

研究の概要

近年、ビッグデータとよばれる大量のデータをコンピュータのマシンパワーで解析することによる問題解決の手法が注目されている。また、そのような場面において使われる統計的手法を身につけた人材に関する期待が高い。現在の教育課程では、数学Ⅰにおいて「データの分析」の単元があり、統計の基礎的な内容をすべての生徒が学ぶ。教科「情報」で、数学と連携することで、簡単なデータ分析を用いて、問題を解決しようとする姿勢を育てる授業実践について報告する。

【キーワード】問題解決 情報の科学 社会と情報 R言語 統計 ビッグデータ 他教科連携

1 はじめに

(1) 本校の教科「情報」について

本校は平成19年に〇〇〇〇高等学校と〇〇〇〇高等学校が統合した高等学校である。情報の専門学科である情報理数科と普通科の二つの学科が設置されている。平成26年度入学生のエデュケーション課程では、普通科では1年生の全クラスが「社会と情報」(2単位)を履修し、3年生では選択科目「情報デザイン」(2単位)を選択することができる。「社会と情報」は2名のチームティーチングの形式で授業を行っている。

また、情報理数科は、公立学校としては全国にまだ20校のみの情報の専門学科である。共通教科「情報」の科目は履修せず、専門教科「情報」の科目を卒業までに25単位以上を履修する。専門科目のうちの必修科目は「情報産業と社会」「課題研究」となっている。

平成26年度現在、本校の情報の教科には、教諭4名、臨時任用的講師1名、非常勤講師1名、実習助手2名が配属されている。

(2) 本校の生徒状況について

本校は、平成26年度現在は、1年生8学級、2年生8学級、3年生7学級の合計23学級規模の学校である。普通科・情報理数科の両学科とも、ほとんどの生徒が四年制大学への進学を希望する。普通科では、8割以上が四年制大学へ進学する。情報理数科では、約7割の生徒が四年制大学へ進学し、クラスの約6割の生徒が、工学部・理学部などの大学に進学している。

2 研究の主題と内容

(1) 研究主題の設定

インターネットの普及により、情報提供のプラットフォームとしてのWWWは生活の中に定着している。現在は、総務省や福井県鯖江市といった自治体や、東京メトロなどの企業が、データをXML形式やWeb APIの形で公開することで活用を促進する、「オープンデータ」と呼ばれる動きも始まっている。このように、いわゆるビッグデータと呼ばれるデータが、個人レベルでも利用可能になっていて、またビッグデータの分析を行い、企業の戦略などに活用する「データサイ

エンティスト」という職種についても話題となっている。しかし、この盛り上がりは一時的なバズワード（Buzz Word）にとどまり、これらのブームが落ち着いていく可能性も大きいと考える。しかし、統計的な手法や統計による分析は、あらゆる活動や計画における基礎的な能力となっていくと考えられる。

平成25年から実施の高等学校学習指導要領では、必履修科目の「数学Ⅰ」の中に「データの分析」の単元が新設され、中学校の数学から継続して統計的内容を学ぶようなカリキュラムへと変遷した。これは、現代の数学教育における統計教育の重要性を反映したものである。

学習指導要領において、共通教科「情報」での「社会と情報」「情報の科学」のどちらの科目でも「問題解決」を扱っている。また、教科「情報」は、学習指導要領第2章第10節に「情報科での学習が他の教科・科目等の学習に役立つよう、他の教科・科目等との連携を図ること」とあるように、教科横断型の教科であると定義されている。特に、「公民科及び数学科などとの関連を図るとともに、教科の目標に即した調和のとれた指導が行われるように留意すること」という記述があるように、数学との関連も強い。

そこで、問題解決に用いる手法として、「数学Ⅰ」の「データの分析」で扱われる統計等の内容を利用し、しかし数学教育ではなく情報教育として問題解決に利用できる指導内容の開発について検討した。

（２） 研究の目標

共通教科「社会と情報」「情報の科学」のどちらにも含まれる単元である「問題解決」の題材として、統計処理を得意とするソフトウェアの利用と、統計の手法を学ぶことで、データを分析して、仮説を立てて物事を推論する能力を身につけさせる授業事例を開発するものである。

（３） 研究の内容

情報の授業として、数学の教員と連携し、ソフトウェアを利用して統計的な内容を学び、またソフトウェアを活用することでデータから推論していく能力を身につけることができる授業実践について研究していく。その際に、数学Ⅰの授業で取り上げられる「データの分析」の単元のうち、平均、分散、標準偏差、相関関係、などの内容を学習していく。

