

# データ分析による問題解決

早稲田大学高等学院 教諭  
武沢 護

次期学習指導要領(2022年から年次進行で実施)における高等学校情報科の概要が明らかになった<sup>1)</sup>。2003年から始まったこの教科は、「情報A」、「情報B」、「情報C」から「情報の科学」、「社会と情報」の選択必修の時代を経て、「情報I(仮称)」(必修修科目)、「情報II(仮称)」(選択科目)と改訂される予定である。そして内容としては、問題発見・解決にあたり、解決すべき問題(事象)を情報の視点から眺めたとき、どのように情報技術を適切かつ効果的に活用するかという力を育むことを目標することになった。

本校では2003年に情報科を設置して以来、指導内容に統計やデータ分析の内容を扱い、そして「数学I」におけるデータの分析の単元や3年「総合的な学習の時間」などとも連携をとってきた。

ここでは、問題解決の一つの手法として高校生が身につけるべき統計やデータ分析を情報科の指導内容としてどのように取り入れてきたかを紹介する。

## 1. 第2学年の授業内容

本校では第1学年、第2学年「社会と情報」(1単位)と分割履修している。そして、データ分析は第2学年の後半に実施している。

表1 第2学年 2・3学期の授業内容

2年 (1単位)		■: 単元 ◆: 評価
2 学 期	■データ分析 ・基礎的な統計量について(平均、分散、標準偏差、相関係数、回帰直線など)	
	◆統計に関する筆記試験	
	■表計算ソフトウェアでのデータ分析I ・表計算ソフトウェアでの統計処理	
	◆表計算ソフトウェアに関する実技試験	
3 学 期	■表計算ソフトウェアでのデータ分析II ◆表計算ソフトウェアに関する実技試験 ◆データ分析に関する課題レポート (探究活動)	

(単元の指導目標)

さまざまなデータをもとに、適切な統計的手法に

よって分析し問題解決できるようにする。

(単元の評価規準)

- ・基礎的な統計的手法を身につけることができる。(知識・技能)
- ・問題解決の場面において、収集、選択したデータをもとに適切な判断ができる。(思考力・判断力)

## 2. 第2学年3学期の授業内容

3学期の授業内容(6回分)について具体的に説明する。ここでは表計算ソフトウェア(Excel)を用いた実習中心の学習活動を展開する。

表2 3学期の授業内容

回	■表計算ソフトウェアでのデータ分析II
1	データベースについての説明
2	表計算ソフトウェアを用いたデータ操作①
3	表計算ソフトウェアを用いたデータ操作②
4	データ分析①(アンケート集計)
5	データ分析②(回帰分析) ◆課題レポートについて(探究活動)
6	◆表計算ソフトウェアの実技試験

### (1) 授業(第2回, 第3回)

次の表(sheet1)にあるような50人のデータに対して、表計算ソフトを用いてデータの操作を行う。

sheet1						
No	性別	地域	氏名	国語	数学	英語
1	男	東京	青木	36	64	89
2	男	千葉	小林	56	32	23
3	女	東京	植村	43	45	44
4	女	埼玉	小原	11	97	76
・	・	・	・	・	・	・
・	・	・	・	・	・	・
50	男	埼玉	渡辺	54	85	67

- ・検索、置換
- ・並べ替え
- ・抽出(フィルタ機能)

- ・参照機能(VLOOKUP)
- ・条件カウント機能(COUNTIF)
- ・データベース関数(DSUM, DCOUNTA など)

例題：次を求めなさい。

- ① 埼玉県出身の英語の合計点
- ② 千葉県出身の国語の平均点
- ③ 数学の点数が 50 点未満の人数
- ④ 東京都出身で英語が 50 点未満の人数
- ⑤ 合計点が 120 点以上 180 点以下の人数
- ⑥ 合計点が 120 点未満 180 点より大きい人数

## (2) 授業(第 4 回)

ここでは、あるアンケート集計結果のデータをもとにクロス集計を行う。授業内容については次に詳しく述べる。

### 【授業案】

(本時の目標)

アンケート集計をもとに、Excel を用いることにより、簡単なデータ分析ができることを学ぶ。

(本時の展開)

○前時までの簡単なデータベースに対する復習、既習事項を確認する。

○実習 1：sheet2 にあるデータに対して、どのような分析が可能かを問う。

- ・「2 項目間のクロス集計」の説明
- ・Excel の機能「ピボットテーブル」を用いたクロス集計の実習
  - 性別×学年、性別×問 1、問 1×問 2、
  - 性別×系、問 1×問 3
- ・クロス集計の結果のグラフ化

