

# 素データを用いた社会調査実習の方法と問題

稲葉 昭英

## 1. はじめに

都立大学の稲葉と申します。私は大学の教員になって今年で9年近くなるのですが、その間ずっと社会調査実習、社会調査法といった授業を担当して来ました。担当はしてきたんですが、いつもなかなかうまくいきません。私なりにいろいろ試行錯誤して改良を重ねてはいるんですが、どうにもうまくいかないんです。うまくいかない原因というのは、まあSORDプロジェクトにも関係してくるんですが、データの公開という問題と非常に関わっている部分があって、このあたりに現在の社会調査実習の問題があるのではないかという話をしたいと思います。

今日の話は大体三つに分かれております。最初に、社会調査実習のカリキュラムの体系についてお話しします。つぎに、これまで、私自身社会調査実習に関するいろんな方法を考えてはそのつど挫折してきたということがありますが、社会調査実習のいろんな方法と、そこでの問題点についてお話しします。最終的にそうした問題を解決するためにということが望まれるのかという話をしていきたいと思います。

## 2. 社会調査実習のカリキュラム体系

まず、社会調査実習の体系ということですが、社会調査実習の目的は、今日の東大の佐藤先生の報告にもございましたが、基本的には社会調査に必要な知識と技術を実習によって身につけるといったところにあると思いま

す。ただ、これが最後の最後で非常に問題になってくるんですが、結局、日本ではデータがあまり公開されておられませんから、社会調査に必要な知識と技術というのは要するに自分でデータをとって自分でデータを分析するという知識と技術を身につけるといったことになってしまうわけです。これまでの調査実習ってというのは自分でデータをとることを前提としてきたわけです。ところが、今後既存のデータというものが公開されて利用可能になっていきますと、必ずしも自分で調査をやらなくても、必要なデータを取り寄せて分析することができるようになってきます。そうすると社会調査実習のあり方もかなり変わってくるのではないかと思います。これは今日の結論でもあります。

現行の社会調査実習の大体のスケジュールですが、授業の内容というのを表1に示しました。社会調査法とか社会調査論といわれる授業では大体どんな場合でも、最初は①のように仮説の構成についての話をします。これは、仮説というものをどうやって作るのかということ講義して学生に実際に仮説を作ってもらうんですが、まずこの部分でつまづく学生が結構多いです。私がこれまで授業を持ったことがあるのは、北大と淑徳大、お茶の水女子大、都立大といったところですが、どこの大学でも共通して、仮説というものがどういうもので、仮説というものを実際どうやって立てればいいのかわからないというのをよく聞きます。

それから②調査の計画についての講義をし

表 1 社会調査実習のスケジュールと関連科目

内容:	
①	仮説の構成
②	調査の計画
③	質問紙の作成
④	サンプリング
⑤	実査
⑥	データの入力・エディティングとクリーニング
	・表計算ソフトの使用法
	・統計パッケージの使用法
(1)	コーディング・エディティング
(2)	表計算ソフトによる入力
(3)	データの読み取りプログラム作成
(4)	読み込みチェックによる入力ミスの発見
(5)	単純集計による論理入力エラーのチェックと訂正
(6)	クロス集計による単純論理エラーのチェックと訂正
⑦	統計解析
	・統計パッケージの使用法
	・統計的方法の使用法
	・結果の解釈
(1)	欠損値処理 if 文、do-end 文、array 文
(2)	単純集計とグラフ
(3)	クロス集計と $\chi^2$ 検定
	$\chi^2$ 検定、第 1 種の過誤と第 2 種の過誤(検定の考え方)
	カテゴリーの再統合・再割り当て (if 文など)
(4)	要約統計量の算出と正規性のチェック
	平均、標準偏差、分散、歪度、尖度
(5)	2 群の平均値の差の検定
	t 検定
(6)	3 群以上の平均値の差の検定
	分散分析、F 検定、多重比較、交互作用、平均値のグラフ
(7)	散布図と相関
	相関係数
(8)	回帰・重回帰分析
	重回帰分析
⑧	レポートの作成

ます。この辺は大体講義でいうと社会調査法、社会調査論、こういった科目が関連します。次に③質問紙を作成するという作業が必要となります。社会調査法とか社会調査論という授業で質問紙の作り方なども一応講義で習っているわけですが、それを実際に実習でやっていく。④にサンプリングというのがあります。サンプリングは一応授業の中では話をするわけですが、実際に社会調査実習の中でサンプリングをやったということは私の場合にはありません。それから⑤の実査。実査も、小島先生のお話にもありましたが、私の場合もいわゆる学外で実際に学生にデータをとらせたという経験はほとんどありません。基本的には大学の中でデータをとらせるという形をします。やはり授業のために学外でデータをとるといのはちょっと抵抗がありますの

で。従って、いくらランダムサンプリングが重要だと講義で説明しても、学内でデータをとるとい方法なので、有為抽出になります。この辺が授業でやることと現実とがなかなか一致しないわけなんですけど、まあ仕方ないかなと思っています。