３ 研究の具体的内容

（１） 学習指導要領における問題解決

高等学校学習指導要領には、「社会と情報」「情報の科学」それぞれの学習内容として問題解決について取り上げられている。

ア 「社会と情報」における問題解決

学習指導要領では第2章 第1節の「内容とその取り扱い」に「問題解決」が登場する。

第1節 「社会と情報」

（４）望ましい情報社会の構築

ウ 情報社会における問題の解決

情報機器や情報通信ネットワークなどを適切に活用して問題を解決する方法を習得させる。

学習指導要領解説には次のような記述がある。

問題の分析の段階では、問題を解決するために必要な事柄を収集・整理する方法を学ばせる。収集や整理の際、情報手段も活用して多様な活動ができるようにすることが重要である。また、数学科の学習と連携して、統計的な手法を活用させることも考えられる。収集方法としては、Web サイトや新聞・書籍からだけでなく、ブレーストーミング、アンケート調査、インタビューなどを行うことが考えられる。整理する方法としては、得られた情報を関連付けて図解したり、表を作成して一覧の形式にまとめたり、適切な種類のグラフを作成したりすることが考えられる。さらに、様々な文章を読み解き、意味や内容を分析し有用な情報を見つけ出す手法であるテキストマイニング(Text Mining)なども考えられる。

イ 「情報の科学」

「情報の科学」では第2章2節で問題解決が扱われている。

(2) 問題解決とコンピュータの活用

ア 問題解決の基本的な考え方

問題の発見、明確化、分析及び解決の方法を習得させ、問題解決の目的や状況に応じてこれらの方法を適切に選択することの重要性を考えさせる。

(3) 情報の管理と問題解決

イ 情報の蓄積・管理とデータベース

情報を蓄積し管理・検索するためのデータベースの概念を理解させ、問題解決にデータベースを活用できるようにする。

(2) 数学と連動する内容について

現在の高等学校学習指導要領の数学では、「数学Ⅰ」において「データの分析」が必須の単元となっており、この単元では教科書や板書といった座学だけではなく、コンピュータを活用した学習が理解に役立つと考えられる。また、数学Ⅰは必修科目であり、ほとんどの学校で1年次に開設されているため、教科「情報」の科目がその学校で何年次に設置されていても、連携をとることは可能であると考えられる。

(3) 数学Ⅰの学習方法

数学Ⅰの学習指導要領のデータの分析の単元には、以下の内容が記されている。

(4) データの分析

統計の基本的な考え方を理解するとともに、それを用いてデータを整理・分析し傾向を把握できるようにする。

ア データの散らばり

四分位偏差、分散及び標準偏差などの意味について理解し、それらを用いてデータの傾向を把握し、説明すること。

イ データの相関

散布図や相関係数の意味を理解し、それらを用いて二つのデータの相関を把握し説明すること。

「学習指導要領 数学編 第2章 第1節 数学Ⅰ 3 内容と内容の扱い」より抜粋

数学Ⅰ担当の教員と指導内容を相談し、情報の授業内でソフトウェアを用いて学習する範囲を決定した。当該学年が「数学Ⅰ」での「データの分析」を未学習の状態である場合には、平均値・分散や相関係数といった基本的な統計量を扱うだけにとどめることとした。

本校の情報理数科では、数学的な内容を扱った学校設定科目「情報数理」を1年から3年まで設置している。本研究では、情報理数科の1年生が全員履修する、数学Aの2単位分の内容を含んだ情報の専門科目「情報数理A」（3単位）の中で、数学Ⅰとも連動させながら実践を行った。

（４） 統計処理ソフトについて

Windows で利用可能な統計処理ソフトウェアは、ベル研究所によって開発された「S 言語」や、現在は IBM が権利を保有している「SPSS」などがあるが、どちらも商用ソフトウェアである。また、統計処理専門ではないが、表計算ソフトウェアとしては Microsoft Excel（以下 MS Excel）がデファクトスタンダードとなっている。しかし、近年、S 言語と同様の処理を行えることを目標としたオープンソースソフトウェアである R 言語が普及しつつある。

R 言語と MS Excel では、どちらのソフトウェアでも基本的な統計量の計算は可能であるが、MS Excel では、操作が複雑になってくる場面も生じる。

そこで、共通教科「情報」で扱うえる程度の情報機器の操作と R 言語の学習をとおして、社会に存在する大量のデータから、問題の発見や分析を行うことができる能力が向上するか、アンケート調査によって調べる。

（５） R 言語と MS Excel の比較

MS Excel はセルに数値や文字を入力し、表を完成させる表計算ソフトで、現在は表計算ソフトのデファクトスタンダードとなっている。セルには、さまざまな関数も入力可能で、計算や文字列操作、また統計の命令もあらかじめ備わっている。

それに対して R 言語は、オープンソースソフトウェアとして開発されており、基本的にはコマンドラインから命令を入力し、各種処理を行うことを基本としている。しかし、GUI での操作を補うために、統合環境である RStudio（<http://www.rstudio.com/>）などの周辺ソフトウェアも開発されている。

