

統計学とデータ分析に対する知識と意識

——社会科学を専攻する大学生の事例から——

坂 無 淳*

要旨 本論文では、統計学とデータ分析に関する2科目での調査結果から、社会科学を専攻する大学生の統計学とデータ分析に対する知識と意識がどのようなものを把握し、どのような教育実践が有効となるかを考察する。その結果、社会科学を専攻する大学生には統計学やデータ分析に対する苦手意識はあるものの、実際はすでにある程度の知識を持っていることがわかった。また、1変数に関する知識は多いが、2変数以上の分析や推測統計学・確率の知識が少ないなどの特徴があった。そのため教育実践としては、高校での教育や他科目との接続を意識しながら、具体的な変数のデータ分析を多く取り入れ、2変数以上による「説明」の面白さを伝えること、また推測統計学に進む段階でのつまづきをやわらげることが重要になるとわかった。本事例はあくまで一事例でしかないが、引き続き大学生が持つ統計学とデータ分析に対する知識と意識を把握することが有効な教育実践につながると考える。

キーワード 統計学 データ分析 社会科学 社会学 大学教育

1 はじめに

本論文の目的は以下の2つである。ひとつめに、社会科学を専攻する大学生の統計学とデータ分析に対する知識と意識がどのようなものを把握することである。ふたつめに、社会科学を専攻する学生に対してどのような統計学とデータ分析の教育実践が有効となるかを探ることである。そのために、本論文では筆者が授業

を担当する福岡県立大学人間社会学部の「データ分析の基礎」と「社会統計学I」という2つの科目での調査結果を分析する。

このような研究を行う必要性を、科目の「入口」と「出口」に関してまとめよう。まず、科目の「入口」に関してである。社会科学を専攻するいわゆる「文系」の大学生は、後述する先行研究でも確認されているように統計学やデータ分析に対する苦手意識を持っていると考えら

* 福岡県立大学人間社会学部・講師

れる。統計学やデータ分析という難しい数学の知識が必要で、かつパソコンで難しい操作を行うというとにかく「難しい」イメージがあるようである。学生が統計学のイメージを持つ手がかりといえば、まず高等学校の数学であろう。高等学校の数学 I では「データの分析」などの内容があり、そこでは平均値や分散、また散布図や相関係数などがある。大学生はこれら統計学の基本的な知識をすでにある程度持っており、過度な苦手意識を持つ必要はない。しかし、大学での統計学がどのようなものかわからない以上、高校の数学全体に対する大まかなイメージが想起されることが予想される。また、パソコン操作に関しても、後述の先行研究からはパソコンやエクセルの操作に対して苦手意識があるようである。本当にそのような苦手意識があるのか。あるとすればその苦手意識はどの程度のものか。このような学生が持っている知識と意識の実態に基づいて、科目の内容とレベルを検討することが有効な教育実践に必要となろう。

つぎに、科目の「出口」に関してである。大学の各科目が単独で成り立っているのではなく、カリキュラムの他科目との関係のなかで設定されている以上、他科目との接続を意識する必要がある。より具体的にいえば、この科目で最低限どのような知識を習得してもらい、次の段階の科目につなぐべきなのだろうか。また当然、その科目以前にどのような科目を受けて、どのような知識をすでに持っているのかの把握も重要になる。特に本論文の2科目は社会調査協会の「社会調査士」の科目（C科目とD科目）となっている。この科目の前にAとB科目、この科目の後にEとFとG科目がある。社会調査士という文脈だけでもそれらとの接続を意識す

る必要がある。ほかにも、統計学やデータ分析の知識は、卒業論文などで使われることがあろう。そのため、社会学、社会福祉学、教育学、心理学など受講生の専門に最低限必要な統計学やデータ分析の基礎が何かを考えつつ、教育を行う必要がある。

このような学生の統計学、あるいは情報機器操作に関する先行研究として、すでに福岡県立大学人間社会学部の事例としては以下があり、ある程度の把握を行うことができる。

福岡県立大学の大学生の情報機器利用実態と情報機器操作スキルの状況について、1年次前期必修科目である「情報処理の基礎と演習」担当者の柴田雅博（2018）が、履修生の人間社会学部1年生160名に対してアンケート調査の実施と分析を行なっている。本論文の関心から一部を参照しよう。まず、自宅で利用できるパソコンがある学生は、科目受講前でも約90%（160人中143人）、受講後は約97%（158人中153人）であり、ほとんどの学生が自宅でパソコンを利用できる。ただし、受講前のパソコンの利用用途としては、文章作成やウェブサイト閲覧が多く、表計算や発表資料作成での利用はほとんどされていない。受講後は文章作成の利用が増えるとともに発表資料作成での利用が増える。しかし、表計算の利用は158人中23人に留まっている。また、学生の情報操作機器操作への自信について、パソコンの基本的な操作スキル、ワードの操作スキル、エクセルの操作スキル、パワーポイントの操作スキル、インターネット利用のスキルを、「充分ある」「ある程度ある」「あまりない」と3段階でたずねた結果、受講前はインターネット利用以外は約半数が自信が「あまりない」。特にエクセルは約70%が「あまりない」と自信が低いことがわかる（受講後は

