

「資料の活用」分野における標本調査の教材開発と実践

竹内洋平¹，愛木豊彦²

現代社会において，資料の傾向を読み取って判断する力が求められている。これを受けて学習指導要領中学校数学科において「資料の活用」が領域として追加された。それに伴い，多くの場面で用いられている標本調査が3年生の学習内容となった。本稿では，標本調査の方法を紹介し，さらに，標本調査の精度について考察する授業案を紹介する。そして，生徒の様子やアンケートをもとに授業の実践結果について報告する。

<キーワード> 標本調査，ヒストグラム，資料の散らばり，乱数サイ

1. はじめに

平成20年度に中学校学習指導要領が改訂され「資料の活用」領域が新しく加わった。それは，中学校学習指導要領解説 ([1]) にもあるように，情報社会において，確定的な答えを導くことが困難な事柄についても，目的に応じて資料を収集して処理し，その傾向を読み取って判断することが求められているからである。それに伴い，標本調査も3年生の学習内容として復活した。

過去の教科書 ([2]) で，標本調査がどのように指導されているかを調べたところ，例として示されている母集団には，次の2つの特徴があることに気付いた。1つ目は，1時間で全数調査が可能なくらい母集団の要素の数が少ないことである。2つ目は，少ない標本でも全体の傾向が分かるような母集団であるということである（このことについては，次節で詳しく述べる）。これらの特徴には，標本調査の有用性を伝えるために，標本調査によって得られる値が実際の値に近いことを，生徒が確かめられるようにという意図があるものと考えられる。

しかし，実際に標本調査を行おうとした場合，母集団の要素の数は多かったり，少ない

標本では標本平均が母集団の平均とかけ離れてしまったりすることがある。従って，生徒が実際に標本調査を行った場合のことを考え，上の2つの特徴を持っていない母集団を題材とした授業を開発することにした。さらに，標本調査の結果を分析する際に，1年生で学習するヒストグラムや資料の散らばりに関する内容を活用できることもわかったため，この点を授業展開に取り入れることにした。

教材を開発するにあたり，次の2つの条件を満たす母集団の組を探した。

(ア)1時間では全数調査が困難な程度に母集団の要素の数が多きこと。

(イ)ある標本調査における標本平均が，母集団の平均に近くなる確率が高いものと，低いものという2つの母集団の組み合わせで，かつ，それらの内容が似ているもの。

そして選んだ母集団が「岐阜市1年間の1日の最高気温」と「岐阜市1年間の1日の降水量」である ([3])。これら2つの母集団の組が条件(イ)を満たしていることについては，次節で述べる。

2. 教材について

本論文で扱う教材は「実際の岐阜市1年間

¹岐阜大学大学院教育学研究科

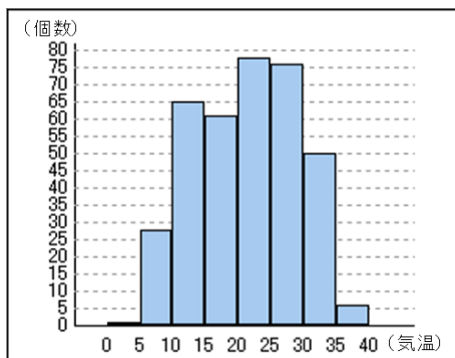
²岐阜大学教育学部

の1日の最高気温のデータからその平均値を求めること」(調査1とする)、「実際の岐阜市1年間の1日の降水量のデータからその平均値を求めること」(調査2とする)である。

以下、2.1節と2.2節で、2つの母集団の組が条件(イ)を満たしていることを示す。そして、授業用に作成した資料Aと資料Bをそれぞれ2.3節と2.4節で紹介する。あわせて、これらに対し、標本調査をした結果も示す。最後に2.5節で乱数サイを用いた無作為抽出の方法を述べる。このように乱数サイを用いた実験を行い、資料の散らばりを考察する授業実践に松野・愛木([4])がある。