それぞれの違いを理解してもらうために、基本的な統計量を求める操作を MS Excel での関数と R 言語での命令で求めた場合の記述方法について比較する。

表 1 主な統計量を求める MS Excel 関数と R 言語の命令

主な統計量	MS Excel の関数	R 言語の命令
合計	SUM(セルの範囲)	sum(値 or 配列)
平均	AVERAGE(セルの範囲)	mean(値 or 配列)
分散	COVAR(セルの範囲)	var(値 or 配列)
中央値	MEDIAN(セルの範囲)	median(値 or 配列)
相関係数	CORREL(配列 1, 配列 2)	cor(配列 1, 配列 2)

ここで相関関係について整理する。「相関係数」とは、2つの変量からなるデータの間に、「一方が増加すればそれにしたがって他方も増加する または他方が減少する」という傾向がみられるとき、2つの変量の間に「相関がある」または「相関関係がある」という。2つの変量からなるデータにおいて、一方が増加すると他方も増加する傾向がみられるとき、2つの変量には「正の相関がある」という。一方が増加すると他方が減少する傾向がみられるとき、2つの変量には「負の相関がある」という。どちらの傾向もみられないときには「相関がない」または「相関関係がない」という。

また、相関係数について述べる。相関係数 r は、 $-1 < r < 1$ の値を取り、 r の値が1に近いほど「強い正の相関」、 -1 に近いほど「強い負の相関」があることを示す。

相関係数の値と相関の強さはさまざまな表現があるが、以下の表現を採用する。

$0.7 \leq r \leq 1.0$	強い相関がある
$0.4 \leq r \leq 0.7$	中程度の相関がある
$0.2 \leq r \leq 0.4$	弱い相関がある
$0.0 \leq r \leq 0.2$	ほとんど相関がない

最後に「因果関係」について述べる。「ある事象 A が原因となり事象 B が起こる」または「事象 A の変化によって、事象 B に変化が起きた」など、複数の事象の変化について原因と結果が判明している関係のことを指す。しかし、別の要因 C が、事象 A 及び B を引き起こすことにより事象 AB 間に相関が生じることがあるが、A と B に因果関係があるとは限らない。このような関係を疑似相関と呼ぶ。

本研究では、相関や相関係数は扱うが、「母集団をどう捉えるか」や「帰無仮説」などの考え方は紹介していない。そのため、統計学としては有意ではない事象も多いが、情報機器を活用して統計量を求めた上で考察するという試みを目的として、そのまま扱っている。

4 実習環境について

(1) 使用 PC

この研究では、生徒個人の R 言語を用いた学習と、グループでのディスカッションとで利用するコンピュータ室を変えている。生徒個人の R 言語の学習には、コンピュータ室のデスクトップ PC を使用した。生徒がグループに分かれてディスカッションをする授業では、ノート PC をグループで1台の割合で使用した。

(2) ソフトウェアの利用形態

今回は、コンピュータ室の PC にソフトウェアをインストールすることはせず、フォルダ単位でコピーするだけで利用可能である Portable 版を利用した。Windows 向けの Portable 版ソフトウェアは、Portable Apps (<http://portableapps.com/>) のサイトにまとめられて、さまざまなソフトウェアが入手可能である。今回は R-Portable (<http://portableapps.com/node/32898>) を利用した。平成 26 年 12 月現在の R-Portable の安定版は Ver 3.1.2 (サイズ 55.9MB) である。

(3) R-Portable の情報

R-Portable をダウンロードし、解凍すると以下のようなフォルダ構成となる。この中の「r-portable.exe」のファイルが実行ファイルで、このフォルダ構成を変更しなければ、フォルダごとコピーすることで、任意の PC 上で実行可能である。

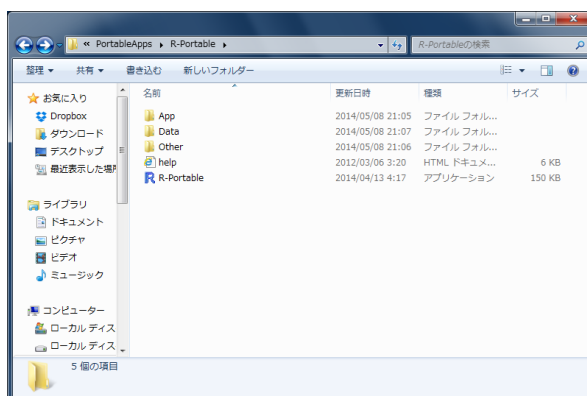


図 1 R-Portable のフォルダ構成

(4) 生徒への配布方法

コンピュータ室の PC にインストールされていないソフトウェアを授業で使う場合は、

- ① コンピュータの環境復元を外してインストール
- ② 集中管理ソフトを利用し、生徒各自のドキュメントフォルダに配布
- ③ サーバの共有フォルダにソフトを配置し、生徒が各自の PC にコピー

などの方法が考えられる。今回は、③の方法で生徒が自席の PC の D ドライブ(環境復元をかけていないドライブ)にコピーし、授業中はローカル HDD の R-Portable を使用することとした。