どの項目も「ある程度ある」と答える学生が増え、エクセルについても約80%が「ある程度ある」と答えている(柴田 2018:193-8)。

本論文が対象とする「データ分析の基礎」は、「情報処理の基礎と演習」と同じ1年次の前期配当科目である。まさに情報機器操作スキルを身につけつつある段階での履修であり、パソコンやエクセルへの自信は低い学生が多いことが予想される。

よりデータ分析に特化した研究として、同じく福岡県立大学人間社会学部の3年次前期の「データ処理とデータ解析Ⅰ」の担当者である石崎龍二・佐藤繁美(2018)が、受講前後にくわえて毎回の授業での調査から学生の到達度を把握・分析している(受講生は76名)。本論文の関心から一部を参照すると、まず、授業の実施前後で確実に記述統計・推測統計の知識やスキルが上昇している。ただし、授業の難易度を5段階でたずねると「難しかった」「やや難しかった」の合計が86.0%と高い。特に記述統計学から推定や検定など推測統計学に進む段階、また1変数の分析からクロス表や相関分析など2変数への分析に進む段階など、内容が大きく切り替わる段階で難しいと感じる学生が多いことがわかる。そのため、仮説検定などの推測統計学や変数間の関連性の分析の指導を丁寧に行う必要があることがわかっている(石崎・佐藤2018:206-19)。

上記の知見は本論文の対象学生にもおおむね共通すると予想される。記述統計学から推測統計学に進む段階、変数間の関連性の分析に進む段階での指導を丁寧にする事で、「データ処理とデータ解析Ⅰ・Ⅱ」や「社会調査実習」など他科目との接続を良好にできる可能性がある。難しいと感じられる部分を全て簡単に説明

するなどはできないだろうが、多くの学生がどこを難しいと感じやすいかを把握し、あらかじめそれを学生に説明しておくことは、必要以上の挫折感や苦手意識を味あわせることなく、スムーズな学習につながると考えられる。

このような関心から以下、本論文ではまず、研究の方法として科目の詳細と調査方法を説明する。次に受講生の知識と意識に関する調査結果をまとめ、その結果を踏まえてどのような教育実践が有効かを考察する。調査結果の分析は、本論文の目的から、明確な仮説にもとづく仮説検証的な分析というよりは、主に結果の記述と探索的な分析を行う。最後に、結論をまとめ、本論文の課題を述べる。

2 方法

2-1 調査協力者と科目の詳細

調査協力者は2018年度前期の「データ分析の基礎」と「社会統計学Ⅰ」の履修学生である。授業前のデータとして、2018年4月のそれぞれ第1回の授業において、後述する質問をパソコン上で回答してもらった。また、授業後のデータとして、2018年7月のそれぞれ第15回の授業において、同様の方法で調査を行なった。

アンケート冒頭には、以下を明示して口頭でも説明を行って任意の調査協力を依頼した。本調査は無記名で行うため個人が特定されることはないこと、回答内容が成績に影響することはないこと、また回答しないことでの不利益はないこと、回答の結果と分析は授業改善と学術目的以外には使用しないことである。

科目の詳細として、2つの科目はともに社会調査士¹⁾(一般社団法人 社会調査協会 2018)、

上級情報処理士（一般財団法人 全国大学実務教育協会 2018）の科目となっている。

「データ分析の基礎」²⁾は1年次前期配当の科目である。社会調査士C科目（基本的な資料とデータの分析に関する科目）に該当する。また、教員免許（中学校社会及び高等学校公民）の科目ともなっている。テキストは廣瀬毅士・寺島拓幸（2010）を使用し、情報処理室を使用して、統計学や社会調査に関する講義とともにエクセルなどパソコンでのデータ分析を行った。