2.1 調査1について

2009年における岐阜市1年間の1日の最高気温をヒストグラムで表したのが、グラフ1である。



グラフ1

この母集団の要素の数は365個、平均は21.4、標準偏差は7.78である。この分布を正規分布とみなすことにする。

母集団の平均(母平均)を μ 、標準偏差を σ 、標本の数を n 、標本平均を \bar{x} 、 $0 < \alpha < 1$ とすると、

$$P\left(-k_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu - \bar{x} \leq k_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha \quad (1)$$

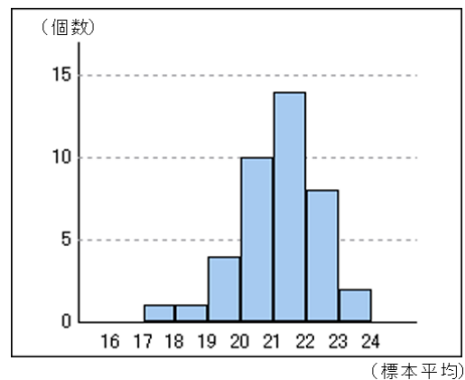
となる k_α を正規分布表から求めることができる([5])。ただし、 $P\left(-k_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu - \bar{x} \leq k_\alpha \frac{\sigma}{\sqrt{n}}\right)$ は、 \bar{x} が $-k_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu - \bar{x} \leq k_\alpha \frac{\sigma}{\sqrt{n}}$ となる確率を表している。ここで、

$$P(-2 \leq \mu - \bar{x} \leq 2) = 0.75 \quad (2)$$

となるような標本の個数を求める。

$\alpha = 0.25$ のとき、 $k_\alpha = 1.15$ である。よって、(2)となるためには、 $2 \leq k_\alpha \frac{\sigma}{\sqrt{n}}$ が成り立てばよい。これを解くと、 $n \geq 20$ となる。

次に、実際に20個の標本を無作為に抽出し、その標本平均を求める。これを40回行った結果をグラフにしたものがグラフ2である。実験結果から、標本数20で標本調査を行うと、グラフ2のように、多くの場合、標本平均が20~22の値をとることがわかった。

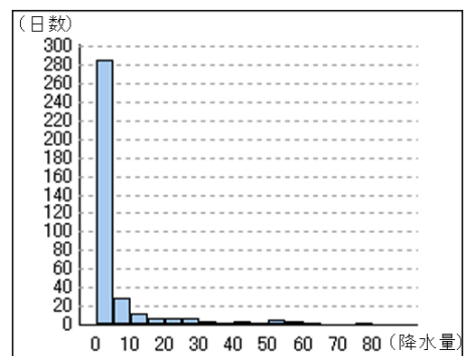


グラフ2

以上の結果から、20個の標本でおおよその母平均が推定できると判断した。

2.2. 調査2について

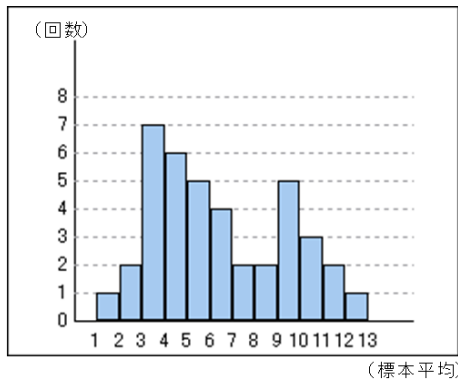
2009年における岐阜市1年間の1日の降水量をヒストグラムで表したのが、グラフ3である。