5 授業展開

(1) 授業例 1 相関係数とデータ間の関係

数学 I のデータの分析には、相関関係を調べ、「正の相関がある」か「負の相関がある」、「相関がない」のどれであるかを考える内容が含まれている。そこで、生徒の視点でも推測が可能である実データを利用し、R 言語を利用してデータ群の相関係数を求め、その相関係数とデータの単位から、どのようなデータであるかをグループ単位で推理させる内容とした。

ウ 事前学習

R-Portable の使用方法について、事前学習として数時間の学習を行った。R 言語では、変

数の他に、配列（ベクタ）を利用することができる。複数の値を配列に格納することで、以降は配列名のみで操作することが可能となる。

例 1 5 つのデータを配列 `sample` に格納してから、配列の平均を求める操作

```
> sample <- c(10, 20, 30, 40, 50)
> mean(sample)
[1] 30
```

例 2 2 系列のデータがある CSV ファイル「data.csv」から配列 `data` に値を格納し、2 系列の相関係数を求める操作

```
> data <- read.csv("data.csv")
> cor(data)
```

エ 利用したデータ

準備された 4 つのデータセットを利用して、データについての「平均値、中央値、最大値・最小値、相関係数」などを R 言語で求めていく。求めた統計量を用いて、プリントの項目についてグループで検討し、まとめていく。相関係数が分かった後は、「相関関係」と「因果関係」についてグループで相談させる。

（ア） データセット 1 高校生の「身長」と「体重」の関係

本校で採用している数学 I の教科書である「高等学校 数学 I」（数研出版）より、相関係数の例題として使われている数値をデータ化して用いた。（表 2）

表 2 男子高校生の身長と体重との関係の例

身長 (cm)	168.4	164.5	171.2	173.0	162.3	170.8
体重 (kg)	59.0	58.4	60.5	65.6	54.2	61.4

相関係数 r は $r = 0.887$ となっており、身長と体重の間には正の相関があることがわかる。

（イ） データセット 2 「柏市の基準地価」と「土地と柏駅との距離」の関係

基準地価は、さまざまな観点によって決定されるが、その際に最寄り駅からの距離が近ければ近いほど利便性が高いため、地価が高くなる傾向がある。そのため、最寄り駅への距離が近ければ近いほど地価は高いという因果関係が成り立つと思われる。

最寄り駅からの距離と基準地価のデータは千葉県の Web サイトの平成 24 年地価調査のデータから、最寄り駅が「柏駅」であるものを抜粋して筆者が作成した。（表 3）

<http://www.pref.chiba.lg.jp/youchi/toukeidata/chika/documents/kasiwa.pdf>

表 3 基準地価と最寄り駅（柏駅）との関係

基準地価 (円)	70,500	73,500	99,000	182,000	165,000	108,000	64,400	105,000
駅への距離 (m)	4,800	2,300	2,100	700	600	2,300	4,600	1,600

相関係数 r は $r = -0.851$ となる。この場合は、因果関係があるので、「相関関係があり、因果関係がある」という関係になる。

(ウ) データセット3 男性の年代ごとの最高血圧・最低血圧と平均年収の関係

厚生労働省の「平成23年国民健康・栄養調査 第2部 身体状況調査の結果」のデータを利用し、男性の平均的な血圧を10歳刻みにしたものを20代から60代まで提示する、平均年収に関しては、国税庁の「平成24年分 民間給与実態統計調査 ー調査結果報告ー」のデータを利用した。(表4)

表 4 成人男性の最高・最低血圧と平均年収との関係

年齢区分	20代	30代	40代	50代
最高血圧 (mmHg)	119.6	124.5	126.4	139.6
最低血圧 (mmHg)	74.3	80.2	82.4	83.6
平均年収 (万円)	313.5	464.5	587.5	626

ここでは、年齢が高くなることによる最高血圧・最低血圧の上昇と、勤続年数や年功序列によると考えられる年収の上昇が、擬似相関としてみることもできる関係が存在する。

つまり、年齢という要素を明らかにしない状態で相関係数を求めてみると「血圧が高い人は年収が高い」という仮説でも相関を見出してしまう。実際に相関係数を求めてみると最高血圧と平均年収の相関係数 $r1$ は $r1 = 0.833$ 、最低血圧と平均年収の相関係数 $r2$ は $r2 = 0.985$ となる。

(エ) データセット4 国民1人あたりのチョコレート消費量とノーベル賞受賞者数の関係

ヨーロッパのCAOBISCO(ヨーロッパチョコレートビスケット菓子工業協会)のWebサイトにある2011年のチョコレート菓子の消費量ランキングとWikipediaのノーベル賞受賞者の人口別ランキングを参考に、筆者がデータセットを作成した。(表5)