社会調査士のカリキュラムでは科目ごとに取り上げるべき内容が決まっており、C科目（基本的な資料とデータの分析に関する科目）は公的統計や簡単な調査報告・フィールドワーク論文が読めるための基本的知識に関する科目である。具体的には単純集計、度数分布、代表値、散布度、クロス集計などの記述統計、グラフ、相関係数、因果関係と相関関係の区別、擬似相関の概念などを含む（一般社団法人 社会調査協会 2018）。

「社会統計学Ⅰ」³⁾は2年次の前期配当科目である。社会調査士D科目（社会調査に必要な統計学に関する科目）に該当する。「社会統計学Ⅰ」は2年次後期の社会調査士の同じくD科目「社会統計学Ⅱ」との連続を念頭においた科目である。また、日本心理学学会の認定心理士の科目ともなっている（公益社団法人 日本心理学会 2018）ため、心理学専攻の学生の履修もある。テキストは、岩井紀子・保田時男（2007）を使用し、情報処理室において、統計学や社会調査に関する講義とともに、実際の社会調査データの分析を行った。具体的には、東京大学社会科学研究所附属社会調査・データアーカイブ研究センター（2018）の提供するメタデータ閲覧・オンライン分析システムであるSSJDA

Nesstarに情報処理室のパソコンから各学生がアクセスし、オンラインでリモート集計を行なった。

社会調査士のD科目（社会調査に必要な統計学に関する科目）は統計的データをまとめたり分析したりするために必要な、推測統計学の基礎的な知識に関する科目である。確率論の基礎、基本統計量、検定・推定理論とその応用（平均や比率の差の検定、独立性の検定）、サンプリングの理論、属性相関係数（クロス表の統計量）、相関係数、偏相関係数、変数のコントロール、回帰分析の基礎などを含む（一般社団法人 社会調査協会 2018）。

2-2 質問項目

質問項目は授業前には以下をたずねた。まず基礎項目として、学年、学科を聞いた。それから自宅でパソコンやエクセルが使える環境があるか以下をたずねた。パソコンを持っているか、そのパソコンはノートパソコンかデスクトップか、OSの種類、そのパソコンでワード、エクセル、パワーポイント、インターネットを使うことができるか。

次に当該科目以外の関連授業の履修状況（統計学など11科目）と資格・免許を取得する予定をたずねた（社会調査士、教員免許 中学（社会）、教員免許 高等学校（公民）、上級情報処理士、マイクロソフト・オフィス・スペシャリストMOS）。

数学・統計学やパソコンに関する意識として、以下をたずねた。高校の数学は好きな科目だったか、高校の数学は得意な科目だったか、高校の数学Ⅰのうち「データの分析」は好きな分野だったか、高校の数学Ⅰのうち「データの分析」は得意な分野だったか。また、パソコン

操作全般には自信があるか、パソコン操作のうちエクセルの操作には自信があるかを聞いた。

また、知識として、以下の用語や手法について、どの程度理解しているといえるかをたずねた（平均値から記述統計学と推測統計学の違いなど12項目）。質問からもわかるように、あくまで学生の主観で回答してもらっている。正確な知識を把握するためには授業前後で試験を行ったり、今回の結果を最終的な成績と対応させるなどすべきだろう。しかし、調査を無記名で行なうこと、また大まかな学生の知識の把握という今回の調査の趣旨のため、それらは行わなかった。そのため、知識といっても上記の用語や手法に対する自信といった意味合いが強いことに注意してほしい。

最後に、自由記述として、授業前は本授業を受けようと思った理由やきっかけ、本授業を受けるにあたって、心配していることや不安点、楽しみにしていることや学びたいことを文章で記述してもらった。

授業後の調査では授業前の質問項目と変化しないと考えられる部分などを除いた。具体的にたずねたのは、学年、学科、パソコン操作全般には自信があるか、パソコン操作のうちエクセルの操作には自信があるか、資格・免許を取得する予定、12項目の知識をたずねた。最後に、受講終了にあたっての考えを自由に記述してもらった。

なお、これらの本調査の質問項目の作成については、先述の柴田(2018)と石崎・佐藤(2018)の質問項目と分析結果を参考とした。

3 結果：統計学に対する知識と意識

それでは、回答者のプロフィールを説明した後に、主な結果のみ4点をまとめよう。

3-1 回答者のプロフィール

表1にあるように「データ分析の基礎」では授業前に54名、授業後に51名が回答した。「社会統計学Ⅰ」では、授業前に68名、授業後に63名が回答した。「データ分析の基礎」は1年生、「社会統計学Ⅰ」は2年生が多い。学科は人間形成学科の学生で心理学や教育学を専門とする学生、また少数社会福祉学科の学生もいるが、どちらの科目も社会学専攻の公共社会学科の学生が多いことが特徴である。