グラフ3

調査2における母集団の要素の数は365個、平均は5.2、標準偏差は12.49である。この場

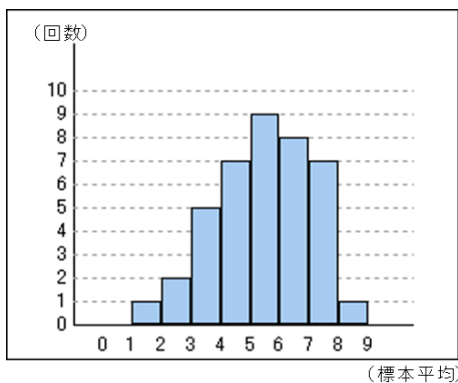
合，調査1と比べて，標準偏差を平均で割った値が大きいので，少ない標本数では標本平均が母平均に近くなるとは限らない。このことを確かめるため，調査1と同じように，20個の標本を無作為に抽出し，その標本平均を求める。そしてそれを40回行い，得られた標本平均をグラフにしたのが，グラフ4である。



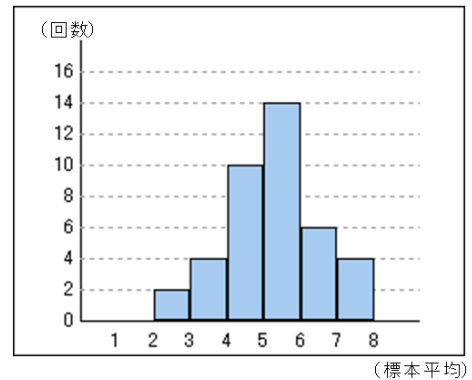
グラフ4

このグラフから分かるように，標本平均の取り得る値の範囲が広く，標本平均は5.2の近くに集まっていない。よって，調査2では，20個の標本で母平均を推定するのは難しいと考えられる。

次に，調査2において標本数を50，100として標本調査を行った。グラフ2，グラフ4を作成したときと同じように標本調査を40回行った結果をまとめたものが，それぞれグラフ5とグラフ6である。



グラフ5

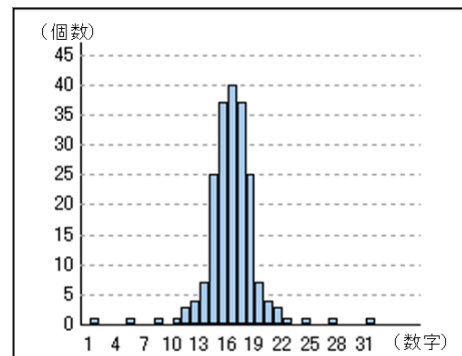


グラフ6

これらの結果から，調査2においては，標本数が20個では，母平均の推定は難しいが，標本数を増やすことで，標本平均が母平均に近づくことがわかる。

2.3. 資料Aについて

調査1の母集団と近い分布をする資料Aを作成した。これをヒストグラムに表したのが，グラフ7である。



グラフ7

この母集団の要素の数は200個，平均は16，標準偏差は2.80である。この分布を正規分布とみなすことにする。

(1)において，

$$P(-2 \leq \mu - \bar{x} \leq 2) = 0.9 \quad (3)$$

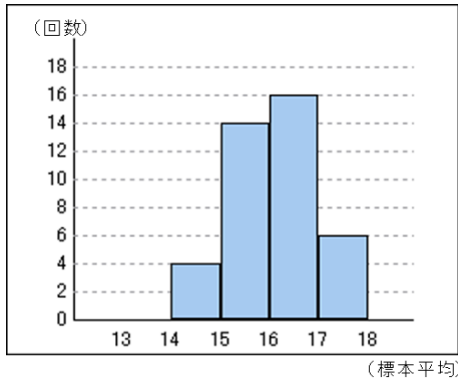
となるような標本の個数を求める。

$\alpha = 0.1$ のとき， $k_\alpha = 1.645$ である。よって，(3)となるためには， $2 \leq k_\alpha \frac{\sigma}{\sqrt{n}}$ が成り立てばよい。これを解くと， $n \geq 5.3$ となる。