表 5 国別のチョコレート平均消費量とノーベル賞受賞者数（人口あたり）の関係

国名（省略形）	SUI	GER	UK	NOR	DEN	SWE	BEL	AUT	FIN	FRA
チョコレート消費量 (Kg/1 人)	10.55	9.57	9.45	9.24	7.85	7.68	7.3	7.05	6.51	3.94
ノーベル賞受賞者数 (人/1000 万人)	30.9	12.58	19.17	21.81	24.92	30.30	9.01	24.72	7.37	9.18
国名（省略形）	ITA	POL	HUN	ESP	IRL	ROM	POR	GRE	JPN	BUL
チョコレート消費量 (Kg/1 人)	3.25	3.17	2.89	2.64	2.02	2	1.91	1.83	1.82	0.68
ノーベル賞受賞者数 (人/1000 万人)	3.28	3.14	9.04	1.71	12.97	1.39	1.89	1.80	1.49	1.38

国別のチョコレート平均消費量と人口 1000 万人あたりのノーベル賞受賞者の相関係数 r は $r = 0.813$ であった。

これは、疑似相関とまでは言えず、多くの国別の指標の中の相関関係を示す一つの例であると考え。データなどからあたかも因果関係があるように読み取ることもできるが、因果関係までは示せないといった事例である。

オ 各グループでの話し合い

それぞれのデータセットについてグループで話し合った意見を各グループに答えさせた。

データセット 1：強い正の相関があり、身長が増えると体重が増える傾向がある

データセット 2：負の相関があり、駅からの距離が短くなると地価が上がる傾向がある

データセット 3：年収が高いほど、年齢も高く、血圧も高くなる傾向がある

データセット 4：チョコレートの消費量が多い国はノーベル賞受賞者が多い傾向があるなどの表現が見られた。

この事例では、恣意的なデータであるが、「正の相関」「負の相関」が現れて、それぞれ因果関係も関係するデータセットと、「正の相関」が見られるが疑似相関となっているものを用意して、相関係数が求められたからといって、それらの値に因果関係が見られるわけではないことを伝えるための演習とした。

グループで話し合いを行った際のワークシートを次に示す。

_____班

メンバー

(役割: 班長, 発表者 1, 発表者 2, 操作, 操作補助)

出席番号	氏名	役割	出席番号	氏名	役割

【データセット 1 について】 データ単位は (cm) と (kg) である。

二つの系列の相関係数は・・・

これから言えること

【データセット 2 について】 データ単位はそれぞれ (円) と (メートル) である。

二つの系列の相関係数は・・・

これから言えること

【データセット 3 について】 データ単位は max,min (mmHg) と sal (万円) である。

3つの系列の相関係数は・・・

これからいえること

【データセット 4 について】 データ単位は cons (Kg/一人) と prize (人/1,000 万人)

2つの系列の相関係数は・・・

これからいえること

(2) 授業例2 2変数の回帰直線

R 言語には、統計学習用として、サンプルのデータセットが同梱されている。今回は、2 変数の相関係数と回帰直線の学習をするために、1920 年代の自動車の速度と制動距離のデータセットを利用した。このデータセット cars には、時速 (speed) とその時速での制動距離 (dist) の数値の組が、50 組用意されている。(表 6)

表 6 自動車の速度と制動距離のデータセットの一部

速度(mph)	4	4	7	7	8	9	10	10	10
制動距離(m)	2	10	4	22	16	10	18	26	34

このデータセットは R 言語のパッケージに同梱されており、直接指定して呼び出すことができる。入力させた R 言語のコマンドは次のとおりである。

```
cars.lm <- lm(dist~speed,data=cars)
```

```
cars.lm$coefficients
```

これにより、この 2 つのデータから回帰直線を引く場合の直線の傾きと切片を得ることができ、この自動車の速度と制動距離には、以下の式で表すことができる回帰直線を引くことができる関係があることがわかる。

$$\text{制動距離}(\text{dist}) = 3.93 \times \text{速度}(\text{speed}) - 17.6$$

また、このときの相関係数は $r = 0.806$ となり正の相関が見られる。



図 2 得られた結果を記入している様子

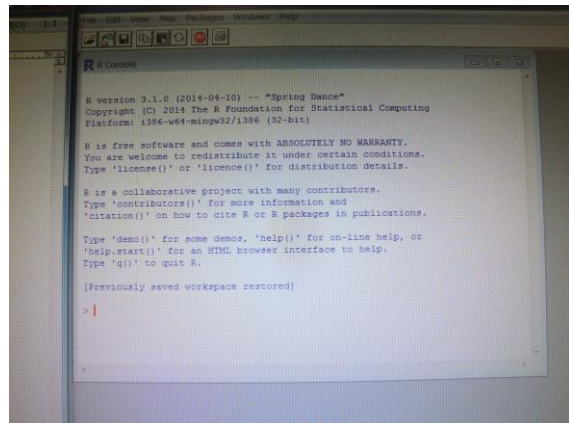


図 3 R 言語使用中の PC 画面

(3) 授業例3 相関があると考えるデータを探す

相関関係の考え方や相関係数の求め方を学んだ後、「あるデータに対して相関がありそうなデータを探す」という題材で授業を行った。

今回は、新聞の報道に「都道府県別の大学進学率と全国学力テストの相関」という記事が掲載されたことをきっかけに、「都道府県別の大学進学率」に関係する（正の相関のある）都道府県別の統計データを見つけることができるか、というテーマに設定した。

