

箱ひげ図の指導に関する一考察 —指導に関して必要な事項の抽出と指導事例に対する検討—

愛知教育大学 青山和裕

1. はじめに

新教育課程における統計教育については、中学校数学の第2学年において箱ひげ図の指導が位置付けられた。高等学校の数学I「データの分析」においてこれまで指導されているものの、同様の指導を中学校において展開すればいいというわけではない。統計的な問題解決が重視されていることや統計指導の充実が求められていることを考慮すると、箱ひげ図の単なる読み書きではなく、問題解決の利用やヒストグラムと併用するなどこれまでとは異なる指導を視野に入れる必要がある。とはいえ、箱ひげ図の指導において扱うべき事項についても明確に定められておらず、またどの程度の指導時数が設定されるかも定かではない状況である。本稿では中学校第2学年に位置づけられる箱ひげ図の指導時数について、高等学校での指導やデータの活用領域の内容構成の面から検討し、その上で箱ひげ図の指導において必要な事項について提案する。最後に知立市立知立中学校での実践事例について必要事項の観点からどの程度達成できていたかを検討する。

2. 箱ひげ図の指導時数について

(1) 現行教育課程での高等学校での指導時数について

高等学校での「データの分析」の指導時数については、9～19時間と教科書によって様々であることが報告されている(裕元・梅田・富田, 2012)。このうち箱ひげ図に充てられる指導時数についてみると、例えば啓林館の「数学I」(高橋他, 2011)では、箱ひげ図は「データの傾向のとりえ方」というセクションに位置付けられ、セクション全体で指導時数5時間となっている。このセクションには、代表値や分散、標準偏差の指導も入っていることを考慮すると、実質的に箱ひげ図に充てる時数は2時間程度だろうと思われる。2時間程度で指導されてきたため、高等学校では基本的な読み書きを指導するにとどまっていたというのが現状である。

(2) 「データの活用」領域の内容構成の変化

中学校での時数はどの程度見込めるのか。実際の指導時数については学習指導要領で規定されるのではなく、各教科書会社の裁量によるため、教科書を待たなくてははっきりしたことはわからないが、内容構成的にどの程度を見込むことができるのかを確認するために、「データの活用」領域の内容構成の変化についてここで整理しておく。

これまで第1学年で指導されていた代表値(平均値, 中央値, 最頻値)や階級という用語が小学校第6学年に移動した。また、誤差, 近似値が第3学年に移動した。代わりに、第2学年で指導されていた確率のうち、「多数の観察や多数回の試行によって得られる確率」いわゆる統計的確率が第1学年に移動した。統計的確率が移動したことでできた枠に箱ひげ図が高等学校から移動してきたという図式となる(文部科学省, 2017)。

第2学年では、他領域も含めて、特に削減された内容や移動した内容はないため、表面的には統計的確率に充てられていた指導時数を箱ひげ図に充てるとするのがまず考えられる。

さて、統計的確率の指導に充てられた時数については、例えば、啓林館教科書(岡本他, 2015)では3時間である。このことから、最低でも3時間の指導時数を見込むことはできそうであるが、3時間では高等学校における指導と大差はなく、読み書きを指導するので精一杯だろう。3時間を最低ラインとしつつ、あとは各教科書でどのように設定されるかという点が問題である。

時数の過不足を論じるためには、指導すべき内容を定める必要もある。基本的な読み書きだけでは不足なのは明らかだとしても、箱ひげ図に関して他に指導すべき項目として何が考えられるのかを次章で整理する。

3. 箱ひげ図の指導において必要な事項について

(1) 四分位数・四分位範囲 (5数要約)

箱ひげ図を学習するために必要な知識として、四分位数がある。量的データを値の小さい順に並べ、同じ数になるように4つのグループに分けた際の境目に位置する値であり、小さい方から順に「第1四分位数」、「第2四分位数(中央値)」、「第3四分位数」と呼ぶ。第1四分位数から第3四分位数の間にデータ全体のおおよそ半分が位置することとなり、ばらつきを中心をとらえる際の指標として用いることができる。第1四分位数から第3四分位数までの差を「四分位範囲」といい、分散や標準偏差を学習しない中学生にとっては、ばらつきを数値化した指標として用いることができる。