実査をやったら、⑥データの入力ですね。入力と同時にエディティングとクリーニングをやるんですが、この辺がまたつまずくんです。データの入力は実際は表計算ソフト、大体 Excel か Lotus ですが、これを使います。情報処理法という授業があるところでは、この授業で得た知識が一応役に立ちます。表計算ソフトでデータを入力して、次に統計パッケージの使用法というのをやるんですが、これも使用する統計パッケージによってここが難しくなったりやさしくなったりします。次の⑦、統計解析というところの統計パッケージの使用法というのも同じです。先ほどの佐藤先生のお話にもありましたが、私も SAS と SPSS の両方を教えた経験があるんですが、現状ではやはり Windows 版の SPSS を使ったほうがかなり楽だと思います。SAS は Windows 版になっても、プログラムを書かなくてはならないんで、このあたりが非常に初心者には疲れるということですね。それから⑦の中で、統計的方法の使用法というのをやるんですが、ここでは当然統計学との関係というのがでてきます。そして結果の解釈、⑧レポートの作成という感じで続きます。

⑥の中身をもうちょっと詳しく見ると、要するにデータの入力から始まる場所ですが、まずコーディングとエディティングというのをやります。まず調査原票からコーディングするんですが、それと並行して調査票の中でいろいろおかしいところをチェックする。これがエディティングです。で、それから表計算ソフトによる入力を行います。入力が終わったらデータの読みとりプログラムというのを作成します。これは Windows 版の

SPSS を使う場合にはほとんど必要ないことですが、SAS なんかの場合にはかなりここでプログラムを作らなきゃならないわけで、このへんが辛いところですね。それから次に、読み込みチェックによる入力ミスの発見、これは具体的にレコード数とケース数がちゃんとあっているかどうかをチェックします。SAS の場合でいうとレコード数とオブザベーションの対応ってやつですね。実習やると必ずここで学生がひっかかります。ここは実は突破するのはなかなか難しく、学生に入力させると必ずデータの桁ずれとか、入力忘れというのがあって、これを検出するというのがなかなか大変なんですね。それから、単純集計による論理入力エラーのチェックと訂正をします。つぎに、クロス集計による単純論理エラーチェックをするといった具合に続きます。

要するに⑥の中で、クリーニングにかなり力を入れているんですね。私は今までこういう方針を取ってきたんですが、この作業っていうのは学生にとって非常に面白くないようです。そりゃそうだと思います。データの入力をして、その入力ミスを見つけて、正しいものに直すということですし、しかもこの作業っていうのは実は統計解析以上に難しいところがあるわけですからね。ここで嫌になって逃げ出したりする学生が結構います。私がお茶大で教えたときに、非常勤で行っていたんですが、学生が泣いてしまいまして、「先生の顔なんか見たくない」って言われました。何で非常勤でそこまで言われなきゃいけないかって言う感じもしたんですが……。

私の知っている方でアメリカの大学院で社会調査法を教えている方がいるんですが、その方に伺うと、アメリカではすでに公開されているデータを使うので、その方はこの部分は飛ばしてしまうと言っていました。要するに、最初はデータ解析の面白さを実感させることを重視する、ということのようです。た

しかに、この辺で不必要な苦勞を味あわせるというのはマイナスかと思います。ただし、日本ではデータがそんなに公開されていませんから、例えば卒論でデータを使おうとすると、自分で調査を企画して、自分で分析するということになってしまって、クリーニングとかエディティングっていうのがかなり大事であるわけで、その辺で私は心を鬼にしてやっているというわけですね。

⑦の統計解析のところですが、これは SAS の場合ですが、まず欠損値処理ですね。次に if 文、do-end 文、array 文とやります。それから単純集計とグラフ、クロス集計とカイ 2 乗検定。だいたいここで、第 1 種の過誤、第 2 種の過誤、検定の考え方なんかをやります。変数の作成、カテゴリーの再統合、再割り当てなどもここでやります。次に 1 変数の分布という話に移ります。具体的には要約統計量の算出と、正規性のチェックとありますが、この正規性のチェックというのは、かなりマニアックな内容ではあるんですけども、私はこれを必ずやるようにしています。要するに相関係数を中心とした、パラメトリックな統計法は正規母集団を仮定するので、分析に使用する変数の分布を検討しておくことが非常に大事なわけですね。私はこのへんをかなり重視しているんですが、これもアメリカの研究者の方なんかに聞くと、あまりその辺はこだわらずに教えてるということでした。要するにロバストネスとかの問題ですが、これもやっていて面白いわけはありませんから、学生から嫌われます。

それで続いて、2 群の平均値の差の検定。この辺までは大体統計学でやってくれているんですが、続いて 3 群以上の平均の差の検定、分散分析、多重比較、交互作用、平均値のグラフ。私は個人的には、分散分析っていうのはかなり大事な方法だと考えているので、ここは結構丁寧にやるんですが、大体統計学の授業でですね、分散分析をやってくれてない

んですね。そのくせ、チェビチェフの不等式とか不必要な難しいのやったりして。そうするとどうしても、社会調査実習の中で分散分析の話をしなればいけないことになる。結構これに時間をとられてしまうというのも問題です。で、散布図と相関。それから回帰・重回帰分析。まあ重回帰までいけば上出来という感じでした、やらないことも多いですね。大体こんなスケジュールです。

