦，2015，「「日本人の国民性第 13 次全国調査」の欠票分析：個人・地点・調査員の特性と調査回収状況の関連」『統計数理』63(2)：229-242.