○実習 2：sheet3 を用いてどのようなクロス集計をすることで何が分析できるかを考えさせる。

○まとめ

### 【授業の実際】

次に、具体的にどのように授業を展開しているかについて述べる。

○実習 1：

次の sheet2 はある学校の男女 30 人の生徒に対する問 1 から問 5 までのアンケートへの回答を一覧表として作成したものである。回答は 5 者択一の形式(1～5 を選択)とした。

この sheet2 のデータについて、Excel の機能「ピボットテーブル」を用いてクロス集計する。さらにクロス集計の結果をグラフ化することで、データ内容を分析する。

sheet2

性	学年	系	問 1	問 2	問 3	問 4	問 5
女	2 年	理	2	3	4	1	1
男	1 年	文	3	2	5	1	3
男	2 年	文	4	5	3	4	4
男	3 年	理	1	2	1	4	1
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
男	3 年	文	4	5	3	1	2
男	1 年	文	3	2	5	5	4
女	2 年	理	5	4	2	2	5

○クロス集計の結果

表 3 学年×性別

学年	性別		
	女子	男子	総計
1 年	2	10	12
2 年	2	3	5
3 年	5	8	13
総計	9	21	30

表 4 問 1×問 2

問 2	問 1					総計
	1	2	3	4	5	
1	1	1			1	3
2	2	1	4	2	2	11
3	1	1	1			3
4	1	1	2	2	1	7
5	1	1		4		6
総計	6	5	7	8	4	30

問：表 4 から読み取れることは何かを述べよ。

### 【情報活用能力育成のねらいの観点から】

次期学習指導要領における情報活用能力の三つの柱、すなわち「知識・技能」(何を理解しているのか、何ができるのか)、「思考力・判断力・表現力等」(理解していること・できることをどのように使うのか)、そして「学びに向かう力、人間性等」(どのように社会・世界と関わり、よりよい人生を送るのか)の観点からこの授業を整理すると次のようになる。

## ○指導事例と情報活用能力との関連

- ・与えられたデータに対してどのような分析ができるか。(知識・技能)
- ・与えられたデータに対してどのように表計算ソフトウェアを活用できるか。(思考力・判断力・表現力等)
- ・与えられたデータに対して、どの項目間のクロス集計に問題意識をもち、解決しようとするか。(思考力・判断力・表現力等)
- ・他の資料においても、クロス集計での分析を積極的に取り入れようとするか。(学びに向かう力、人間性等)

## ○情報活用能力の育成に関する工夫

- ・どのような情報手段の活用が有効かを考えさせる。(主体的な学びの促進)
- ・与えられたデータに対して、どのような分析により新たな価値が発見できるか考えさせる。(深い学びの促進)

## ○実習 2 :

次の sheet3 はアメリカ合衆国のある地域の小学校 4～6 年生の児童 90 名が、成績(問 1)・スポーツ(問 2)・容貌(問 3)・お金持ち(問 4)の 4 つを重要と思う順にランク付けしたものである。

男女、学年、人種、地域(都会・郊外・田舎)を各学区からの標本が約 1/3 ずつになるように層化抽出した。このデータをクロス集計することでデータ分析する。

sheet3

性	年	人種	地域	問 1	問 2	問 3	問 4
M	4	W	R	2	3	4	1
F	4	W	S	2	1	4	3
F	4	B	U	3	4	1	2
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
M	6	B	R	1	2	4	3
M	6	O	U	2	1	3	4
F	6	W	S	4	3	1	2

## ○活動のポイント

十分な時間を確保して、生徒同士がいろいろ議論しながら作業を進め、どのようなクロス集計が新し

い価値を生み出しているかを考えさせる。

## (3) 授業(第 5 回)

ここでは回帰分析を取り扱う。回帰分析は時間(説明変数)の変化に伴って、ある値(目的変数)が変化していく様子を関数の形で関係づけて分析することである。例えば、現在のデータをもとにして、将来のデータの予測することに適用する。このようなことは社会の中ではよくあることである。

ここでは、表計算ソフトウェアを用いて線型回帰すなわち近似直線を  $y = ax + b$  として、その係数  $a$ ,  $b$  を Excel の関数を利用して求める。

## ○実習 :

次の sheet4 はある商品の年間の月別の売上高を示した表である。これを回帰分析する。

sheet4

月	4月	5月	6月	7月	8月	9月
売上	318	412	462	482	512	618

月	10月	11月	12月	1月	2月	3月
売上	660	712	698	724	733	745

Excel では、回帰直線の傾きや  $y$  切片は次の関数で求める。

- ・傾き = SLOPE(目的変数, 説明変数)
  - ・ $y$  切片 = INTERCEPT(目的変数, 説明変数)
- グラフで表すと次のように描くことができる。