また、四分位数に最小値、最大値を加えた5つの数字でデータのばらつき具合を把握することもでき、これら5つの数字をまとめて「5数要約」と呼ぶ。5数要約という用語は指導要領には記載はされていないが、これら5つの数字は指導において必然的に扱うため、5数要約という用語も教科書に記載されるかもしれない。

(2) 箱ひげ図の定義・書き方・読み方

5数要約を図に表したのが箱ひげ図である。第1四分位数から第3四分位数までを箱で表し、中央値

D データの活用	
← 現行「D資料の活用」の名称を変更	
<p>データの分布の傾向</p> <ul style="list-style-type: none"> ・ヒストグラムや相対度数の必要性と意味 <p>多数の観察や多数回の試行によって得られる確率</p> <ul style="list-style-type: none"> ・多数の観察や多数回の試行によって得られる確率の必要性と意味 (←中2) <p>(用語に累積度数を追加)</p> <p>(用語から、代表値、(平均値、中央値、最頻値)、階級を削除)</p> <p>(→小6)</p> <p>(内容の取扱いから、誤差、近似値、$a \times 10^n$の形の表現を削除)</p> <p>(→中3)</p>	
<p>データの分布の比較</p> <ul style="list-style-type: none"> ・四分位範囲や箱ひげ図の必要性と意味 (追加) ・箱ひげ図で表すこと (追加) <p>場合の数を基にして得られる確率</p> <ul style="list-style-type: none"> ・確率の必要性と意味 ・確率を求めること <p>(「確率の必要性と意味」を一部移行 (→中1))</p>	

図1: 「データの活用」領域の内容構成の変化(第1, 2学年分)

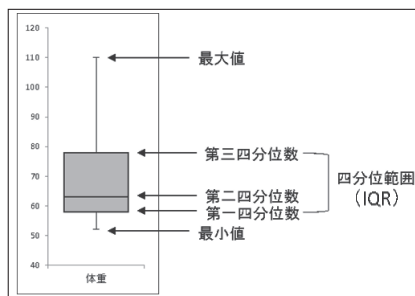


図2: 箱ひげ図と5数要約の対応 (画像はBellCurve『統計WEB』より)

で箱の中に線を引く。第1四分位数から最小値までと第3四分位数から最大値までのそれぞれのひげを伸ばすと箱ひげ図となる。

ひげの始まりから終わりまでの長さが「範囲」となり、箱の長さが四分位範囲となる。上下のひげと箱の上側・下側の4つの区画にそれぞれ4分の1のデータが存在することとなり、分布の様子を視覚的にとらえることができる。

このような箱ひげ図の基本的な作りや読み方について指導するのは当然であるが、箱ひげ図からデータの分布を読み取る際には、様々な誤認が起りやすいことが報告されている。例えば、小口(2010)は図3のような調査問題を用いて大学生を対象に箱ひげ図からの読みとりに関する調査を行っている。

A, B 2つのクラスの得点をまとめた箱ひげ図を用いて、問9では30点以下の学生の比率が大きいのはどちらかを問うている。30点はどちらのクラスでもひげの途中に位置し、ひげの中にどのようにデータが点在するかは把握不能であるため、どちらのクラスに30点以下の学生が多く存在するかは読み取り不可能であるはずなのだが、正解できたのは調査において15.8%のみで、56.1%の学生が「クラスB」を選択してしまっている。30点以下のひげに注目するとクラスBの方が長いというのが誤ってしまった原因として考えられる。

また問10では、80点以上の比率について問うている。どちらのクラスでも80点は中央値に位置するため、80点以上の学生はおおよそ半分いるというのが妥当な読み取りである。この問題で正解できたのは8.8%で、73.7%の学生が「クラスA」を選択してしまっている。問9と同様に箱やひげが長い方が、その部分に該当するデータが多く存在するように誤認してしまうということである。

(3) 箱ひげ図の利点

箱ひげ図の利点については、定義と表裏一体だが5数要約を視覚的に容易に把握できることまず挙げられる。他には、あとで触れる榊原実践において用いられたように、7クラス分のデータを比較する際など、多群比較をする場合に箱ひげ図は便利である。箱ひげ図を導入する際の文脈として、多群比較をする場面を設定し、ヒストグラムや度数分布多角形では分析しにくいという流れから箱ひげ図を導入する流れが主流となると思われる。