### 3. 実習方法のパターン

先に述べたスケジュールで実際うまくいくのかというと、まあ正直言ってあんまりうまくいっていないですね。今までどういうやり方をしてきたかという、表2をご覧ください。まずは大きくわけて、パターンは、私の場合は3つほどあるような気が致します。まず、既存の素データを利用して⑦、つまり統計解析ですね、これを中心とした実習を行ってから、自分たちで実際にデータを収集して分析

するというやり方です。佐藤先生がやられている方法とかなり近いと思います。カリキュラムとしてはかなり万全だとは思いますが、いかんせん時間がかかりすぎるわけです。それから、学生の疲労が非常に大きいですね。

二番目のやり方は既存の素データを利用して、要するに統計解析を中心とした実習だけを行う。したがって、自分たちはもうデータをとらないというやり方です。これはまあ、時間がかからない。それと一応手順の習得というのは可能なわけですが、解析以外の実習がなくなります。そうすると質問紙を作るとか、クリーニングとかエディティングとかの作業がなくなるということになります。仮説を作ってその仮説から質問項目を作るという作業もなくなるわけですね。その辺が評価のわかれるところです。

三番目の方法は既存の素データを使用せず、自分たちで収集した素データを利用して、これに対して統計解析を行う。これも一通り

表2 素データを用いた社会調査実習の方法

方法	長所	短所
1. 既存の素データを使用して⑦を中心の実習をおこなってから自分たちのデータを収集・分析 (1)公開されているデータ利用 (2)研究用データ利用	カリキュラムとしては万全	時間がかかりすぎ、学生の疲労大
2. 既存の素データを使用して⑦を中心の実習のみ (1)マイクロデータ (2)マクロデータ	時間かからない、一応の手順の習得可能	解析以外の実習なし、学生の関心いまいち  →解析法が限定される
3. 既存の素データを使用せず、自分たちで収集した素データを使用して⑦中心の実習 (1)グループでデータ収集 (2)全体でひとつのデータ収集	一通りの実習が可能  データの規模が大きい	スケジュールが難しい、統計パッケージの使用に難 雪崩的離脱 データの規模が小さい フリーライダー、 バーンアウト

の実習が可能なんですけど、非常にスケジュールをたてるのが難しいです。この理由は後で説明します。この方法のもう一つの問題は、この方法を使うと統計パッケージの使用法とこの方法を理解させるのが非常に難しい。これは要するに、全員が同じデータを扱うわけではないので、結果の読み方とかを緻密に教えることができないんですね。

こういう3つのパターンが社会調査実習についてあるんですが、それぞれのパターンについてもう少し詳しく見ていきたいと思いません。

### 3.1 既存の素データの分析からオリジナルデータの収集・分析まで

まず、既存の素データを利用してその分析をしてみても、それから自分たちのデータを分析するという方法ですが、これまでの私の経験ですと素データとしてGSS (General Social Survey) データを使っています。これは私が北大にいた時のことで、その時は、新國先生が情報処理教育センターに勤めていらっしやいまして、非常にお世話になりました。あとフロアにいらっしやっている大津先生にもお世話になりました。あと本日参加されている北大の西浦君は私の授業を聞いていたという、考えてみれば非常にこわい人たちの中でこのようなお話をしなければならぬのですが、GSS データを使った時は、コード表を各自に配布してデータの共有という形でデータを利用しました。GSS データというのはアメリカのデータですが、このデータのいいところはサンプル数が非常に多いところにあると思います。まあ多いといってもGSSというのはそんなに多くもないんですが、一応三重クロス集計ができるくらいはあるということですね。で、変数が比較的多いので、多様な仮説が立てられる。この辺が非常に便利だと思います。しかし学生の反応はいまいちでして、要するに外国のデータであるために、自分自身の調査票を作

るときに調査票の構成なんか参考にならない。それと身近な感じがせずリアリティがない。まあ、こういう言い方されると身も蓋もないんですが、それから、入力とかクリーニングとかの実習というものが、公開されているデータそれ自体を使ってしまうと、そこではできないということになりますね。

続いて既存の素データを利用したもう一つのパターンとして、今度は授業担当者の研究用のデータを利用するというのがあります。これは要するに私自身が行った調査のデータを使ったということですね。これは北大でもお茶大でもやったことがあるんですが、この場合は調査票とコード表、データについては北大の場合は共有で、お茶大の場合にはフロッピーディスクでそれぞれ学生に配布しました。フロッピーはデータを Lotus のワークシートファイルの形式で配布しました。実習の作業というのは、Lotus の場合には Lotus のファイルで、K3 フォーマットという、CSV ファイルと同じですけど、ファイルに変換してからデータ解析するというのをします。お茶大は、学生さんは非常に優秀なんですけど、大学のシステムがどうしようもなく、たったこれだけの作業だけでもすごく時間を取られました。ワークステーションも何度も壊してやろうか思ったくらいひどくて、本当に学生が可哀想だと思いました。で、このやり方ですと、私自身の行った国内のデータを使ったわけですから、一応質問紙の構造とかデータの特徴っていうのは、担当者である私が熟知しているわけですね。で、国内のデータであるため、そこそこ学生もリアリティや関心が持てるという側面があります。自分のデータをこういう風に言うのもなんですけど、質問紙の作成とか仮説の設定の仕方なんかは学生にとって参考になるというメリットもあると思います。それから、一番大きいのは、そのデータを使用して発表した論文ですとか学会報告資料なんかを使用できるということ