学生のパソコン環境について、大学でパソコンが利用できることは当然として、自宅でもパソコンやエクセルを利用できる学生が大半であることがわかった。自宅でパソコンを持っている学生は「データ分析の基礎」では90.7%（54人中49人）、「社会統計学Ⅰ」では100%（68人中68人）であった。ワード、エクセル、パワーポイント、インターネットについても利用できる学生が大半であり、エクセルについてのみ述べると、その自宅のパソコンでエクセルを使うことができる学生は「データ分析の基礎」では90.4%（52人中47人）、「社会統計学Ⅰ」では95.6%（68人中65人）であった。

他科目の履修状況については表2のように、1年生が大半の「データ分析の基礎」ではこの科目が初めて大学で統計学やデータ分析に触れる科目であることがわかる。2年生が多い「社会統計学Ⅰ」では、前学期や今学期に「データ分析の基礎」など他科目を履修している学生が多い。ただし、学生の履修状況には多様性があ

表1 回答者のプロフィール

データ分析の基礎 (授業前)					データ分析の基礎 (授業後)				
学年	人数	割合	人数	割合					
1年生	49	90.7%	47	87.0%					
2年生	2	3.7%	1	1.9%					
3年生	3	5.6%	3	5.6%					
4年生	0	0.0%	0	0.0%					
合計	54	100.0%	51	94.4%					
学科	人数	割合	人数	割合					
公共社会学科	50	92.6%	48	88.9%					
社会福祉学科	1	1.9%	0	0.0%					
人間形成学科	3	5.6%	3	5.6%					
合計	54	100.0%	51	94.4%					

社会統計学 I (授業前)					社会統計学 I (授業後)				
学年	人数	割合	人数	割合					
1年生	5	7.4%	0	0.0%					
2年生	61	89.7%	61	96.8%					
3年生	2	2.9%	1	1.6%					
4年生	0	0.0%	1	1.6%					
合計	68	100.0%	63	100.0%					
学科	人数	割合	人数	割合					
公共社会学科	47	69.1%	46	73.0%					
社会福祉学科	1	1.5%	0	0.0%					
人間形成学科	20	29.4%	17	27.0%					
合計	68	100.0%	63	100.0%					

表2 他科目の履修状況

データ分析の基礎				社会統計学 I			
科目	前学期まで履修	履修中	履修していない	科目	前学期まで履修	履修中	履修していない
統計学	3	2	49	統計学	43	1	24
情報処理応用演習	4	1	49	情報処理応用演習	46	0	22
社会調査法	1	3	50	社会調査法	45	2	21
社会調査の設計	0	3	51	社会調査の設計	0	43	25
社会統計学 I	3	3	48	データ分析の基礎	46	2	20
社会統計学 II	0	2	52	社会統計学 II	0	4	64
データ処理とデータ解析 I	0	4	50	データ処理とデータ解析 I	1	1	66
データ処理とデータ解析 II	0	1	53	データ処理とデータ解析 II	0	0	68
社会調査実習	0	3	51	社会調査実習	1	45	22
文化人類学B	0	1	53	文化人類学B	0	2	66
質的調査法	0	1	53	質的調査法	0	0	68

注：n=54

注：n=68

ることもわかった。

3-2 知識

- ① 1変数に関する知識は多いが、2変数以上や推測統計学・確率の知識が少ない

統計学やデータ分析に関する知識についての平均値を表3にまとめた。数値が高い方が理解度が高い(5=よく理解している、4=ある程

度理解している、3=どちらともいえない、2=あまり理解していない、1=まったく理解していない)⁴⁾。また、それぞれの知識12項目を合計した合計得点が表の一番下の「知識(合計)」である。

表からはどちらの授業でも授業前からすでに1変数に関する分析の理解度が高いことがわかる。具体的には平均値、中央値、分散、標準偏

表3 統計学の知識の平均値

	データ分析の基礎 (授業前)	データ分析の基礎 (授業後)	社会統計学 I (授業前)	社会統計学 I (授業後)
平均値	4.7	4.8	4.7	4.8
中央値	4.7	4.8	4.7	4.8
分散	3.7	4.0	3.8	4.1
標準偏差	3.4	3.8	3.7	4.0
四分位点	3.9	3.9	3.8	4.1
正規分布	2.2	3.4	2.9	3.8
標準化	1.8	3.0	2.8	3.2
グラフの使い分け	3.1	4.2	3.4	3.8
散布図	3.6	4.0	3.6	3.8
相関係数	3.3	3.9	3.5	3.9
クロス集計表	1.3	3.5	2.9	3.8
三重クロス集計表	1.2	2.9	2.1	3.6
単回帰分析	1.3	2.7	2.1	2.7
重回帰分析	1.2	2.6	2.0	2.7
t検定	1.3	2.0	2.5	2.9
分散分析	1.4	2.2	2.3	2.8
記述統計学と推測統計学の違い	1.4	2.2	2.3	2.9
知識 (合計)	43.3	57.4	52.9	61.5
n	54	51	68	63

