

調査設計とデータに関する基礎分析

神山真由・菅原純一・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2年度(2020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受講者は2つのグループに分かれていた(藤井奈々子ほか 2017; 伊藤・廣谷 2018; 米田・松野 2019; 藤井竜哉ほか 2020)。しかし、今年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。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)。本実習においての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している。また、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、本実習では1種類の調査票を600部送付している。表1に示した調査の設計の調査項目については、年齢・性別といった基本属性項目、職場環境に関する項目、部活動・サークルに関する項目、パーソナリティに関する項目、兄弟姉妹に関する項目、態度(政治的なものなど)に関する項目、COVID-19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91.8%となっている¹⁾。通年で実習を行うようになってから最も高い回収率である。過去の通年での実習における回収率を付記しておくと、2016年度70.3%(藤井奈々子ほか 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% (藤井竜哉ほか 2020) である。ただし, 先までに述べた通り 2 種類のグループには今年度は分かれておらず, その点に違いがあることは述べておく。

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。表 2 に示されるように今年度は COVID-19 の影響により, 大学自体が始まる時期が遅れ, 調査に取り掛かったスケジュールは例年よりも後ろ倒しのものとなっている。

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。
調査機関	2020 年 8 月 5 日～9 月 15 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 ・職場環境に関する項目 ・部活動・サークルに関する項目 ・パーソナリティに関する項目 ・兄弟姉妹に関する項目 ・態度に関する項目 ・COVID-19 に関する項目
有効回収票(率)	551 (91.8%)

表 2 調査スケジュール

年月		授業・調査実習の内容
2020 年	4 月 28 日	実習授業開始
	7 月 6 日	調査会社へ依頼
	7 月 22 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 28 日	調査協力者リスト納品
	7 月 31 日	質問項目締切
	8 月 1 日	調査票最終確認
	8 月 5 日	調査票封入・投函
	9 月 15 日	調査票回答締切
	9 月 21 日	データ入力・クリーニング開始
2021 年	1 月 12 日, 19 日	最終報告会
	2 月 18 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、ここでは、調査対象者の属性として、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）。返送ありの調査対象者は、調査対象者全体と比較して、年齢、性別、学歴、従業上の地位・雇用形態に若干の違いが見られる。

表 3 記述統計

	全体	返送あり
	%	%
年齢		
20代	38.7	37.7
30代	61.2	62.3
性別		
男性	50.0	49.5
女性	49.8	50.5
学歴		
中学校	5.3	4.7
高等学校	24.7	24.3
専修・高専・短大	37.0	38.3
大学・大学院	32.8	32.7
従業上の地位・雇用形態		
正規雇用	59.3	59.9
非正規雇用	19.0	18.5
自営業・家族従業者・内職	5.7	5.3
無職・その他	15.8	16.3

次に、このような返送の有無による調査対象者の属性の違いが、母集団においても当てはまるのかを検討するために、返送の有無と年齢、返送の有無と性別、返送の有無と学歴、返送の有無と従業上の地位・雇用形態について、それぞれカイ二乗分析を行った。

まず、返送の有無と年齢については、以下の結果を得た（表 4）。

表 4 年齢と返送の有無のカイ二乗分析の結果

	返送の有無		χ^2
	なし	あり	
20代	24	208	2.792 †
	10.3	89.7	
30代	24	343	
	6.5	93.5	

上段：実数, 下段：パーセント

†p<0.1

カイ二乗検定の結果、年齢と返送の有無に 10%水準で有意な関係が認められた。20代で返送なしの者が 10.3%, 返送ありの者が 89.7%, 30代で返送なしの者が 6.5%, 返送ありの者が 93.5%となっており、30代の方が20代よりも返送する傾向が高いといえる。

次に、返送の有無と性別についてカイ二乗検定を行い、以下の結果を得た（表 5）。

表 5 性別と返送の有無のカイ二乗分析の結果

	返送の有無		χ^2
	なし	あり	
男性	27	273	0.794
	9.0	91.0	
女性	21	278	
	7.0	93.0	

上段：実数, 下段：パーセント

カイ二乗検定の結果、性別と返送の有無に有意な関係は認められなかった。
続いて返送の有無と学歴についてカイ二乗検定を行い、以下の結果を得た（表 6）。

表 6 学歴と返送の有無のカイ二乗分析の結果

	返送の有無		χ^2
	なし	あり	
中学校	6 18.8	26 81.3	8.343 *
高等学校	14 9.5	134 90.5	
専修・高専・短大	11 5.0	211 95.0	
大学・大学院	17 8.6	180 91.4	

上段：実数, 下段：パーセント

* $p < 0.05$

カイ二乗検定の結果, 年齢と返送の有無に 5%水準で有意な関係が認められた。標本では, 学歴が中学校の者で返送ありが 81.3%, 学歴が高等学校の者で返送ありが 90.5%, 学歴が専修・高専・短大の者で返送ありが 95.0%, 学歴が大学・大学院の者で返送ありが 91.4%, であり, 学歴が高くなるにつれて返送する者の割合が高くなる傾向にある。