すね、これが最大のメリットだと思いました。というのは、いわゆる実習というやり方で授業をやったときに、一番学生に言われるのは、実際にどういう風に結果を表現したりまとめたりすればよいのかわからない、というあたりのことなんです。こうしたことってというのは、授業なんかで個々の方法を習っても実際にはなかなか身につけることはできなくて、むしろ実際のデータを使って書かれている論文とか学会報告のレジメなんかを見せた方が、はるかに理解できるような気がします。ただ私が行った調査データにもいくつか欠点がありまして、その一つはサンプル数が少ないということなんです。GSSなんかと比べると非常に少ない。従って、使用できる解析が限定されてしまう、というデメリットがあります。

### 3.2 既存の素データの統計解析

次に、二番目の既存の素データを利用するだけの方法ですが、マイクロデータの実習というのがあります。これは要するに、実習をするためにデータを収集するわけです。そのデータを使って実習を行うわけですね。この場合は学生各自に調査原票、コード表とフロッピーディスクを渡します。学生は調査原票から入力して、ファイルの結合、コピー、ファイル変換、データ解析という順で作業を進めるわけですね。まあ、かなり泥くさいことをやるということになります。ですが、大概こういうデータは学内の学生から取ったデータなので、まあ学生自身が興味を持ってくれるということはあるようです。それからもう一つのメリットとしては、一応調査原票の入力から始まりますので、入力とかクリーニングとか、データの読み込みチェックというものができるといことですね。ただし、やはり変数は限定されます。まあ対象が全部学生ですので、変数は限定せざるをえません。しかも調査原票からクリーニングさせるというのは、いい経験だと思ってやるんですけど、

ここで脱落する学生が多いということはありませんね。それから社会的に有意義な仮説を立てにくい。学内の学生を対象にする調査といたしますと、質問も限定されてくるわけですね。よくあるのは小遣い額とか、恋人の有無とか、それはそれで関心を持ってくれる分はいいんですが、あまり面白い仮説にならないことが多くて、ちょっとマンネリな感じもあってよくないなあという気はします。で、サンプルが少ないというのも非常に大きい問題で、第2種の過誤も非常に大きくなってきますから、なかなかよい統計解析ができなくなるということになります。

それから、既存の素データとしてマクロデータ、具体的には都道府県別データなんかを使ったことがあります。これは都道府県別データのコピーというのを渡して、実際に入力、解析をやってもらうんですが、これは結論からいうと、大失敗しました。これは、マクロデータの解析法が非常に限定されてしまうということが大きいです。マクロデータの場合はほとんど量的な変数になりますし、都道府県別データというのはサンプルも非常に少ないんで、扱うのが非常に難しい。質問紙法との接点がありませんから、社会調査法なんかの講義と接点がなくなってしまうという問題もあります。まあ、メリットとしては、時間的に教えるのは楽です。

### 3.3 オリジナルデータの統計解析

最後に三番目のやり方ですが、これは要するに、学生自身がとにかくデータを集めて、それを分析するという方法で、とくに共通のデータは使わないというやり方です。3の1で、グループごとにデータを収集する、というのがあります。これはいろんな大学でやってみたんですが、質問紙も学生が作成するということになりますので、学生の興味関心が一応得られます。この方法をとると学生はそこそこやるんですが、デメリットも非常に大きいような気がします。まず、質問紙がどう

しようもないと、実習そのものが成立しない。これは後でご説明しますが、淑徳大なんかでは、質問紙を作る授業とそれを入力して解析する授業が別なんです。そうすると、教員が質問紙のチェックをちゃんとしていないのか、「それじゃあどうやったって駄目だよ」という質問紙がやたらに多くなるんですね。分散分析をやろうとしても分散分析にかけられる質問項目が1つもない。そうすると当然授業は成り立たなくなります。仕方がないからそういう変数だと仮定してやってみなさいとか、かなり無理なことをしなくてはならなくなります。

それからもう一つの問題は、グループでやるとフリーライダーが必ず出現するということです。同時にバーンアウト、つまりバーンアウトしてしまう人のことですが、これもまた出現します。で、表に雪崩的離脱って書いてあるんですが、これはグループで離脱してしまう可能性があるということです。私、淑徳に在職中によく経験がありましたけれども、要するにグループのリーダーがデータを持って逃亡したり、学校に来なくなってしまうとかいったことがあります。グループのメンバーが泣きそうな顔をして、データを持っている奴が学校に出てこない、学校やめるかもしれない、とかいって泣きついてきたなんていうことがありました。こんな感じでどっと離脱する人間が一度にでてしまう。それから、どうしてもグループの中で、一生懸命やる子と、適当にいいところだけもらってというフリーライダーがでてきちゃって、公平性がないという声も聞こえてきます。それと学生が自分たちでグループを作ってデータの収集をやると、非常にサンプル数が少なくなります。それから例題用のデータを使わないですから、アウトプットの読み方も一律には説明できないんですね。つまり全員が共通のデータを使っているときには、例えばクロス集計っていうのはこうやってこう見れば