差、四分位点（パーセンタイル）、グラフの使い分けで平均が3を超えている。また、2変数に関しては散布図、相関係数の理解度が高い。これらは高校の数学で取り上げられているためだと考えられる。

一方で2変数以上の理解度は低い傾向がある。具体的にはクロス集計表、三重クロス集計表、単回帰分析、重回帰分析、t検定、分散分析である。これらは大学に入って本格的に触れる学生が多いためだと考えられる（例外は先述の散布図、相関係数）。また、推測統計学・確率に関する部分とは、表では正規分布、標準化、記述統計学と推測統計学の違いである。また、t検定、分散分析のほか、クロス集計表、三重クロス集計表、単回帰分析、重回帰分析も推測

統計学の内容を含む。

② 知識の授業前後の変化

表3から授業前後の変化を見よう。「知識（合計）」をみると、「データ分析の基礎」の授業前後では43.3→57.4と理解度が上がっている。同様に「社会統計学 I」の授業前後でも52.9→61.5と理解度が上がっている。また、1年生が多い「データ分析の基礎」よりも2年生が多い「社会統計学 I」での理解度が高いこともわかる。個々の用語や手法の授業前後の変化については表を参照してほしいが、どの項目も理解度が上がっていること、先述のクロス集計表など2変数以上の分析や正規分布・標準化など推測統計学・確率に関する理解度も上がって

いることがわかる。

3-3 意識

③ 履修のきっかけは資格取得と将来のため

履修学生には社会調査士、上級情報処理士を取得することを考えている学生が多かった。社会調査士を取ろうと思っている学生は「データ分析の基礎」では77.8% (54人中42人)、「社会統計学 I」では63.2% (68人中43人)であった。上級情報処理士を取ろうと思っている学生は「データ分析の基礎」では64.8% (54人中35人)、「社会統計学 I」では58.8% (68人中40人)であった。自由記述での授業をとろうと思ったきっかけには多様な記述があるが、「社会調査士の資格を取りたいと思ったから」など資格に関するもののほか、科目の内容やおもしろさに関するもの(「統計学の授業を受けて、統計がおもしろいなと思ったからです。また、心理コースなので、統計は今後大切になってくると思ったからです」)、「将来の仕事のスキル習

得のため」など卒業後のスキルに関するものがあった。

④ 数学とパソコンに対しての苦手意識を持つ学生が多い

履修にあたっての不安点として、自由記述では数学や統計学への苦手意識(「数学が苦手であること。高校時代は文系でした」など)と、パソコンやエクセルへの苦手意識(「パソコンの操作があまり得意ではないのでついていけるか不安です」など)をあげる学生がいた。ただし、パソコンの苦手意識は高くても、向上させたいという意識が高いようである(「統計学について理解できたうえでデータを分析できるようになりたい。またパソコンをうまく活用出来るようになりたいです」など)。

これらの苦手意識を数値で確認したのが図1である(授業後や「データ分析の基礎」もほぼ同様のため人数の多い「社会統計学 I」の授業前のみ掲載)。箱ひげ図の箱と横線は第1四分

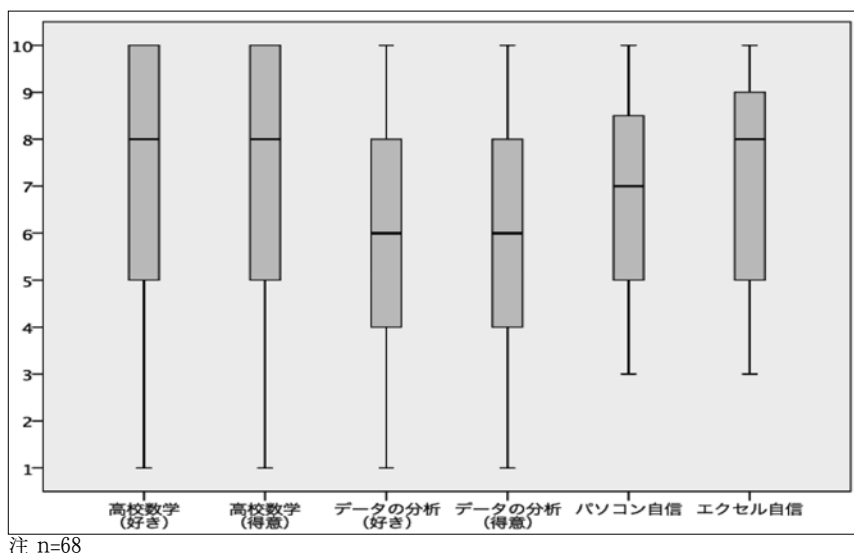


図1 数学・データ分析・パソコンへの意識 (社会統計学 I 授業前)