最後に, 従業上の地位・雇用形態と返送の有無についてカイ二乗検を行い、以下の結果を得た (表 7)。

表7 従業上の地位・雇用形態と返送の有無のカイ二乗検定の結果

	返送の有無		
	なし	あり	χ^2
正規雇用	26	330	4.261
	7.3	92.7	
非正規雇用	12	102	
	10.5	89.5	
自営業・家族従業者・内職	5	29	
	14.7	85.3	
無職・その他	5	90	
	5.3	94.7	
上段：実数, 下段：パーセント			

カイ二乗検定の結果、性別と返送の有無に有意な関係は認められなかった。以上の結果をまとめると、性別と学歴が、返送の有無と有意な関連を示した。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（返送あり=1，返送なし=0）とした二項ロジスティック回帰分析を行った。結果は、表8の通りである。

表 8 二項ロジスティック回帰分析の結果²⁾

	Coef.	S.E.	Exp(Coef.)
切片	2.344 **	0.392	10.43
年齢 (ref. 20代)			
30代	0.474	0.306	1.61
性別 (ref. 女性)			
男性	-0.285	0.336	0.75
学歴 (ref. 大学・大学院)			
中学校	-0.759	0.540	0.47
高等学校	-0.023	0.391	0.98
専修学校・高専・短大	0.675	0.412	1.96
従業上の地位・雇用形態 (ref. 正規雇用)			
非正規雇用	-0.582	0.403	0.56
自営業・家族従業者・内職	-0.829	0.551	0.44
無職・その他	0.247	0.542	1.28
-2対数尤度	319.411		
N	599		

**p<0.01

結果を確認すると、切片のみが1%水準で有意であり、その他は10%水準においても有意な差は確認されなかった。以上のことから本調査において、調査対象者の調査票返送の有無は回答者の年齢、性別、学歴、従業上の地位・雇用形態と関連性がないことが明らかとなった。

4 さいごに

本実習に残された課題としては、3点あげられる。

第一に、現在行っている紙媒体による調査方法についての再考である。紙媒体による調査票を印刷し、製本し、郵送すること、また返送された調査票をまとめ、記載内容を手作業でデータ化することも本実習の大切な学びの一部ではあるが、一方で、昨今、インターネットを介した調査が主流になりつつある現状も考慮しなければならな

いだろう。そうした実態から、本調査においても自由記述欄をつうじて、インターネットを介した調査を望む声が複数寄せられている。もちろん、インターネット環境が整備された人だけに調査依頼することには調査を進めるうえでの問題がともなうが、総務省（2020）の報告によると、2019年度のパソコン保有率が69.1%、モバイル端末全体では96.1%となっている。そうした調査結果からも一考を要する案件だと考える。さらにCOVID-19の蔓延がもたらした社会変動や人々の意識変容も見据えながら、検討する必要があるだろう。

第二に、過去のデータの活用である。これまでも本実習において、過去のデータの活用が課題であることは指摘されてきた（下瀬川・池田 2015、藤井奈々子ほか 2017、伊藤・廣谷 2018）。しかし毎年、受講生の関心に沿って調査票を設計・作成するため、意図的に質問を組み入れない限り、時系列的な分析は難しいのが現状である。よって、TA主導で定点観測を可能とする質問を決めておき、調査票に組み入れて時系列的に分析することが、本実習の調査をさらに生かすことにつながるものと考えられる。

第三に、回収率に関する時系列分析を実施し、調査方法にフィードバックすることである。各年度の個票データを入手できなくても、過去の報告書を活用して分析を行えば、調査に利するデータを得られる可能性がある。特に回収率に関しては、前述したとおり、改善傾向にある。しかし調査協力者頼みの調査方法では、回収率を維持するのは困難であるため、方策を講じる必要がある。たとえば、回答率を高める方法としては5つの要因が存在するとされる（吉村 2017）。その5つは①事前挨拶状の送付または電話による事前連絡 ②個人名をあらかじめ特定して接触する ③謝礼の提供 ④催促状（郵送調査の場合）や拒否後の再接触（翻意を依頼）⑤見やすく答えやすく工夫された調査票、である（吉村 2017: 81）。こうした方法をふまえ、調査協力を得られやすい体裁を整えるなど、調査の精度を保つうえでの努力は必須のことと考える。

以上のような課題はあるものの、本実習では、受講生は社会調査の理論、統計的手法の修得だけではなく、テーマ設定や仮説構築、データ分析、報告書執筆、さらには、調査票の作成、封入・投函作業、データ入力といった調査の実務も行うことができた。今後も社会調査の理論と実践を体得するより良い実習となるよう、TA間での引継ぎ等をつうじて、課題を克服していく必要がある。

[注]

- 1) 調査対象者 600 のうち、551 が有効票、1 が無効票となった。3 節における分析はその無効とした 1 名分を欠損値として扱い、分析がなされている。
- 2) 年齢、性別、学歴、従業上の地位・雇用形態について多重共線性を考慮し、相関などを確認したが強い関連は見られなかった。

[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.
- 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.
- 下瀬川陽・池田岳大，2015，「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 27 年度東北大学教育学部・教育学実習「統計的調査」報告書』：1-6.
- 総務省，2020，『令和元年通信利用動向調査の結果』（2021 年 3 月 22 日取得，https://www.soumu.go.jp/johotsusintokei/statistics/data/200529_1.pdf）
- 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.
- 吉村治正，2017，『社会調査における非標本誤差』東信堂.