いいんだよっていう形で説明できるわけですが、個々のグループによってデータが違う場合には、一律にそのやり方を説明できない。そうすると、授業をする方にとっては非常に不都合だということになります。それからですね、こういう方法をとると、学生同士の時間調整が非常に難しい。この方法をとったときに非常にでてくる苦情の一つです。東大なんかだとこういうのはないみたいですが、都立大なんかはこれはすごくあるんですね。都立大っていうのは八王子にありまして結構都心から離れていて、通学時間がかかる学生が多いんですね。そうすると、なかなか授業時間以外に学生同士が集まって何かするというのが、難しいという構造があります。それから、グループで作業をするためにグループ間の進度の差異がまちまちになってしまうという問題もあります。例えばあるグループはデータの収集が終わって、クリーニングも終わっているんですが、あるグループはまだ質問紙ができていないということによくあります。そうなってくると授業が全然進まないんです。

もう一つの問題は、調査公害ということですね。これは、いくつものグループが学内で調査やると、一人の人間が10票も20票も質問紙に答えるということになります。都立大学ですと人文学部の場合に、学生は一学年に120人くらいしかいないので、さすがに答える学生も可哀想だなという感じがします。それから3の2、これは、授業全体で一つのデータを収集するという事なんですが、この方法は質問紙の構成のチェックなんかもこちらでかなりできますし、アウトプットの読み方なんか、全員で一つのデータをいじるので、一律に説明が可能です。ある程度のサンプル数の確保も可能です。ところがデメリットとしては、やはり全員の関心を質問紙に反映させることが難しいということがあげられます。要するにみんないろいろなことを言って

くるわけで、それを一つの質問紙に反映するってというのは非常に難しいということですね。全体のとりまとめが非常に大変になってきます。ここでもフリーライダーがでてきます。学生の中で結果的にリーダー格になる子がいるわけですが、こういう方法をとるとそういう子が背負い込まざるを得ない部分が大きくて、結構辛い思いをしたなどということをよく言われます。

#### 4. 問題点とその解決法

というわけで、いろんなやり方を試みましたが、これがいいという方法はなかなかないということですね。ただ言えることは、マクロデータというのは使い方が非常に難しいという感じがします。それから、社会調査実習ってというのは他の授業との関連が非常に問題になってくる。参考までに淑徳大学社会学部のカリキュラムを表3で紹介したいと思います。このカリキュラムは、私の在職中にできたもので、かなりの部分を私が作ったという自負があるんですけども、淑徳大学では現在こういう風に行われております。まず基礎情報処理という授業を1年の時に必修でやります。授業は全部半期です。基礎情報処理は、パソコンの基本操作とアプリケーションの使用法、具体的に言うとワープロは一太郎、表計算はLotusの使い方なんかを実習でやります。社会統計学の授業というのも必修でやはり1年から履修することができます。社

会学方法論というのは選択なんですけど、これは仮説の構築法なんかをやります。社会調査I、II、IIIというのも必修です。社会調査Iは社会調査法の講義です。社会調査IIは事例調査をやって、そこから仮説を作って、調査票を作って、データの収集までを実習形式でやります。調査票ってというのは、統計的な調査のための調査票作りなんですけど、これが終わると社会調査IIIに続きます。社会調査IIIってというのは、社会調査IIで集めたデータを入力して、クリーニングして、SASを使って分析してレポートを出すというものです。で、ここで使うSASはそんなにたいしたことではありませんので、SASの上級者編として選択の授業で応用情報処理IIという授業が置かれています。

こうしてみると淑徳大学の社会学部のカリキュラムってというのは一見すると完璧なんですけど、やっぱりなかなか上手くいきません。実際の学生の声とそれに対する対応策を表4に示しましたが、調査票を作る前に解析を教えて欲しいという苦情が非常に多いですね。そういう意味では、やはり学生が自分でデータを集める前に、例題用のデータをいじって具体的にどういう解析を使うのかという勉強してから、例えば質問紙を作るなりするというやり方をした方がいいのではないかということになります。それから、グループでの時間調整が難しく、なかなか思い通りにできないという問題があるわけですが、これも考

表3 淑徳大学社会学部のカリキュラム (すべて半期1コマ)

講義名	履修年次	内容
基礎情報処理	必修、1年～	パソコンの基本操作と講義、一太郎、Lotus
社会統計学I	必修、1年～	記述統計学
社会統計学II	必修、1年～	推測統計学
社会学方法論	選択、1年～	科学方法論、仮説構築法
社会調査I	必修、1年～	社会調査法の講義
社会調査II	必修、2年～	事例調査実習→仮説構成→調査票作成→実査
社会調査III	必修、2年～	データの入力(Excel)→クリーニング→統計解析(SAS)→レポート
応用情報処理II	選択、3年～	統計解析(上級編)の講義・実習(SAS)