このテーマについては、社会学や教育学の大きな問題となる命題ともいえるため、正解と言える物は存在しない。そこで、生徒に各自が考える「相関があると思う都道府県別のデータ」を見つけてもらい、CSV データに加工した上で、R 言語上で相関係数を求めて、その値を元にレポート形式で報告する内容とした。

文部科学省の平成 25 年度学校基本調査から、高校 3 年生の大学進学率（%）、全国学力テストの国語 A 問題正解率（%）、算数・数学 A 問題正解率（%）の 3 つの都道府県別の統計データを MS EXCEL にまとめ、生徒に配布した。ここから、大学進学率-国語 A 正解率、大学進学率-算数・数学 A 正解率、国語 A 正解率-算数・数学 A 正解率の 3 つの都道府県別のデータについて、相関係数を求めさせた。（表 7）

表 7 都道府県別の大学進学率と全国学力テスト正解率との相関

項目の組み合わせ	相関係数
進学率 - 国語 A 正解率	0.133
進学率 - 数学 A 正解率	0.336
国語 A 正解率 - 数学 A 正解率	0.818

この後で、生徒に「各都道府県の大学進学率に相関があると考える都道府県別の統計データを見つけ、相関係数を求めてレポートとして提出」という課題を出題した。

この問いに関しては、完全な正解はない問題だと考える。各都道府県の高校生の大学進学率の違いは、さまざまな要素から生まれ、単一のデータと因果関係は証明できない。しかし、擬似相関となる項目を含めて、各都道府県内の大学設置数、人口規模、世帯平均年収、など相関が見られる項目もあると考えて出題した。

生徒が提出したレポートには、以下のような仮説と相関係数を調べてきたものがあった。

「大学進学率は、学習塾・予備校費用が多い県ほど良い」	相関係数 0.4
「大学進学率は、各県の大学の数と相関の関係にある」	相関係数 0.58
「大学進学率は、高校在学者数が多い県ほど高い」	相関係数 0.57
「大学進学率は、各県の人口と相関がある」	相関係数 0.58
「大学進学率は、一人当たりの学校教育費と正の相関の関係にある」	相関係数-0.21
「大学進学率は、在学者一人当たりの学校教育費と相関がある」	相関係数-0.36
「大学進学率は、都道府県別の大中小の企業数と相関がある」	相関係数 0.57

また相関係数を求めることはできていなかったが、「大学進学率は、各県の農業就職率と負の関係にある」という仮説を立てて途中まで調べている生徒もいた。

レポートに課した時間が短かったため、さまざまなデータについて相関を求めることができなかったが、正解の無い問いに対して、自らが解と思われるものを探すという授業形式ができた。

6 まとめ

(1) アンケート結果

授業以前と授業以後での「問題解決能力に関する自己評価」をアンケートで調査した。問題解決力の自己評価項目を確認するためにアンケートに記載した質問は以下の 5 つである。

設問 1 問題解決能力があると感じる

設問 2 データについての読み取り能力があると感じる

設問 3 コンピュータを利用して、物事を分析する力がある

設問 4 問題解決について、コンピュータを利用して取り組むことができる

設問 5 問題解決をするために、コンピュータは有効である

これらについて、授業前と授業後にアンケート調査を行い、その結果をまとめた。

ア 授業例 1 についてのアンケート (H25 年度 1 年生)

表 8 授業例 1 の授業以前の自己評価アンケート

	設問1	設問2	設問3	設問4	設問5
あてはまらない	4	5	7	1	1
あまりあてはまらない	24	19	20	17	3
ややあてはまる	13	16	12	14	19
あてはまる	0	0	2	8	18
わからない	0	1	0	1	0

表 9 授業例 1 の授業以後の自己評価アンケート

	設問1	設問2	設問3	設問4	設問5
あてはまらない	1	3	2	1	0
あまりあてはまらない	14	14	14	5	2
ややあてはまる	20	20	19	26	13
あてはまる	4	2	6	8	26
わからない	2	2	0	0	0

イ 授業例 2 についてのアンケート (H26 年度 1 年生)

表 10 授業例 2 の授業以前の自己評価アンケート

	設問1	設問2	設問3	設問4	設問5
あてはまらない	3	4	3	2	0
あまりあてはまらない	14	16	12	12	3
ややあてはまる	20	15	20	16	12
あてはまる	0	1	0	6	21
わからない	1	2	3	2	2