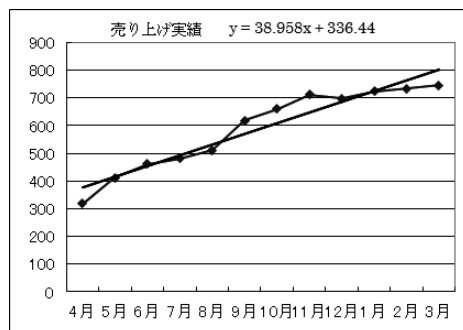


図 1 売上曲線と回帰直線

この回帰分析に関する数学的な理論や計算は授業では割愛しているが、ここでのポイントは複雑な変化を示す現象が簡単な関数でモデル化できることを学ぶことである。



図2 総合演習の様子

#### (4) 課題レポートと卒論(3年)の関係

総合演習として、各自がテーマを設定し、それに関するデータをインターネットなどから見つけ、そのデータを加工したりグラフ化したりして、その分析(コメント)を含んだレポートを作成する。

○レポートタイトル例：

- ・外国人が日本で食べておいしいと感じた日本食
- ・朝食の重要性
- ・東京都23区における平均年収と犯罪率の関係
- ・Jリーグの観客動員数と収容率
- ・何故、暑いとコーヒーが飲みたくなるのか
- ・埼玉県狭山市における人口の変化
- ・自然公園の公園数、年間利用者数の推移
- ・ネットワーク利用犯罪の現状 など

○活動のポイント(探究活動)

ここでは、データの選択も含め与えられた問題を解くのではなく、自ら興味のあるデータを探し問題を設定し、自らの観点を決めて分析させることが重要となる。第2学年のこの時期にデータ分析を行うことで第3学年の総合的な学習の時間における卒業論文を執筆するためにツールの一つとして位置付けている。

### 3. データ分析とR言語でのプログラミング

本校では、今年度から2年生の授業内容としてデータ分析に加え、R言語のプログラミングを導入することになった。この理由はRが統計解析用に開発されたフリーウェアでありプログラミング言語としてインタプリタ型のインターフェースをもつ優れたものであると判断したからである。

具体的には、いままでExcelで実施してきた2節の内容(統計分析、データの抽出など)をRを用いて取り扱うだけでなく、プログラミングの基本的な文法(条件分岐、繰り返し、ユーザー定義など)を学習させることにする。

#### 4. 今後の課題

情報科の中で育成する問題解決能力とは何か。これが例えば、数学における問題解決能力であれば、かなり明確でありかつ限定的である。なぜかというそれは、問題が数学に関わることであり、解決するための手法・方略は数学的思考力(Mathematical Thinking)に基づく力、例えば、規則性を見つける、類推する、既知の公式・定理を活用できるなどの力であるからだ。このように考えるならば、情報科で育成されるべき問題解決能力とは、ある事象を情報の観点からモデル化でき、情報技術、例えば計算論的思考力(Computational Thinking)によって解決できるものとなる。この視点からみるならばデータ分析とそれを実行する統計手法そしてプログラミング技術は高校生にとって問題解決能力を育むものとして有力なものになるであろう。

さらに、教材の選択(データや事象など)も非常に重要になってくる。なぜならばそれは生徒たちの学習の動機づけに非常に影響するからである。よい教材が生徒の学習意欲を高めることは自明である。

次期学習指導要領ではカリキュラム・マネジメントがことさら強調され、教科横断型の学習がより一層求められることになったが、このことは別段目新しいものではなく、いままでもさまざまな学校で実践されて来ているものである。

もし、それらが取り組みとして無意識なものであれば、それを学校全体として意識化する必要があるかもしれないが、重要なことは一律な学習指導要領にとらわれずに、各学校の実状や目の前にいる生徒の実態に即して教員集団がいろいろ工夫して主体的なカリキュラムを構成することにある。

#### 参考文献

- 1) 文部科学省、「次期学習指導要領に向けたこれまでの審議のまとめ(素案)のポイント参考資料」, [http://www.mext.go.jp/b\\_menu/shingi/chukyo/chukyo3/053/siryu/\\_icsFiles/afiedfile/2016/08/02/1375316\\_2\\_1.pdf](http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/053/siryu/_icsFiles/afiedfile/2016/08/02/1375316_2_1.pdf)

(アクセス日：2017年3月27日)