ただし、多群比較での利点ばかりを強調すると、2グループまでならヒストグラムを用いればよく、3グループ以上の場合に箱ひげ図を用いるなどという分析の本質とはかけ離れる利用方法が確立していく可能性があるため注意が必要である。分析においては、データを多面的にとらえ

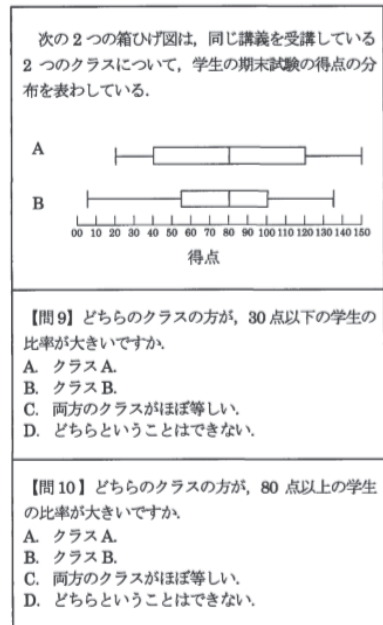


図3：小口(2010)の調査問題

て総合的に判断することが求められるため、この場合はヒストグラムを用いればよく、こっちは場合は箱ひげ図を用いるのが妥当などという分類方法は存在しない。

さらに別の利点としては、データの標準的な値がどの範囲なのかをとらえることができるという点も挙げられる。通常、平均と標準偏差が用いられることが多いが、正規分布の場合では、平均 $\pm 2 \times$ 標準偏差の範囲にデータの95%が存在することが知られている。子どもの成長に対して年齢に応じた標準的な身長や体重、成人に対してもBMIの標準的な範囲などはよく示されることが多く、ばらつきのある値に対して標準的な範囲でとらえようとするのは世の中で共有された視点である。このような視点を中学校での統計指導に持ち込もうと考えると、分散・標準偏差を指導しないため扱えないこととなる。代表値は代表的な値を中心の1点でとらえようとする視点であるため、ばらつきは考慮されない。そのデータの標準的な範囲はどこなのかという視点で分析する際に、中学校で用いることができるのは四分位範囲だけである。

例えば中学生の1日の勉強時間や睡眠時間、お小遣いでもいい。データを集めればばらつきがあるのは当然であるが、中学生の標準的な値はどのあたりだろうと考えた際に、四分位範囲に注目してとらえるというのは妥当な分析である。このような使い方の場合には、多群比較を必要とせず、1グループや2グループのデータでも箱ひげ図を活用することができる。もちろん、ヒストグラムや度数分布表を用いて、中心を含むいくつかの階級でとらえたり、累積相対度数なども併用すればよい。

(4) 箱ひげ図の欠点

一方で欠点としては、小口(2010)の調査でも明らかになったように、箱やひげの中でデータがどのように点在しているか分からないことが挙げられる。箱やひげの中で均等に散らばっているのか、端の方に偏っているのかはわからないのである。また、分布の形状が多峰性をなしていたり、二極化していたりするかもわからず、その箱ひげ図がいくつのデータで作られたものなのかも、添え書きや説明がなければわからないのである。例えば、あるクラスのデータとそのクラスを除いた全校生徒全体のデータを箱ひげ図に表したとすると、2つの箱ひげ図のもととなるデータ数は全然違っているはずだが、箱ひげ図の見かけからはそのことを読み取ることはできない。

箱ひげ図は作るのも容易で視覚的にとらえやすいという利便性がある反面、5数要約によって元のデータの持つ情報量が極端に圧縮されてしまっているために、見えなくなることも多いという欠点がある。

(5) 箱ひげ図とヒストグラムの比較

これまでまとめた箱ひげ図の利点・欠点は、分析する場面でヒストグラムも併用することで生徒も気づきやすくなる。いろいろな手法を用いて多面的にデータを分析することも求められているため、ヒストグラムとの併用は望ましい指導であることは間違いないが、最初に触れた指導時数との兼ね合いで、実際にできるかどうかは難しいところでもある。