表4 社会調査実習に関する学生からの要望とそれに対する対応策

要望	対応策
調査票を作る前に解析法を教してほしい	例題データは使用した方がよい
グループでの時間調整が難しく予定通りにできない	個人作業化へ
仮説の立て方、検証方法が分からない	データを用いた実例、マニュアルの整備
自分たちが関心のもてるデータを使してほしい	データの選択可能性
フリーライダーが多くて不公平だ	極力、グループ作業よりも個人作業に
質問紙作り、実査で疲れすぎる	実習において質問紙の作成・実査をやめる？
授業時は毎回の課題をこなすのがやっとなで、結果の読み方や解釈がわからない	マニュアルの整備、データを用いた実例

え方はもちろんいろいろあるわけですが、極力、学生が個人個人でやれるような形で、つまりグループ作業という形ではなくて、個人作業でやれるような実習がいいのではないかと思います。

それから仮説の立て方、検証法がわからないというのがあります。これはですね、決して学生が理解不足だとか理解度がなにかいうことではないと思うんです。これはやはり実際にデータを用いて書かれている論文とか、学会報告資料とかを見せるべきではないかと思います。それとマニュアルの整備も大事なことであると思います。あと、自分たちが関心の持てるデータを使って欲しいという要望も学生からかなりあります。そういう意味では、データの選択可能性というのはあった方がいい。それからフリーライダーが結構いるという問題ですが、これも先ほどと同じで、グループ作業よりも個人作業にしていた方がいいかと思います。それから質問紙作りで疲れすぎてしまうという問題もあります。これは、冒頭に述べたように、これまでの社会調査実習というのは、自分で質問紙を作って自分でデータを収集するというのを念頭に置いてきたからなんです。アメリカなんかでは必ずしもこういう形ではやられていないわけですね。つまり今までデータというものが公開されていなかったの、自分でやるしかなかったし、そのために必要な方法を教えなくてはならなかったわけですが、今

後公開されていくようになれば、場合によってはこの部分はカットするという考え方もでてくるかと思います。それから、授業では毎回の課題をこなすのにやっとなで、結果の読み方や解釈がわからないという声もあります。これはやはりマニュアルの整備、それから、データを用いた実演が必要ですね。特に SAS とか、SPSS の使い方自体よりもアウトプットの読み方、解釈の仕方ですね。こういうことに一番要望が高いです。

というわけで、いろいろ問題点を指摘してきたわけですが、じゃあそういう問題はどうかやって解決していったらいいのかということになります。あんまり解決法にはならないんですが、まず、例題用のデータを使用すべきであると思います。つまり学生が仮にデータを取るにしても、データを取る前に例題用のデータを社会調査実習で使用した方がよいのではないかと、ということです。これは統計パッケージの使用法とか、アウトプットの読み方なんかを学習するためには、どんな場合であれ、全員共通の例題データを使用した方がやりやすい、というのが私なりの結論です。それからこれを使用しないとなかなか授業のスケジュールが組みにくいというのももう一つの理由ですね。

じゃあ例題用のデータを使うときにどんなデータを使うかということですが、まず一つの条件はサンプル数が多いということです。特に三重クロス以上の集計ができるサンプル

が欲しいなという風に思います。サンプル数が少なくなると第2種の過誤も非常に大きくなりますから、なかなか仮説が検証されにくくなりますので、サンプル数が多い方がいいかと。それからまあ、できればやはり日本語のデータの方がいいであろうと思います。さらに担当者が質問紙の構造を熟知しているということも、非常に大事ではないかと思えます。それから、学生が関心を持てる、卒論なんか参考に becoming というようなデータがよいと思えますね。そのデータを用いた論文とか報告があること、これも結構大きいんじゃないかと思えます。そういったものがあると、いわゆるプロの社会学者がそのデータを使ってどのように論文を書いてアウトプットをまとめて、どんな風にそれを文章で表現するのかというあたりがわかるんじゃないかと思えます。社会学に関して言うと、この辺を書いたテキストがあまり出回っていないんじゃないかという気がします。心理学なんかでは、『実践心理データ解析』などという本がこんな感じの本でして、こういうのがいいのではと思うんですが、社会学ではあまりこうした本は見かけませんね。

この意味で、全員共通で使うデータというのは、担当者が手がけた国内のデータが利用可能であれば、それがいいんじゃないかという気がします。それからもう一つの問題は、社会調査実習のどこから作業を経験させるべきか。私の場合には、データの入力から始まって、エディティングとか、クリーニングとかをやらせてきたわけですが、これは本当にやっている面白くないですし、われわれがデータ解析する時もこの部分っていうのは本当につまらないですね。プロの社会学者でも、この辺はかなりいい加減にやっている人も多いです。これをどう考えるか。クリーニングのなされたデータが公開されて、それを利用するのが可能という状況になってくれば、場合によってはこの部分は省略してしまうとい