表 4 意識・知識間の相関（社会統計学 I 授業前）

	高校数学 (好き)	高校数学 (得意)	データの分析 (好き)	データの分析 (得意)	パソコン 自信	エクセル 自信	知識 (合計)
高校数学 (好き)	1						
高校数学 (得意)	.947	1					
データの分析 (好き)	.486	.437	1				
データの分析 (得意)	.487	.470	.882	1			
パソコン自信	.164	.131	.205	.282	1		
エクセル自信	.303	.216	.218	.299	.854	1	
知識 (合計)	-0.121	-0.059	-0.115	-0.235	-0.052	-0.185	1

n=68

位点、中央値、第3四分位点を示している。10は嫌いだっただ、苦手だった、あるいは自信がないなどネガティブな回答、1は好きだった、得意だった、自信があるなどポジティブな回答を示している。どの項目も数値が高いネガティブな回答が多く、中央値も8など高校の数学やパソコン操作には苦手意識を持つ学生が非常に多いことがわかった。ただし、高校数学の中でも「データの分析」に関しては、中央値が6であるなど苦手意識が比較的低い分野であることもわかった。

これら数学やパソコンへの苦手意識の間や、先述の知識と意識の間には相関はあるだろうか。ピアソンの相関係数をみたのが表4である。数学に関する変数間には正の相関があるのは当然である。しかし、それら数学に関する変数とパソコンやエクセルに関する変数の相関は低い。また、12項目の知識の合計である「知識(合計)」と意識の間の相関係数は小さい(知識は数値が高いと知識が多く、意識は数値が高いと苦手なため相関係数は負となるが絶対値は0に近い)。そのため、これら数学やパソコンへの苦手意識、また知識はそれぞれ別のものとして捉える方が適切だと考えられる。

4 考察：どのような教育が有効か

以上をふまえて、どのような教育実践が有効となるかを考察しよう。

① 具体的な変数についてのデータ分析を多く取り入れること

調査からは高校の数学は苦手であり好きではないが、数学のうち「データの分析」は比較的苦手意識が低い学生が多いことがわかった。そこで授業では難しい数学はほとんど使わず、あくまで「データの分析」の延長にある分野であることを強調して、学生の苦手意識を低めることができると考えられる。また、高校の「データの分析」の知識との関連性を再起させるような内容が有効だと考える。社会科学を専攻する学生に比較的「データの分析」が好まれる理由として筆者が推測するのが、数学の他の分野よりも「具体的」であるという点である。ある具体的な変数と別の変数の関係が事例となるような具体性の高さ(抽象度の低さ)に関心を持つ学生が社会科学系専攻には多いのかもしれない。そうすると、社会科学専攻の学生向けの科目では、受講生が関心を強く持つような具

体的な変数を含んだ実際の社会調査データなどを例として使うことが有効になると考える。時間が許せば実際の社会科学系の論文でどのような変数が統計的に分析されているか、その具体的な使用例を多く紹介することで、より学生の関心を高めることができると考えられる。

② 1変数の「記述」とどまらず2変数以上による「説明」の面白さを伝えること

結果からは学生は1変数の分析に関しては比較的、知識があることがわかった。一方で、2変数以上については知識が少ないことがわかった。高校までの「データの分析」では、主に1変数の記述的な分析が中心であるが、2変数の相関係数の分析も含まれている。社会科学系の実際の研究では1変数の「記述」で終わることは少なく、多くの場合はある変数を別の変数で「説明」する仮説を立て、それをデータで検証するという形をとる。さらに第3変数などをコントロールすることが重要になる。こう考えると、基礎としての1変数の分析は当然必要であるが、2変数以上の関係について分析する「説明」の重要性や第3変数のコントロールの重要性を強調する必要がある。学生が実際に先行研究を読む際には、その研究の仮説を把握し、結果を理解するためにこの点が不可欠になるためである。また、卒業論文などで実際に自身で調査や実験を設計し、分析を行う際にもこの点が不可欠になるためである。

