

マイクロデータの入手と教育・研究への応用

佐藤 博樹

1. 東大文学部の社会調査実習の事例

東京大学社会科学研究所の佐藤博樹です。お手元のレジュメにありますように、「マイクロデータの入手と教育への活用」というテーマで30分ほど報告をさせていただきます。私がお話しさせていただきます内容は、昨年4月からピンチヒッターとして東京大学の文学部の武川正吾助教授と一緒に社会調査実習の授業を持たせて頂いています。午前中のセミナーでは、データセットをどう収集し、二次分析のために利用できるようにする仕組みをどのように作っていくのか、そういう議論があったわけですが、そうした既存のデータセットを教育に使うことをやってみたわけです。私は、盛山先生と違ひまして、社会学の分野で社会調査をきちっとやってきた研究者ではありませんので、どのように社会調査を教えればいいのかよくわからない点も少なくなかったのですが、私なりに少し考えたことがありますので、その経験をお話ししたいと思います。それが前半です。後半では、社会調査実習などの教育に使えるデータセットとしてどのようなものがあるかということについて、お話しできればと考えております。

社会調査で何を教えるべきかということなのですが、社会調査にもいろいろあり、ヒアリングなどによる事例研究や調査票による大量調査などがあるわけですが、今回、授業でやりましたのは、調査票による大量調査の具体的な実施方法を教えることです。サンプリング方法ですとか、調査票、設問の作り方な

どに関しても教科書があるわけですが、今回は教科書を使って教えるのではなく、仮説を立てそれに即した調査票を作り、調査を実施し、データを分析するという一通りの調査の流れを経験してもらうことをやろうと考えました。ところが、学部生がメインの授業ですので、“自分の問題関心に即して仮説を設定し、その仮説を検証するためにアンケート調査票を設計し、回答結果を分析していく”などということは直ぐに自分たちではできないわけです。学生は、大学院生を入れて40人弱でしたが、家族に関心がある人もいれば、犯罪に関心がある人もいれば、教育について関心がある人もいます。いろんな分野に関心がある学生がいる。そういう異なる関心を満たす調査票を作って調査を実施させる必要がある。どうしたかといいますと、まず大量アンケート調査を分析した調査報告書を取り上げ、どのように仮説を立て、設問をどのように分析しているかということ、とりあえず学んでもらうことをしました。その時、家族、犯罪、教育などいろんな分野やテーマに関心を持ってる学生がいますから、できるだけいろんなテーマや関心に対応できるようなデータセットを取り上げて、その調査報告書を読ませることを考えたわけです。

それでイギリスで実施されています British Social Attitudes Survey を取り上げたわけです。ご存じの方も多いかと思いますが、資料の最後の方にこの調査の紹介がありますが、1983年から毎年行われている調査です。いわゆるインタビュー調査と自記式の調査を

組みあわせたもので、インタビューは1時間ちょっとだったと思います。British Social Attitudes Survey は、時系列に比較できる設問とそれぞれの調査年のモジュールつまり特定のテーマを決めた設問が含まれています。調査の分野は広範囲にわたり、相当幅広いテーマを取り上げています。この2年度分の報告書を読むことにしました。具体的には、関心のあるテーマに即して学生をチームに分け、例えば、犯罪、教育などですね、そして各チームの関心のあった報告書の章を読ませ、報告させる。British Social Attitudes Survey の報告書をですね。その際、ここに疑問があるとか、もう少しこういう分析ができるんじゃないかとか、あるいはこういう仮説が立てられるんじゃないかというような、報告書を読みながら自分たちの仮説を作らせる。各グループ7、8人くらいですが、そのグループの関心に近い報告書のレポートを読ませて報告させ、疑問点の抽出や仮説の構築みたいなものをやらせたわけです。その上で British Social Attitudes Survey のデータセットを取り寄せてありましたので、データセットは SPSS のファイルになっていましたが、データセットを各グループに渡して、自分たちの疑問や新しく立てた仮説に即して再分析をするということをやったわけです。

既存の報告書を読んだ上で、自分なりの問題関心を持って既存のデータを分析する、そういうことをやることによって、社会調査によって得られたデータをどのように分析していくのか、どういうようなことが分かるのか、ということを多少なりとも理解させることが第一段階の作業でした。それからさらに、イギリスの設問を念頭に置きながら、日本に置き直したときに、どういうテーマが立てられるのか、そういうことを議論したわけです。それに平行して British Social Attitudes Survey の膨大な調査票を学生に翻訳させました。全部です。調査票を翻訳するというこ

とを通じて、調査票の設計の仕方ですね、どのように調査票が設計されているのかを理解させることをしたわけです。二次分析の際に、調査票の翻訳がある方が便利だろうと考えたこともあります。こういう手続きを踏んだ後で、British Social Attitudes Survey との比較を一部では可能とし、日本の状況に即した設問を設計し、日本で実施する調査票を作って、調査を実施したわけです。事前に英語の調査票の翻訳をさせてあったため、日本で実施する調査票の作成では、我々教員が手を出さなくてもかなりの完成度のものができました。