うやり方も考えられるんじゃないかと思えます。

じゃあ学生が例題用データを使ったあとで、自分で解析するデータはどういうものか。先ほど述べましたように、グループ作業っていうのはなるべく避けた方がいい。それから、これまでは調査票の作成とか実査というのを社会調査実習の中で行うことを前提としてきたんですが、この考え方自体を再考する余地があると思えます。つまり、自分自身で調査をするという考え方自体を極力避けた方がいいんじゃないかと思えます。既存のデータを使い尽くす、あるいは自分でやるにしても、新しい質問紙は作らない。そういう形で極力、これまでになされた調査っていうのを使い、これまでに質問されている形を踏襲する、というやり方をした方が実はいいのかもしれないということですね。で、もし今後データの公開というのが可能になれば、むしろ WWW を用いてどうやってデータを探すのか、あるいはどうやってダウンロードするのか、そういうことを教えた方がはっきり言って現実的かもしれません。そうなってくると、従来は自分でデータを収集して分析するための技術を習得するということが社会調査実習の目的だったわけですが、むしろこれからは、自分で欲しいデータを探して、獲得して分析する技術の習得という形に、社会調査実習のやり方が変わってくるかもしれない。そうすると実は今まではデータの入力とか、クリーニングとか、エディティングとか、かなり泥臭い部分をやらざるを得なかったんですが、場合によってはそれをカットしてしまってもよいのではないかと思えます。まあ反論もあるかと思えますが、そういう立場が出てくるんじゃないかと考えています。

そうすると、学生が分析するデータですが、例えば研究用とは別に、実習教育用にデータが公開され、アーカイブ化されていると便利なのではないかと思えます。要するに、現在

公開されているデータというのは、大体が研究者向けで、大学院生以上が対象という形になっていますよね。一部には学部の学生用に公開されているのもあるようですが、そういったものは少ない。一方で、実習のために収集しているデータというのが結構あるんですね。大体そういうのは有為抽出で、研究用には問題があるわけですが、ただ、実習教育用には再利用されてもいいんじゃないかと思えます。研究用とは別にこういうデータをアーカイブ化するということも考えられるということです。

それから、小島先生の、なかなか人生を感じさせるような報告がありましたけれども、マニュアルを充実させるのは社会調査関係の講義では、非常に大事かと思えます。実際に、私家版のマニュアルというのは、非常にあちこちで作られているわけですね。ただし、そうしたマニュアルはその授業を取っている人と、その担当者にしか行き渡っていないわけです。こういうのは交換とか、参照できると非常に便利なわけです。実際に小島先生もおっしゃっていましたが、SASとかSPSSのマニュアルよりも、その大学でSASとかSPSSを動かすためにどういう手順でコンピュータを動かせばいいのか、そこところが結構大事で、逆にそういうマニュアルってというのが結構必要とされるんじゃないかと思えます。そういう意味で、この私家版のマニュアルというのは、WWW上で公開するなりアーカイブ化されるといいのではないかと、思うんです。実際、私は都立大学でそういう呼びかけをしているんです。センターの職員がSASとかSPSSのマニュアルを作るのは非常に大変なので、むしろすでに作られているマニュアルを利用した方がいいんじゃないかと。これは私が淑徳大学をやめたあとに作ったオレンジ色のSASのマニュアルですが、こういうものも極力交換できるといいんじゃないかと思っています。SPSSにつ

いても、こういうマニュアルを都立大学で作ってまして、普及していったらいいんじゃないかなと思っています。

## 5. 今後の調査実習教育のあり方

まあこのように、これまでは自分自身で独力で調査してそのデータを分析するという形で実習をしてきたわけですが、これからは公開されているデータを利用して分析していくというように変化していくという方向が予想できます。ただしそうすることの問題もあると思います。一つは、質問紙の作成とか実査とかエディティングとかクリーニングとか、こういう作業が実習に組み込めない。この部分が一番社会調査実習で大事だという風に考える人もいるわけですね。実際に、質問紙をどうやって組み立てるか、仮説に見合った質問項目をどうやって組み込むかということが、社会調査実習をやる上での最大の目的であって、むしろ統計解析なんかはあまりやらなくてもいいという立場もあるわけです。ここがやはり難しいところだと思います。ただし、仮に自分で質問紙を作ると、仮説の検証や解析にかえてマイナスになるということも、現状の問題としてあることも事実なんです。要するに、個人個人で質問紙を作っても、それはそれでいい経験になるんでしょうけれども、結果として講義で習ったことの半分も生かせないというようなこともあるわけです。そうすると、つまるところ社会調査実習の目的を、自分で質問紙を作る、データを入力して分析する、という泥臭い部分の経験を大事にするのか、それとも、とにかくデータを統計解析することに重点を置くのか、それでかなりあり方が変わってくると思います。これが大体私の社会調査実習についての経験と考えです。どうも有り難うございました。

## 稲葉講演に対する質疑

司会 (盛山) : どうも有り難うございました。詳しいお話で、いろいろ参考になる方も多いと思います。時間が押してますが、1人か2人ご質問いただいて、そのあと大谷先生のお話を聞きたいと思いますので、まず稲葉先生の今のお話に対するご質問どうぞ。

小島 : 茨城大学の小島ですけれども、社会調査全般ということでお話し頂いたんですけれども、先生のご経験、あるいは学生がどういうことを知っているのか聞きたいんですが、社会調査全体がまあうまくやれるというのは、一体どの学年、どの学部、あるいは大学で何年くらいかかれば可能だと先生は考えておられるのか、あるいはその考えを学部の学生にどのように伝えているのか、ちょっと教えてもらいたいのですが。