表 11 授業例 2 の授業以後の自己評価アンケート

	設問1	設問2	設問3	設問4	設問5
あてはまらない	2	2	2	1	1
あまりあてはまらない	14	16	14	11	1
ややあてはまる	20	17	20	17	13
あてはまる	1	1	1	8	23
わからない	1	2	1	1	0

(2) 分析

授業前と授業後のアンケートについて比較する。

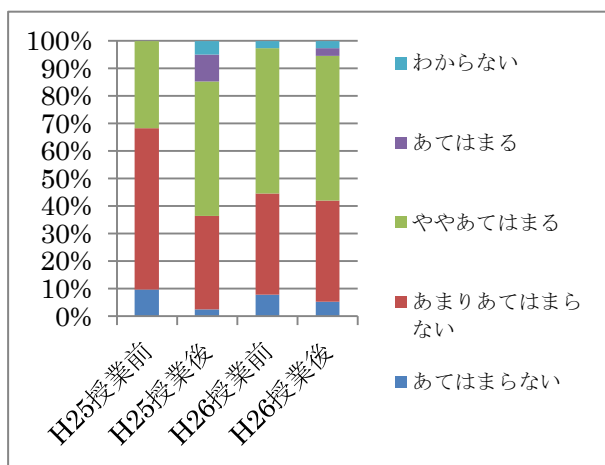


図 4 設問1の回答の推移

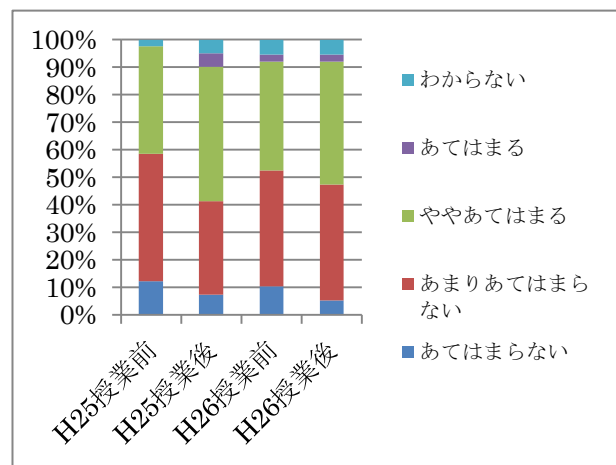


図 5 設問2の回答の推移

設問1の「問題解決能力があると感じる」についての回答は、2つのクラスとも「あてはまらない」「あまりあてはまらない」が減少し、「あてはまる」が微増した。

設問2の「データについての読み取り能力があると感じる」について、2つのクラスとも「あまりあてはまらない」「あてはまらない」が減少した。

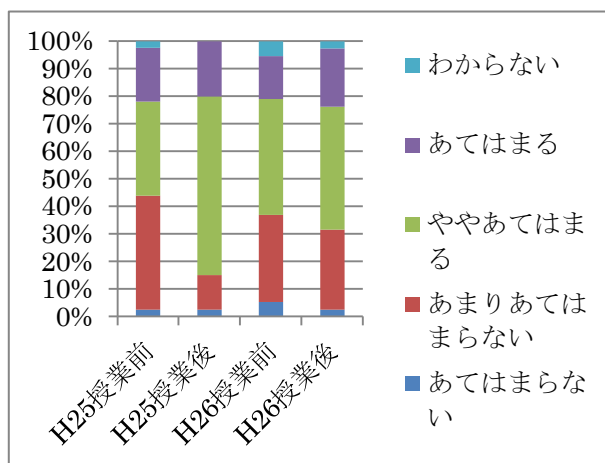


図 6 設問3の回答の推移

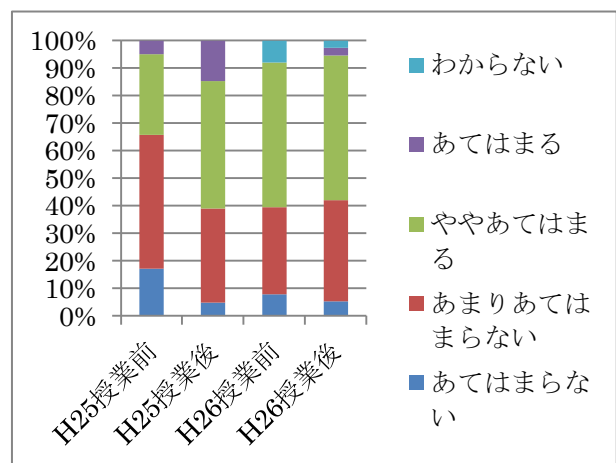


図 7 設問4の回答の推移

設問3の「コンピュータを利用して、物事を分析する力がある」についての回答は、2つのクラスのどちらも「あまりあてはまらない」の回答が減少している。

設問4の「問題解決について、コンピュータを利用して取り組むことができる」についての回答は、授業例1を行ったクラスでは「あまりあてはまらない」の回答が大きく減少している。