調査は、文京区の選挙人名簿から500名をサンプリングし、360名程度を回収しました。調査票をコーディングし、入力し、SPSS のファイルにして、それを各グループに戻して、当初の問題関心に即して分析させました。British Social Attitudes Survey との比較を念頭に置いて作られた部分もありますので、国際比較を一部では行いました。

つまり、テーマを立て、そのテーマの仮説を実証する。仮説を実証するためにどういう操作変数を設定し、それをどうやって質問票に起こしていくかということを、既存の調査を分析することを通じて理解させ、それを前提に日本での調査票を設計し調査票を作るといように二段階で社会調査実習を進めたわけです。

2. 実習時の課題

以上のような授業を進める上で、課題として感じた点は次のようなものでした。これを次にお話したいと思います。British Social Attitudes Survey は、エセックスのデータアーカイブから入手することができます。それは教育用にも使えます。ただ、教育に使う場合は、学生が第三者にデータセットを渡さないようなどの指導が教員の責任となります。もちろん学生から「第三者にデータを

提供しません」といった誓約書を取るわけですが、教員も指導の義務を負うわけです。日本には、これまで既存のデータセットの二次分析の慣行がありませんでしたので、学生に対してデータセットの利用ルールを徹底することが、非常に大事だと思いました。

それともう一つ、学生がデータセットを誤解しないで使っているかどうかの確認が難しいことです。既存データの利用にはいろいろな指示があります。例えば、復元してデータを使えと書いてあるんですね。ですから、復元したデータを使わなくてはならない。じゃあ、学生の全員がそれをやっているかどうかですね。各グループには院生を張り付けてありましたので、院生に確認するようには言いましたが、教員がすべてを直接確認するのは極めて難しいです。もちろんこれは学生の数にもよると思いますが。

三つ目は、東大のSPSSの数が足りないので、かなり苦労しました。SASは入っているんですが、はじめて統計処理を行う多数の学生に教えるのは、SPSSのほうが指導が楽ですし、取り寄せたデータセットもSPSSファイルだったため、SPSSを選択しました。SASよりもSPSSのメニュー画面で処理する方が、初心者には使いやすいと思います。既存データを分析するためのとっかかりとしては、よかったのではないかと思います。二次分析を行う場合、どの程度学生が汎用パッケージを使えるか、ということがかなり大事になります。

3. 教育・研究で使用できるマイクロデータ

第3は、社会調査実習などの教育に使いやすい汎用データソフトとして何があるのかということです。家族とか、教育とか、社会移動とかテーマを特定化した授業の場合は、データセットを決めやすいのですが、いろいろな関心を持った学生に教える授業では、それ

ほど選択の幅は広くありません。学生がみんな違ったデータセットを使うことも論理的には可能ですが、それでは教員が十分指導できません。そのため一つの授業ではせいぜい2つ程度のデータセットしか取り上げられないと思います。そうした汎用的なデータセットには、前述のBritish Social Attitudes SurveyやアメリカのGeneral Social Surveyがあります。General Social Surveyは、Web上からダウンロードできます。今、ダウンロードできるのは、72年から94年までのデータセットだと思います。設問の検索もWeb上でできますので、調査の全体の構造を学生に調べさせるということもできます。

海外のデータセットを利用する場合の問題は、日本が対象に入っていないということです。我々のSSJデータアーカイブがオープンすれば、例えば生命保険文化センターの「日本人の価値観調査」などは汎用的なデータセットですので、授業にも使えるのではないかなと思っていますが、なかなか授業に使えるような公開されている日本に関するデータセットがない。ただ、日本が入っているデータセットとしてはInternational Social Survey Programmeという国際比較プロジェクトをあげることができます。この調査プロジェクトができた背景ですが、イギリスのBritish Social Attitudes Surveyのようにそれぞれの国で、汎用的な時系列のデータセットが作られています。しかし、各国のデータセットを相互に比較することがほとんどできないのです。それぞれがそれぞれの社会的なコンテキストの中でテーマを設定し調査を実施しているためです。そこで、初めから国際比較を組み込んだ調査を実施しようとの考えから始まったのがInternational Social Survey Programmeです。最近では、20数カ国が参加しています。日本からは93年から参加しています。先ほどのBritish Social Attitudes Surveyには、インタビュー調査と別

に自記式の調査票があると説明しましたが、自記式の調査票の一部が、International Social Survey Programme のための調査です。アメリカでは、GSS に ISSP の調査がっています。日本では、NHK の放送文化研究所が実施しています。ただ日本の場合、GSS とか British Social Attitudes Survey がありませんので、独自にやっているわけです。93 年以降のデータセットは、日本との国際比較が可能なデータセットになっています。このデータセットはどこから入手できるのか。一つは、ICPSR に大学が加盟すると、そこから入手できます。日本のデータセットを含めて入手できます。ICPSR に、皆さんの大学が入っていない場合は、ドイツの ZA から個人でも直接入手できます。ZA はケルンにあり、ホームページ上でデータセット申し込みます。NHK の放送文化研究所の「放送研究と調査」という雑誌に、日本の調査結果が、単純集計ですが掲載されています。ISSP のデータセットは、日本が入ったデータセットとい