稲葉 : それは難しい質問ですね。要するに、いわゆるちゃんとした社会調査が、いつ頃になったらできるか、ということですか。やはり学生にはなかなかできないよ、としか言わないですね。まあここで社会調査実習受けたからといって、必ずしも完璧な社会調査実習ができるとは限らない、というのは言いますね。いつも言うことは、基本的にここで失敗したら、卒論で生かせばいいという風に、ネガティブな言い方しかしませんね。それからいい社会調査をやるには、社会学の理論とか先行研究をよく読んでいないと、なかなかできないですね。2年生くらいでこのテクニカルなところを習っても、やはり本質的な社会学の理論なり、研究の流れなりをつかんでいないと研究はできないと思います。そういう

意味では、学部の卒論で必ずいいものが出せるとも言い切れないし、そういう言い方もしないですね。ちょっと変な答ですが、まあ要するにいつまでに完璧なものができる、というようなことはあまり言っていない、というのが正直なところです。

小島 : それじゃお聞きしますが、私が社会調査っていったらデータ解析まで含めると、その調査票を作る前に、解析法を教えて欲しいなどという学生なんかもあって、そういうのを考えると、私の経験でも大学院の博士課程を終わるくらいになってやっとわかるのかなって感じがするし、それを勉強するのにとんでもない時間を取られます。そうすると、学部の2年生、3年生でやるなんてことは、よっぽどやさしいことしかできないんじゃないかという風に思ってます。それから実際に社会調査なんかでも、先生の説明でも、インタビューにしてもノンサンプリングエラーなんか全然触れられてませんからね、そこまで含めるとね、こりゃあ、学部の2、3年生では無理だろうと私はいつも思っているんですが。

稲葉 : そうですね。もちろん学部の2、3年でやれることってというのは限界があるわけです。しかし、社会学という学問を選んだ学生は、データをとってそれを分析するということに社会学ならでは、という非常に大きな期待感を持っていることが多いです。この期待にやはり教員は応える義務があると思います。だから完璧ではないけれども、こういうことが一通りの手順なんだ、ということをお教

えることに意義があるんじゃないかと思います。それからいわゆるノンサンプリングエラーについては、講義の中では一応話はします。実際取るデータは有為抽出ですから、学生にはランダムに取ったと想定して分析しなさい、ということになってしまいますが、**佐藤(東大)**：非常に参考になるお話で勉強させて頂きましたが、4ページの最後のところで、社会調査の全体の流れの中で全部教えられればいいけれども、学生のニーズとか時間の関係で、一部カットするとありますが、それをどう考えるかなんですが、私も学部の授業では、実査とかクリーニングの作業を落として、既存データの分析に授業のウェイトをおいてもいいのではないかと考えています。ただ大学院生の場合は、やはり実査やコーディングなどを経験することによって、調査票でこういう作り方をしては駄目だって初めてわかるわけですね。大学院生については、少なくとも一回は実査やクリーニングをやらせての方がいいと思うんですが、学部生については、全体のカリキュラムとか時間を考えると、それらをカットするのはやむを得ないかなと思ってます。

**小内**：札幌学院大学の小内ですが、うちの大学でも来年から、240人くらいの学生の調査実習が必修になるんで参考までに教えて頂きたいのですが、3ページのカリキュラムをやっていく際の、スタッフと学生数というのがどういう風になっているのか説明して頂けますか。

**稲葉**：淑徳大学社会学部の社会学科では、必修なんですけど、学生一学年だいたい180人くらいですか。で、コンピューターの教室は、一応パソコン本体が40台あるんですね、そこで授業やるんですけども、一クラス30人くらいでやるんです。それとTAが一人つきます。教員が非常勤を入れて何人かでこの授業

を担当するという形ですね。ただ、これも必修になってまだ一期くらいしか過ぎておりません。はっきり言ってあまりうまくいってるとはいえないですね。やはり授業同士の連携というのは非常に難しく、社会調査Iでならった社会調査法が身につけていなくて、社会調査IIで作った質問紙がメチャメチャだったりすると、社会調査IIIっていうのは、もうやる前から成り立たなくなっちゃうっていうのが難しいところですね。教員同士が授業の教え方とか進み具合なんかをお互いに相談できればいいんですが、なかなかそういう機会がなかったというのもあって、それがちょっとネックになってるかなあとと思います。ただ、1クラスに30人規模で、TA1人つけばまあ教員一人でも解析を教えるのは何とかなると思います。で、SPSSならば多分もっと楽です。SASなんですね、淑徳は、あとついでに申し上げておくと、やっぱりSASとSPSSではもう桁違いに授業が違うなあという感じで、SASのおそらく半分くらいの時間でSPSSではできちゃうと思います。SPSSで単純集計つくるなんてほんと簡単ですが、SASは非常にそこだけで大変だなあという感じがします。

**小内**：この社会調査I、II、III、と分かれていますけど、それぞれが6クラスできたんですか？

**稲葉**：社会調査Iというのは講義なんで、これは1クラスです。II、IIIというのは、これはどちらも半期ですが、II、III、は6クラスくらいですね、それぞれ。

**司会(盛山)**：いろいろと検討しなきゃいけないことがあると思いますが、時間が迫っておりますので、これで、稲葉先生のお話を終わらせて頂きます。どうもありがとうございました。