③ 推測統計学に進む段階でのつまずきをやわらげること

同様に学生は記述統計学に関しては比較的知識がある一方で、推測統計学・確率については知識が少ないことがわかった。この点は先行研

究の石崎・佐藤(2018)でも確認されていることだが、記述統計学から推定や仮説検定など推測統計学に進む段階で指導を丁寧に行う必要がある。具体的な工夫としては、母集団の推測という目的を忘れないよう指導する点があると考えられる。推定や仮説検定では手続きや計算、またパソコンでの操作が多く、それらの細かい点に気が行きがちである。しかし、それらの細かい手続きのみに着目するのではなく、あくまで母集団の推測という目的を忘れないよう指導することが有効だと考える。また、仮に推測統計学に進む段階でつまずいたとしても、多くの人がその段階でつまずくということを事前に説明することも、苦手意識を高めないために重要だと考える。仮に一度でよくわからなくても、何度も話を聞き、実際に自分で計算・分析する機会を持ち、他科目でもそれらを繰り返すことで、推測統計学が理解できるということを伝えることが有効になると考える。

④ 社会調査データと実験データの違いを意識すること

本調査の受講生は大半が社会学専攻の学生であり、一部心理学や教育学などを学ぶ学生がいた。このような場合、どちらの学生にとっても、社会調査によって集められたデータと実験によって集められたデータを扱う際にはいくつかの違いがあると伝える必要があると考える。具体的には、社会調査データと実験データには、あくまで大まかな傾向でも主に2つの違いがあると考えられる。まず、社会調査データでは比較的多数のデータを取る一方で、実験データでは少数のデータを取る傾向がある。そのため、データの提示方法や推測統計学での扱いに違いが出る場合がある(具体的には推定や検定を行う際

に多数のデータでは正規分布を想定できるが少数のデータでは正規分布を想定できないなど)。次に、第3変数のコントロール（統制）について、社会調査データでは比較的多数のデータを取って、事後的に第3変数を統制するという方法を取ることが多い（具体的には三重クロス表や重回帰分析など）。一方で、実験データの場合、対象者をランダムに実験群と対照群に割り当てて第3変数を統制する実験計画を行うことが多い。そのため、よく使われる分析方法も異なる傾向がある（例えば社会調査データの場合、重回帰などの回帰分析がよく使われるが、実験データの場合はt検定や分散分析など平均値の比較がよく使われるなど）。統計学やデータ分析といってもこのように分野によって事情が異なるということを説明することが、上位の科目や卒業論文との接続に役立つと考えられる。

⑤ パソコンではなく考え方の科目であることの強調

授業前の調査結果からは、パソコン操作への不安や、一方でパソコンの利用方法を学びたいという積極的な意識が見られた。受講の理由としても資格に必要だからという理由とともに、将来に役立つからという理由が比較的多くあった。授業前の調査は第1回の授業で情報処理室に集まった直後に行なった。パソコンに囲まれている状況からエクセル、パワーポイント、ワードなどパソコンを多用する授業であると意識された可能性がある。また、授業後の調査ではパソコン操作やエクセル操作への自信が高まった、あるいはもう少しエクセルを使用する時間を長くしてほしいという意見があり、エクセルの操作への意欲は高い。ただし、この2

科目では、エクセルはデータ分析の道具の一例でしかない。単にエクセルの使用方法という技術的なことをメインに教えるのではなく、どのようにデータ分析を行えるのかという事例を示すのが目的である。この点は授業の中でも強調したが、さらに強調して伝える必要がある。また、パソコン操作や実際の分析については、他科目でも重点的に扱っていることを伝え、科目間の連携をとることも重要だと考える。

5 まとめ

以上、本論文では、筆者が授業を担当する福岡県立大学人間社会学部の「データ分析の基礎」と「社会統計学I」という2つの科目での調査結果から、社会科学を専攻する大学生の統計学に対する知識と意識がどのようなものかを把握し、その結果から、社会科学を専攻する学生に対してどのような統計学とデータ分析の教育実践が有効となるかを探った。

結果と考察については先述の通りで重複となるため繰り返さないが、最後に本論文の限界と今後の課題を3つ述べたい。ひとつめに、本論文の事例の限定がある。本論文の結果は、あくまで一大学の2科目の現時点の事例でしかない。また社会科学といっても社会学専攻の学生が多く、また教員も社会学が専門である。そのため、本論文の結果を単純に日本の社会科学を専攻する学生全体へ一般化することはできないだろう。あくまで今回対象とした大学、学部、学科に限定した上で、この知見から教育実践のヒントを得ることが妥当だと考える。ふたつめに、大学教育の前段階である高等学校の数学や地歴・公民など他科目の教育の把握と検討が不十分である。中学校や小学校においても統計学

