

Web上の都道府県別データから2変量の関係を見る

ー現行学習指導要領下の生徒に対する実践からー

愛知県立半田高等学校 菅野 栄光

1 はじめに

教科書に用意されているデータではなく、現実社会における生のデータを学習教材として活用する場合、その方向性は2つある。1つは、教師の方で、生徒が興味を持つだろうと思われるデータを提示し、それに関わるテーマ、すなわち、「…はどうなっているのでしょうか？」なる問いを与え、それを探ることを軸にして授業を展開していくものである。もう1つは、何を調べるのか、何を課題とするのか、それらを学習者自身が設定していくものである。

学習者自身による問題設定（Problem Posing）を取り入れた教育活動の効用は、様々なところで指摘がなされている。残念ながら、高等学校においては、授業進度や効率との兼ね合いからその拡散傾向を嫌い、さほど実践は行われていないのが現状である。小・中学校で培われた自ら問題を探し設定し解決するといった態度を、高等学校においても継続して育成していくことは重要であろう。今回、統計教育の分野でそれを行うとしたらどういった展開が考えられるかを探る。ただし、あくまで数学Ⅰの「データの分析」という単元の学習において、[データの相関]の内容を指導するという流れの中でのことである。全く自由に調べ発表するといった調べ学習とは一線を画する。

本モデルは、総務省統計局のWebサイトにある都道府県別データから生徒自身が2変量を選び、相関係数を求め、グループ単位で集約の上、他のグループに提示し検討していく、といったものである。共同作業を通じて、多くの変量の組み合わせについて相関係数や散布図を見ていくことでその理解を深めることを主目的とする。因果関係などの厳密な解釈にまで深入りし過ぎないことに留意したい。

2 使用するデータ

総務省統計局Webサイト「なるほどデータforきっず」の「おやくだちデータ倉庫」内の都道府県別データ
<http://www.stat.go.jp/kids/datastore/index.htm>

このデータ倉庫には都道府県別のデータが、面積、人口、一人当たり県民所得など31項目用意されている。このうちの2項目（2変量）を選んで相関を調べるとすると、単純に考えて ${}_{31}C_2=465$ 通りあることになる。しかしながら、高校生たちが興味を持つ組合せはそう多くない。予想されるのは、以下のようなものである。

[一人当たり県民所得] × [乗用車保有台数]， [一人当たり県民所得] × [道路舗装率]，
[一人当たり県民所得] × [一戸建て住宅の割合]， [一人当たり県民所得] × [大学生の数]，
[一人当たり県民所得] × [水道普及率]， [一人当たり県民所得] × [産業別事業所数]
[人口密度] × [企業の本所の分布]， [人口密度] × [建物出火件数]，
[人口] × [ごみの量]， [ごみの量] × [建物出火件数]，
[面積] × [牛乳（生乳）の生産量]， [山地面積] × [牛乳（生乳）の生産量]，
[面積] × [土地の価格]， [土地の価格] × [一戸建ての住宅の割合]

いくらかの実践の結果、[一人当たり県民所得]を軸にしてもう1つの変量を変えて比較してみるというグループが多いことが観察できた。裕福な県だと〇〇も多いはず…、といった思考が働いたものと思われる。

[総務省統計局Webサイト内の都道府県別データを利用するメリット]を以下にあげる。

- ・データをダウンロードしてすぐに表計算ソフトで利用できる。
- ・どの項目（変量）においても必ず都道府県だけの数、すなわち47個のデータがそろっているので、2変量の相関を考えることが可能。これが国別のデータになるといつも比較できるとは限らない。
- ・生徒が考える2変量の組合せがある程度予想できるので、教師側の準備もしやすい。データの出所を全く自由にするに際限なく拡散していく可能性があるが、これくらいの緩い制約であれば、生徒側の自由に自分らの興味で調べているという能動的な部分も引き出され、バランスの取れた実践が期待できる。

3 授業について

(1) 授業計画【2時間】

教科書による相関係数についての定義や、少ないデータでの手計算による算出などの指導が済んだ後に位置づける。

前半の1時間では、教師による表計算ソフトの扱いについての説明と、生徒たちがWebサイトの都道府県別データから2変量を選び相関係数を求め、グループごとに質問用紙（次時に他のグループに対して提示する）を完成させるところまでを行う。

後半の1時間では、前時に完成させた質問用紙を他のグループと交換・回答し、それらをもとにグループごとにコメントをする。1つのグループの時間は適宜調整し、全部のグループが簡単にでもコメントするだけの時間と、全員で相関係数についてのまとめを行えるだけの時間を確保する。