よって，標本平均が母平均に確実に近い値

をとるように、標本数を10とする。

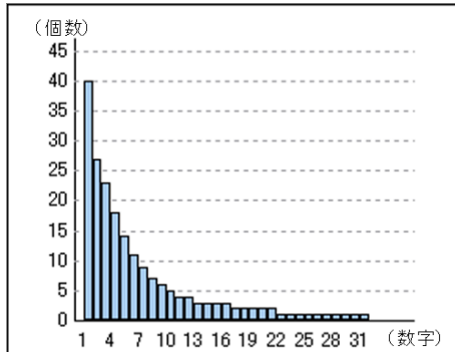
10個の標本を無作為に抽出し、その標本平均を求める。そしてそれを40回行ったデータをグラフにしたのがグラフ8である。この標本調査を数多く行ったが、どの場合でも標本平均はグラフ8のように母平均に近い値をとった。



グラフ8

2.4. 資料Bについて

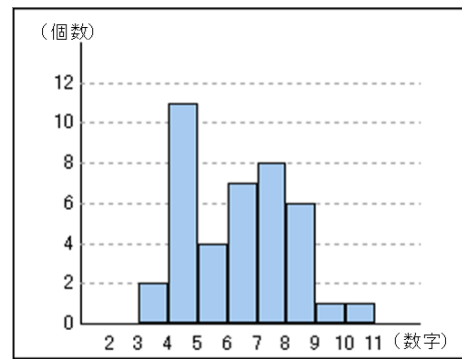
調査2の母集団と近い分布をする資料Bを作成した。これをヒストグラムに表したのが、グラフ9である。



グラフ9

この母集団の要素の数は200個、平均は6.5、標準偏差は6.69である。この場合、資料Aと比べて、標準偏差を平均で割った値が大きいので、少ない標本数では、標本平均が母平均に近い値をとりにくい。

このことを確かめるために、標本数を10として標本平均を求める。そして資料Aの調査と同じように、これを40回行ったデータをグラフにしたのがグラフ10である。



グラフ10

多くの実験から、資料Bに対しては、グラフ10のように、母平均6.5の近くに標本平均が集まらないことがわかった。

2.5. 無作為抽出の方法

調査1、調査2、資料Aの平均を求めること、資料Bの平均を求めることの4つの標本調査において、正20面体の乱数サイを用いて無作為抽出を行うことにした。

調査1、調査2の母集団の要素の数はどちらも365個であるため、乱数サイを3回振って、百の位、十の位、一の位を決める。必要な百の位は0, 1, 2, 3の4種類なので、乱数サイの出た目が1~5のときは百の位を0, 出た目が6~10のときは1, 出た目が11~15のときは2, 16~20のときは3とすることにした。十の位、一の位を決める場合は、乱数サイの出た目が1~9のときはそのまま用い、出た目が10~20のときは、その数の1の位を用いることにした。ただし、000や366から399までの数は使わない。

資料A、資料Bの母集団の要素の数はどちらも200個である。乱数サイは正20面体であるため、乱数サイを2回振って、上2桁と一の位を決めることにした。上2桁を決める場合、乱数サイの出た目をそのまま百の位と十の位の数とする。一の位を決める場合、乱数サイの出た目が1~9のときはそのまま用い、出た目が10~20のときは、その数の1の位を用いることにした。ただし、000は使わず、201から209までの数は、1~9として用いる