うだけでなく、毎年のモジュールでテーマが変わるため、いろいろなテーマについて国際比較ができるデータですので、是非、授業に使われてみてはいかがでしょうかと思います。

最後に、日本の大学院教育と言っているのかもしれませんが、一部の分野を除いてまだ、既存のデータセットを分析して、例えば修士論文を書くというような研究スタイルが確立されていません。自分で独自に調査をしないと論文にならない。こうしたことから、主要なデータセットを紹介するセミナーを社会科学研究所の日本社会研究情報センターで開こうと考えています。いろいろなデータセットの紹介、このデータセットを使えばこんな研究ができるといったセミナーを、若手研究者向けに開きたいと考えております。

どうもあっちいたりこっちいたり、まとまりのない報告でしたが、この1年間の、社会調査実習を行った感想をお話しさせていただきました。どうもありがとうございました。

佐藤講演に対する質疑

司会(盛山)：どうもいろいろお世話になっておりました恐縮です。ありがとうございました。ご質問ありましたらどうぞ手を挙げて下さい。

小島：茨城大の小島ですが、話伺ってまして、大学生の違いだと思ったんですが、あの British Social Attitudes Survey は有名な調査なんですけれども、これは当然英語ですよ。英語の調査が使えるというのは東大の学生だからできたのではないのでしょうか。

佐藤(東大)：そうです、英語です。

石井：札幌学院の石井です。カリキュラムの内容を教えてください。

佐藤(東大)：正直言って、かなり苦しい時間割でした。1年の授業なんですけど、夏休みまでの初めの3回ぐらいが既存の論文の報告ですね。その後、どういう再分析をするかというのを報告させました。つまりデータセットをどのように再分析できるかについて3回ぐらいやりました。人数多いですから、8チームぐらいに分け、チームごとにどういう分析をするかという報告をさせました。そんな

分析しちゃダメだとか、この変数が大事ではないとか議論しました。そして2週間ぐらい空けてまた、再分析結果を報告させる。そういうことで大体夏休みになりました。夏休みの直前から調査票作りを初め、9月、10月にかけて日本での調査票を作らせました。平行して英語の調査票の翻訳もやりました。11月に調査を実施し、データを入力して、翌年の1月になってから日本のデータの分析です。前半は既存データの再分析。後半は日本での調査票の作成、実査、分析でスケジュールは相当きつかったと思っています。

田中：他の方が調査されたデータを用いて、学生諸君も含めましてそれを研究のデータとするということは非常に結構なことだと思うんですけども、そのことがなかなか十分行われないというところに、収集されたデータの客観性に関する認識の違いが各研究者の間にあるように見えるんですが。社会学者の側では、データの客観性ということに対する認識はどのような状態なんでしょうか。もちろん、データさえあればそれは客観的という意味ではなく、客観的に対象をつかむにはデータのそれなりの条件が必要だとは思いますが。そういう事柄などについて教えて頂ければと思います。

佐藤(東大)：調査の客観性ですが、まず、調査のサンプリング方法と回収率、調査の仕方の問題になります。いい加減なサンプリングや低い回収率では、精度が問題になります。それでデータセットの価値が決まってくる。

もう一つは、個人に対する調査の場合、行動を聞くのか、そのオピニオンを聞くのかでも違います。行動に関すること、例えば、買い物の回数などはある程度まで客観的に測定可能ですが、オピニオンやバリューについて

は、質問の仕方によって回答が違ってくるが多くなる。後者の場合は、質問の仕方でも回答が異なるとも言えます。客観性っていうのは非常に難しいです。

司会(盛山)：難しい問題ですね。

稲葉：今日の私の報告でも、お話ししようと思っているんですが、社会調査実習とか、社会調査法とかっていうのは必ず他の授業との関係っていうのが問題になります。特に講義の統計学をしっかりとやっているかどうか、やっぱり大きいと思うんですね。私の経験ですと、大体統計学の授業っていうのは分散分析までやらないことが多くて、やっても回帰分析くらいで終わっちゃうっていうのが多いです。チェビシェフの不等式とか大変難しいのをやるくせに分散分析の肝心なところをやってくれていないな、っていう感じが非常に強くて、社会調査の実習なんかやると、その点非常に問題になるなという印象をもっているんですが。他の授業でこういう風にやってくれてるともったいない、そういうのがありましたら。

佐藤(東大)：統計的な分析については盛山先生の授業でやられていると思います。私は部外者としてお手伝いさせていただいておりますので、社会学のカリキュラム構成は詳しく分かりません。盛山先生、いかがですか。

盛山：いや、それはぜんぜん上手くできておりませんで、そういう状況の中で佐藤先生にお願いすることになっています。この問題、大変気になっていて、日本の大学における社会調査関係の授業システムはどうも、特に文科系に関しては上手くいってないんじゃないかという感じが非常に強くて、そこら辺はまた別の機会に検討させて頂く時間があればいいかなと思っているところであります。