いずれの時間も、インターネットに接続されたコンピュータを一人一人が使用できる環境においての授業を想定している。

(2) 授業のねらい

- ・数多くの例に触れることによって、相関係数や散布図の意味を理解する。
- ・グループごとに問題を設定したり、コメントしたりすることを通じて、互いの知的なコミュニケーションを図る。
- ・自分たちの考えを分かりやすく説明したり、伝えたりする力を高める。
- ・相関関係は必ずしも因果関係を表していないことに気づかせ、安易な解釈をしない姿勢を身につけさせる。
- ・与えられたテーマに基づく分析ではなく、自由な設定において、主体的にデータの分析に向かう探究心を養う。

(3) 授業展開

【1時間目】

	主な発問など	予想される反応	留意点
導入	「相関係数について、実際の都道府県別の統計データをもとに学んでいきましょう。各項目とも47個ずつデータがあるので手計算では大変です。表計算ソフトを利用します。」 「データをダウンロードしたら、2つの項目のデータを都道府県名でソートをかけた上で合わせます。相関係数はCORRELという関数を使って求めます。散布図も表示できるのでやってみましょう。」	数学Ⅰ履修時点で高等学校情報科の方で先に表計算ソフトの操作について学んでいる可能性は低い。中学校段階までの経験の差によって反応は様々であろう。	必要最低限のことを効率よく伝える。 ソートのかけ方 関数CORRELの使い方 グラフウィザードによる散布図の表示
展開	「同じグループの人たちで協力して、色々な項目どうしを組み合わせで相関係数を求めてみて下さい。」 「正の相関がありそうなのか負の相関がありそうなのか予想してからCORRELを使うといいよ。」 「散布図をながめて、全体とはかけ離れた位置にあるものがあったら何県なのかチェックしておきましょう。」 「相関係数を求めた2項目の組合せから8組の組合せを選んで《他のグループへの質問用紙》に記入して下さい。」	「どれでもいいのですか？」 「関係なさそうなものどうしがいい。」 「これどう見ても正の相関がありそうだね。」 「これ絶対北海道。」 「やっぱり東京って極端なところにあるのかも。」 「どれがいいかなあ。」 「変わったのがいいよ。」	自分たちの興味を持った項目について調べればいいことにし、自主性に任せる。 相関係数という1つの数値による指標だけでなく、散布図を意識させることによって、外れ値などの存在を意識させる。

	「たくさん求めたグループは、意外性のある結果が出た変わった組合せばかりでなく、無難な結果というか、うん、なるほど、といえるような結果の組合せも入れてバランスよくしておいて下さいね。」	「想定範囲内だけど、これも入れておこうか。」	次時に相関関係と因果関係の違いについて議論したいので、因果がいえるような例も含ませておきたい。
まとめ	「今日は現実のデータを組み合わせて、多くの相関係数やその散布図を見てみました。だいたいイメージもふくらんだと思います。」 「次の時間には、質問用紙を違うグループと交換して、相関係数がどのくらいなのか予想するといったことを行います。」	「たくさんさんの散布図を見たおかげで、相関係数の数値によってどのような分布の違いがあるのかわかってきました。」 「よし、いっぱい当ててやるぞ！」	

【2時間目】

	主な発問など	予想される反応	留意点
導入	「前の時間に各グループとも相関係数を求めた中から8組の組合せを取り上げてもらいました。」 「それでは《他のグループへの質問用紙》を別のグループと交換して下さい。」	「これは自分たちも調べました。」	自分たちも同じ組み合わせを試した場合は、その経験をもとに答えればよいとする。
展開	「もらった質問用紙を見て、グループのみんなで意見交換した上で、正の相関があるか、相関がないか、負の相関があるか、大きく3段階で予想して下さい。」 「自分たちが問題を作ったときの経験なども生かして判断してみましょう。」 「8組の相関係数予想の中で一番自信があるものを選んで、その理由も書いておいて下さい。」 「それでは問題を作ったグループに用紙を戻してチェックしてもらいましょう。」 「チェックが終わったら、各グループにコメントしてもらいます。次のようなことについて簡単に説明して下さい。 ①何問正解だったか。 ②回答グループが自信ありとした問題についてはどうだったか。自分たち（出題者側）の解釈と同じか。 ③逆に、出題者側として意外な結果だと思われた問題について。」	「予想がグループの中でバラバラのときはどうしたらよいですか。」 「[人口] × [ごみの量] はきっと正の相関があるよね。ごみを出すのは人間だし…。」 「[人口] × [ごみの量] については回答グループの解釈の通り、自分たちも人が多い県ほどたくさんのごみを出すのだと考えました。」 「[一人当たり県民所得] × [乗用車保有台数] はもっと強い正の相関があるかと思ったら、それほどではありませんでした。その県の交通事情などきっと他にも要因があるのかもしれない。」	多様な意見が出た場合は、グループ内でそれぞれの考えを述べ合わせて、説得力のあるものを採用させる。 相関関係があるもののうち、明らかに因果まで言えそうなものがあれば選ばせる。 時間配分に注意する。後の方のグループになると重なる例などが出てくるので、状況によっては説明を簡略化させる。 2変量以外の要因が存在する可能性に気づかせる。