ことにした。

3. 授業実践について

3.1. 授業の概要

場 所 : 岐阜大学教育学部附属中学校

日 程 : 平成 22 年 3 月 1 日, 3 日, 4 日の
3 日間

対象生徒 : 3 年 1 組 37 名

教材名 : 「標本調査」

各時間の内容とねらい

< 第 1 時 >

・ 内容

岐阜市 1 年間の 1 日の最高気温の資料において、標本調査を行う。

・ ねらい

標本調査において、標本の意味と標本を無作為に取り出す必要性がわかり、実際に標本調査をする活動を通して、標本調査の有用性を感じることができる。

< 第 2 時 >

・ 内容

岐阜市 1 年間の 1 日の降水量の資料において、標本調査を行い、最高気温と降水量の標本調査の結果をまとめて、それぞれの資料の特徴を考える。

・ ねらい

資料のデータの数も標本数も同じなのに、少ない標本数でおおよその平均がわかるものと、わからないものがあることに気づき、それぞれの資料の特徴を読み取ることができる。

< 第 3 時 >

・ 内容

資料 A と資料 B の標本調査を行い、なぜ資料 A では少ない標本でも標本平均が母平均に近く、資料 B ではそうでないのかについて考える。

・ ねらい

平均のまわりに集まっている分布の母集団に対する標本調査は、平均のまわりに集まっていない母集団に対する標本調査より、少な

い標本数でおおよその平均がわかることを実感し、それは平均に近い値が標本となる確率が高いからであることを理解する。

3.2. 授業の流れと生徒の活動の様子

< 第 1 時 >

標本調査は生徒たちにとって初めての経験なので、標本という言葉の意味や無作為に取り出すことの必要性を授業者が説明した。そして問題を提示し、乱数サイを用いた標本調査の方法を説明した後、実際に調査 1 の標本調査を行った。生徒たちは、順調に乱数サイを使って 3 ケタの数を決めることができた。そして各自で標本平均を求め、それをクラスで集め、1 つのグラフにまとめた。その結果が写真 1 である。

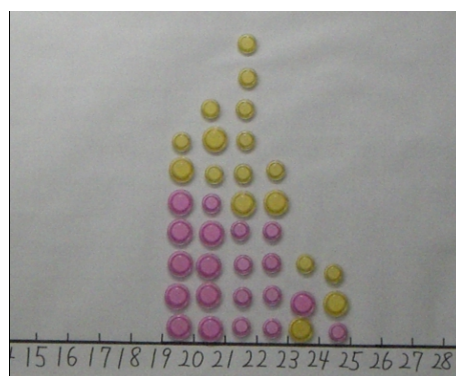


写真 1

このグラフから、生徒たちは平均を 20 ~ 22 と推定していた。

< 第 2 時 >

初めに調査 2 の標本調査を行った。調査の方法は調査 1 と同じであり、生徒たちは第 1 時よりも素早く調査を行うことができていた。そして各自で標本平均を求め、それをクラスで集め、1 つのグラフにまとめた。その結果が写真 2 である。

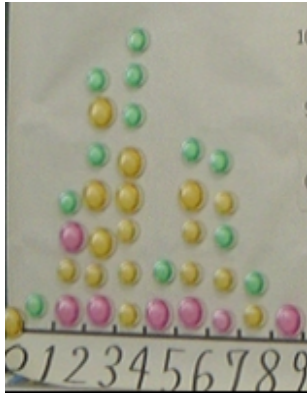
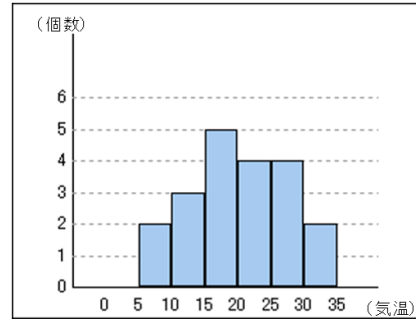
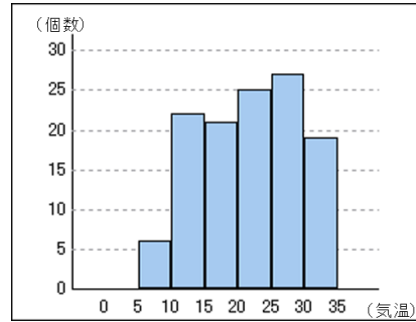


写真 2

このグラフから、調査 1 とは異なり、平均を推定することは難しいのではないかという意見が出た。ここで、授業者が調査 1 と調査 2 において、標本数を 50, 100 とした場合の標本調査の結果を提示した(写真 3)。この結果は授業者が事前に標本調査を 40 人分行ったものである。