やデータ分析に関することは教えられている。高大接続や大学入試のあり方も大きく変動している現在、社会科学系の大学入学者がどのような知識と意識を持って入学してくるかを把握することはますます重要になろう。みっつめに、統計学とデータ分析の教育が各学生の専攻分野においてどのような位置を占めるかのより詳細な検討が課題となる。本論文では社会調査士や学部・学科のカリキュラム上の位置づけについては言及した。くわえて各学問分野の中での統計学とデータ分析の教育の位置づけの検討が必要になろう。例えば本論文で履修学生が多かった社会学では2014年に日本学術会議から「大学教育の分野別質保証のための教育課程編成上の参照基準 社会学分野」が公表されている（策定経過や大まかな内容は笹谷春美（2018）ほか）。社会学に限らず他分野でもこれらの参照基準が公表されている。これらの参照基準は一例であるが、各学問分野で統計学とデータ分析の教育がどのように位置づけられているかの把握と検討が引き続き必要だと考える。

[注]

- 1) 福岡県立大学の社会調査士科目はこのほかにA科目として「社会調査法」、B科目として「社会調査の設計」、E科目として「データ処理とデータ解析Ⅰ・Ⅱ」、F科目として「文化人類学B」と「質的調査法」、G科目として「社会調査実習」がある。
- 2) 「データ分析の基礎」のシラバスにおける授業の概要は以下の通りである。「受講生がデータ分析の基礎的な方法を身につけることを目標とする。現代社会の状況把握や問題解決など目的に応じたデータを入力して内容を読み取り、データ分析を行なって他者との効果的な議論を行うための基本的な方法を学ぶ。

各種の公的統計や調査報告書の入手方法とその適切な読み方、データを数値や図表で表現する方法、その他の基礎的なデータ分析について講義する。また、受講生自身が基礎的なデータ分析を行って、データを活用（理解・分析・提示など）する方法を学ぶ。」

- 3) 「社会統計学Ⅰ」のシラバスにおける授業の概要は以下の通りである。「社会調査の結果や統計データをまとめ、分析するために必要な社会統計学の基礎について学ぶ。記述統計学・推測統計学の基本的な知識と分析手法を中心に、統計的データを適切に読み、まとめ、分析するための方法を学ぶ。そのことを通じて現代社会について様々な角度から適切な分析と議論ができるようになることを目指す。」
- 4) これらの変数は順序尺度であり、厳密には平均値やピアソンの相関係数の分析などを行えないとする立場がある。ただし、社会学など社会科学系の分析ではこれらを行うことも多く、本論文でもこれに従う。

[謝辞]

本論文の作成にあたって、アンケート調査に協力していただいた受講生に感謝いたします。

[参考文献]

- 廣瀬毅士・寺島拓幸, 2010, 『社会調査のための統計データ分析』 オーム社.
- 一般社団法人 社会調査協会, 2018, 「社会調査士カリキュラム詳細」, (2018年10月19日取得, http://jasr.or.jp/for_students/get-sr/curriculum_sr/).
- 一般財団法人 全国大学実務教育協会, 2018, 「上級情報処理士」, (2018年10月19日取得, http://www.jaueb.gr.jp/zaigakusei/license/information_process_high.html).

- 石崎龍二・佐藤繁美, 2018, 「統計演習科目における学生の自己評価に基づいた教育効果の検証 (2017)」『福岡県立大学人間社会学部紀要』26(2): 205-20.
- 岩井紀子・保田時男, 2007, 『調査データ分析の基礎——JGSSデータとオンライン集計の活用』有斐閣.
- 公益社団法人 日本心理学会, 2018, 「認定心理士とは」, (2018年10月24日 取得, <https://psych.or.jp/qualification/>).
- 笹谷春美, 2018, 「『参照基準 社会学分野』の策定経過と概要」『社会学評論』69(2): 203-16.
- 柴田雅博, 2018, 「福岡県立大学人間社会学部における初年次情報リテラシー教育の効果 (2017年度)」『福岡県立大学人間社会学部紀要』26(2): 191-204.
- 東京大学社会科学研究所附属社会調査・データアーカイブ研究センター, 2018, SSJDA Nesstarホームページ, (2018年10月19日 取得, <https://nesstar.iss.u-tokyo.ac.jp/webview/>).