まとめ	「各グループがあげてくれた2項目の組合せの例から、相関関係を単純に原因や結果に結びつけてはいけなかったことがわかりました。」 「他にも、例えば、昔から『貧乏子たくさん』などと言われていますが、実際、県民所得と出生率の相関を調べると負の相関があります。みなさんならこれをどう説明しますか。」 「現実問題としては様々な要因が複雑に影響を及ぼしており、単純に2変量の関係だけでは説明がつかないことが多いのですね。」 「しかし、複雑なデータ分析を行うための足がかりとして、2変量の関係を数値として捉える相関係数やその散らばり具合を目で確認できる散布図といったものが有用であることは理解できたのではないのでしょうか。」	県民所得と出生率の関係 「都会の人は一人暮らしが多い。」 「女性が働く子育てが大変。」 「自分らで選んだ項目についてたくさん試すことができたので面白かったです。」 「自分のグループと他のグループの多くの相関係数と散布図を見たので、だいたいこれくらいの数値ならこんな散らばり具合といったイメージが湧くようになりました。」	時間がなければ、この例は取り上げず、グループでの発表例をもってまとめに代える。 余裕があれば、多変量解析などの話をしてもよい。 コンピュータの助けを借りて多くのケースを試すことができた。このことが1つの目標であったという振り返りをさせる。
-----	--	---	--

4 おわりに

高校生に限ったことではないが、主導権を生徒に渡した時の反応はすこぶるよい。都道府県別のデータだけ用意して、「これだけの項目についてデータがあるのだけど、自分たちの興味のある項目どうしを選んで組み合わせで関係を調べてみましょう」、そう投げかけるだけでかなりの生徒は熱心に取り組んでくれる。生徒の内から湧いてくる興味を原動力に何とかコンパクトに授業展開ができないものかを探った。時間を惜しみなく使えば、より充実した中身の濃い調べ学習が可能になるかもしれないが、今回は数学という教科の一単元の学習という位置づけで考えたモデルである。

グループによるデータの解釈を巡った話し合いや意見交換は、次期学習指導要領でも目標とされている「自分の考えを分かりやすく説明したり、互いに自分の考えを表現し伝え合ったりすることなどの指導を充実する」ことにつながるであろう。

また、「データのばらつきや偏りなどデータ間の関係について、適宜コンピュータなどを用いてデータを整理し、数学的に考察し説明ができるようにする」との目標を受けて、コンピュータ室で生徒が表計算ソフトを活用しながら探求していく授業モデルとした。情報科との連携を強化したいところであるが、設定される学年の違いなどからそう簡単ではないかもしれない。数学の時間だけで必要最低限のことを効率よく教え、コンピュータを道具として活用できる状態にもっていきたい。「多くのデータを扱う場合には、コンピュータなどを積極的に活用するようにする」とある通り、できる限り上手に活用したいものである。

本授業モデルにおいては、表面的には、ただ相関係数をたくさん計算させ、それをゲーム感覚で当てっこさせただけに映るかもしれない。しかし、数学B（新課程）の学習前の段階でΣ記号も使えず相関係数の概念を導入していることもあり、多くの例、多くの散布図に触れさせ、相関係数の数値によって2変量にどのような関係が成り立っているのを具体例で実感させることは意味のあることだと思われる。付け加えて、現実のデータを扱うことにより、相関関係が必ずしも因果関係につながらないという注意を喚起させることも大事なことであろう。世の中には、強い正の相関関係を示す散布図を見せ、さもその2つの事柄に強い因果関係があるかのように錯覚させ、自分たちにとって都合の良い主張を行う、といったことがなされている場面も少なくない。そのようなことに対して、科学的、批判的に見るができるような態度を育成することは重要であろう。

《出題グループ 班, 回答グループ 班》

比較する2項目	強い, やや強い 正の相関 $r > 0.7$	弱い相関, あまり相関なし $-0.7 \leq r \leq 0.7$	強い, やや強い 負の相関 $r < -0.7$
① [] × []			
② [] × []			
③ [] × []			
④ [] × []			
⑤ [] × []			
⑥ [] × []			
⑦ [] × []			
⑧ [] × []			

↑ 出題グループが組み合わせた2項目について記入して下さい。
(相関係数については自分たちで控えておくこと)

↑ 回答グループが予想して○をつけて下さい。

回答者グループは①～⑧の予想の中で一番自信があるものをあげ、その理由（どのように分析したか）も書いて下さい。