グラフ 11



グラフ 12

同じように、調査 2 に対しても、生徒ごと(グラフ 13)と班ごと(グラフ 14)にグラフをかく。

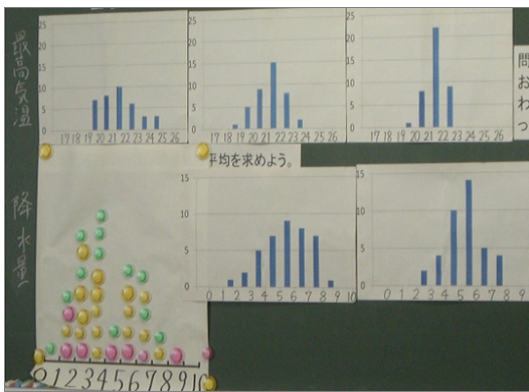
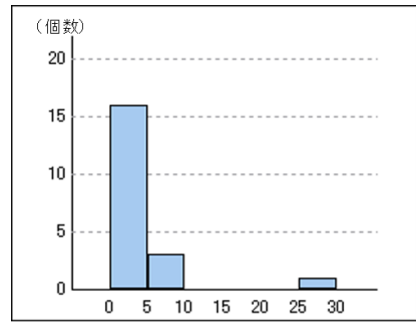


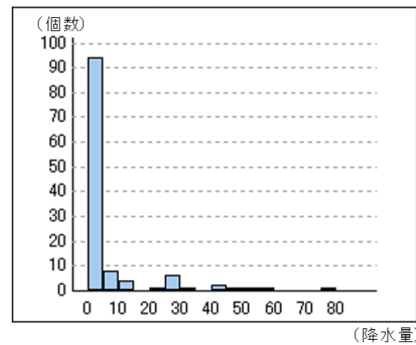
写真 3

これらのグラフから、標本数を増やすと調査の精度は上がることがわかる。また、最高気温の標本調査では標本数が 20 でも平均が推定でき、降水量ではできないということにも気付いた。そこで、その理由を考えることにした。

まず、既に行った調査 1 に対する標本調査で、乱数サイを振って得られた最高気温の値(標本平均ではない)を、各自でまとめる(グラフ 11)。さらにそれを班ごとにまとめる(グラフ 12)。



グラフ 13



グラフ 14

このようにして得られたグラフは、標本平

均の場合とは異なり，資料を表すもとのグラフ(グラフ1，グラフ3)と同じような形となる。この活動を通し，生徒たちはそれぞれの資料のグラフの形と標本調査の結果との関係について考察し始めた。そして，グラフの特徴として，「最高気温のグラフは，まん中に均等なところがある。山のような形をしている。」，「降水量のグラフは，0が多い。滑り台のような形をしている。」という意見が出た。しかし，平均に着目して資料の特徴を考える生徒はいなかったため，授業者から「平均に着目してわかることはありますか。」と問いかけた。すると，生徒から「最高気温の方は平均の所に集中している。」，「降水量の方は平均以外のところに多くある。」という意見が出た。そこで「平均のそばにたくさんデータがあると，少ない標本数でおおよその平均が推定できそうである。」としてまとめた。

<第3時>

第2時の最後で予想したことがいつでも言えるかどうかを確かめるために，同じような資料を使って，もう一度実験することにした。そのための資料として，資料Aと資料Bを生徒に提示した。そして乱数サイを使った標本調査の方法を説明して，資料Aと資料Bに対する標本調査を行った。標本数が10個で，乱数サイも2回振れば決まるので，生徒たちは前時よりも短時間で標本調査を終えることができた。そして各自標本平均を出し，それをクラスで1つのグラフにまとめた。その結果が写真4である。