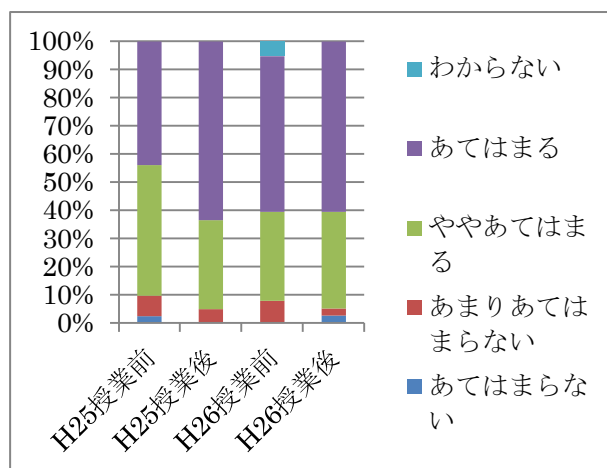


図 8 設問5の回答の推移

設問5の「問題解決をするために、コンピュータは有効である」についての回答は、2つのクラスとも「あてはまらない+あまりあてはまらない」の合計は減少し、「あてはまる」が増加した。

授業以前に「問題解決能力の有無」を問う設問に対しては、自己評価を低めに回答する傾向があるが、「問題解決の手順」の最初の手順となりうる「データの分析」の手法を体験することによって、「問題解決能力の有無」についての自己評価が変化していると考えられる。

(3) 生徒の感想

授業例1の感想

- ・中学までのように自分で計算して解くのではなく、コンピュータを使って問題を解くのはとても楽しく便利だと思った。
- ・紙に直接書き込んでデータを分析するより、コンピュータを活用した方が便利であることを実感した。
- ・コンピュータを検索などにしか使わなかったもので、このようなことができてすごいと思った。

授業例2の感想

- ・経験したことのない授業だったのでとても楽しかった。
- ・今まで数学でPCを使うということがなかったけれど、1年間やってきていろいろ活用方法があるのだと思った。
- ・データを使って物事を予想したりする授業は、身近なものを調査してみようという気持ちにさせてくれた。

どちらのクラスでの感想でも、数学の統計分野でコンピュータを使うことへの新鮮な印象や、

データの分析をすることが身近な問題解決の一步であるにとらえる感想が見られた。

このように数学Ⅰの「データの相関」の題材でも、実際のデータに対して分析を行うことで、「統計の手法が問題解決に結びつく」という意識を高めることができたと考える。

(4) 今後の課題

本研究では、学校設定科目「情報数理 A」の授業時間を利用して実践を行った。この科目は、数学 A の内容を含む時間 2 単位分と情報的な内容を中心とした時間 1 単位分の内容で授業を進行している。また、1 年生を対象としているため、コンピュータ教室の各種説明や基本的なコンピュータ操作を含んでいたため、実践例をあまり重ねることができなかった。

いろいろな学校の情報の授業でこの実践を活用していただくためには、MS Excel へのアドインとして R 言語を利用できる「RExcel」や R 言語専用の統合環境「RStudio」など、GUI を利用した容易な操作体系を持つソフトウェアの選定や、統計として扱う授業の素材データなどについてさらに研究を続けなければならないと考えている。

7 おわりに

情報の学習指導要領では、「問題解決」を取り扱うと定められているが、多くの教員が「高校生にふさわしい問題解決とは何か」という疑問を持ち、日々研究を重ねている。

インターネットでのショッピングや、ポイントカード・電子マネーを利用する際のコンビニエンスストアなどでの個人情報の取得も当たり前となってきた。また、ビッグデータの名のもとに、個人情報やセンシティブな情報（機微な情報）が蓄積され、消費者としては意識をしなければならない社会となってきた。しかし、これからの社会を生きていく者として、統計的なものの味方や ICT 能力を活用して、さまざまな問題を解決して行かなければならない。そのような社会を生きぬく力を身につけるための一環として、この研究が教科「情報」の授業担当者の参考になればと考える。

参考文献

- 『不透明な時代を見抜く「統計思考力」』神永正博（ディスカバー・トゥウェンティワン）
- 『統計データは面白い！－相関図でわかる経済・文化・世相・社会情勢のウラ側－』本川裕（技術評論社）
- Web サイト「社会実情データ図録」 <http://www2.ttcn.ne.jp/honkawa/>
- 『統計学がわかる ハンバーガーショップでむりなく学ぶやさしく楽しい統計学』向後千春・富永敦子（技術評論社）
- 『データ分析ってこうやるんだ！実況講義 身近な統計数字の読み方・使い方』吉本佳生（ダイヤモンド社）
- 『データサイエンティスト養成読本』（技術評論社）
- 『EXCEL&R 対応 統計処理ポケットリファレンス』涌井良幸、涌井貞美（技術評論社）
- 『別冊宝島 最強のスキル「統計学」必須のツール「ビッグデータ」』（宝島社）