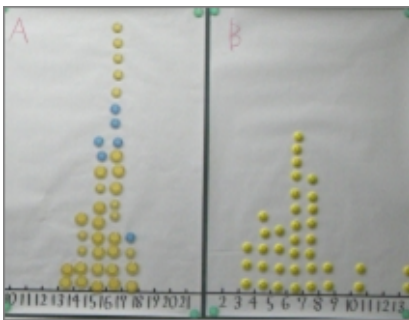


写真4

この結果から，生徒たちはやはり資料Aの方が正確に推定できていると感じていた。そして，その理由について考察したところ，ほとんどの生徒が，平均に近いデータが標本となる確率に着目し，考えを進めていた。その中には「資料Aの平均 ± 3 のデータが標本となる確率は $\frac{146}{200}$ ，資料Bの平均 ± 3 のデータが標本となる確率は $\frac{88}{200}$ である。」，「資料Aの平均 ± 1 のデータが標本となる確率は $\frac{113}{200}$ ，資料Bの平均 ± 1 のデータが標本となる確率は $\frac{33}{200}$ である。」などと，実際に確率を求める生徒もいた。よって「平均に近い値が多い資料は，平均に近い値が標本となる確率が高いので，おおよその平均が少ない標本数でわかる。」としてまとめた。

4. アンケート結果と考察

本授業のねらいは次の2つである。

- (a) 標本調査の有用性を実感し，少ない標本数で全体の平均が推定できる場合とできない場合があることを実感する。
- (b) 少ない標本で標本調査をするためには，資料の平均の周りに値が多いかどうかを確認することが大切であることを理解する。

これらのねらいが達成できたかどうか確認するために，次の2つの質問をし，解答を得た。

(1) 「標本調査の良さは何ですか?」

- ・全体を調査する時間と手間を省くことができること。
- ・だいたいどのくらいの数になるか簡単にわかること。
- ・少ない標本で，全体のおおよその平均がわかること。

ほとんどの生徒から上記のような意見を得ることができた。このことから，ねらい(a)は達成できたのではないかと考える。

(2) 「少ない標本で標本調査をするために，大切なことは何ですか?」

- ・平均に近い値が多いことを確認すること。

- ・予備調査で、おおよその資料のグラフを確認すること。
- ・あてずっぽうではなく、事前に標本調査が適しているかどうかを調べること。
- ・資料の情報をある程度知っておくこと。
ほとんどの生徒から上記のような意見を
得ることができた。このことから、ねらい(b)は
達成できたのではないかと考える。
- 最後に、生徒の授業の感想を一部紹介する。
- ・標本調査自体を知らなかったけど、これだけ
で平均値を出せると思うと便利だと感じた。
- ・標本調査をやってみて、ただ平均を出して
終わるのではなく、いろいろ考えることが
できたのでよかったです。
- ・乱数サイを用いて標本調査して、すごく楽
しかったです。
- ・ぜひ他のものでも標本調査をして調べてみ
たいと思いました。

5. 今後の課題

第2時において、時間の関係上、標本の数を増やすと調査の精度が上がっていくことを教師が教える形となってしまった。これは標本調査における大切な数学的考え方なので、生徒自身から引き出せるような授業展開にしていきたい。

また、第3時において、資料A、資料Bの標本調査を行った結果が、明確な差があるとはいえないものとなってしまい、少し生徒が困惑する場面があった。この2つの資料をさらに改良し、標本調査の結果に明確な差を出すことができるものにしていきたい。

今回実践を行ったのは、1年生で「資料の活用」を学習していない生徒たちである。今後、「資料の活用」を学習した生徒を対象に、この内容を指導する場合、ヒストグラムや代表値の使い方をどのようにすべきか、再度検討する必要がある。

引用文献

- [1] 文部科学省, 2008, 中学校学習指導要領解説 数学編 .
- [2] 赤攝也・井上義夫 他 19名, 平成8年, 新版中学校数学3, 大日本図書株式会社 .
- [3] 気象庁ホームページ .
<http://www.jma.go.jp/jma/index.html>
- [4] 松野利香・愛木豊彦, 2009, 「資料の活用」における授業の提案と実践, 岐阜数学教育研究, 第8号, pp. 22-34.
- [5] 高松俊朗, 1984, 数理統計学入門, 学術図書出版 .