(6) はずれ値の見極めと扱い

量的データでは、他のデータと比べて著しくかけ離れた値(はずれ値)が存在することがある。

はずれ値は計測ミスや回答ミスで生じる場合もあり、それがわかるのであれば分析の際に除外するのが普通である。ミスではなく実際に存在する場合もあり、その際にはむしろ特異な例として分析の対象として掘り下げるようなこともある。はずれ値とみなすかどうかの判定基準はデータや文脈依存で明確に定めることができないが、四分位範囲を1つの目安として用いることがある。

箱ひげ図を作る際には、教科書の定義的には、最小値や最大値までひげを伸ばすとなっているが、それだとはずれ値までひげを伸ばすことになってしまう。それに対して、はずれ値を見出す目安として、第1四分位数から四分位範囲の1.5倍を引いた値と第3四分位数に四分位範囲の1.5倍を足した値を設定し、その範囲の外側にある値をはずれ値として拾い出すという方法がある。もちろんそれだけで決めるのではなく、そうしたデータが存在することを確かめて除外するかどうか扱いを決めるとするのが手順であるが、わかりやすい指標として使うことができる。

この内容は中学校学習指導要領及び解説には記載がないため指導内容ではない。

(7) 箱ひげ図を用いた問題解決活動

問題解決活動は小・中・高等学校を通じて新教育課程で求められるところであるが、指導時数が十分になれば到底無理である。しかも理想的な問題解決においては、箱ひげ図のみを用いるのではなく、小学校で学習したグラフやヒストグラムも含めて総合的に活用するような問題解決である。他教科との連携や総合的な学習の時間を用いるなど策を講じなければかなり難しいだろう。

(8) 箱ひげ図の指導において必要な事項について

以上のことから箱ひげ図の指導に必要な事項について、本研究においては次のように設定した。

- 四分位数・四分位範囲（5数要約）
- 箱ひげ図の定義・書き方・読み方
- 箱ひげ図の利点
- 箱ひげ図の欠点
- 箱ひげ図とヒストグラムの比較
- はずれ値の見極めと扱い
- 箱ひげ図を用いた問題解決活動

4. 榊原実践での箱ひげ図の指導について

前章で設定した指導事項を視点にして実際の指導について検討してみる。知立市立知立中学校において榊原健先生が第2学年を対象に2019年2月～3月にかけて指導提案を試みてくれたので、それを参考に検討してみることとする。

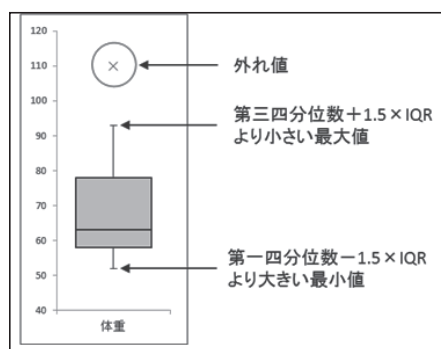


図4：はずれ値が表示された箱ひげ図
(画像はBellCurve『統計WEB』より)

(1) 時数と単元構成

榊原実践における単元構成は5時間である。最初に触れた指導時数に対する検討から5時間というのは現実的に無理のない範囲であろう。第2学年1組(習熟度別のため18人)を対象として実践である。

表1：榊原実践の単元構成

第0時	ジャストストップ！20seconds！の実験をしよう
第1時	データをヒストグラムにして比べよう！
第2時	自分たちのデータを箱ひげ図にし、比べてみよう！
第3時	どのような人がジャストストップ！20seconds！が得意か考えよう！
第4時	夏に過ごしやすい都市はどこか箱ひげ図を駆使して考えよう！
第5時	仮説が正しいかどうかを箱ひげ図を駆使して検証しよう！

「ジャストストップ！20seconds！」という活動が単元の主軸になっている。ストップウォッチを目視せず自分の感覚で20秒をねらって止めるというものである。数学の授業の時間を使わず、朝の時間などを用いて第2学年の全クラスのデータを集めて授業の題材としているため、第0時としている。

ヒストグラムや箱ひげ図を作成する際には、宮崎大学・藤井先生の開発した「SimpleHist」と「SimpleBox」を用いているが、教師が事前にグラフを作成しておき、生徒にはそれをワークシートなどに印刷して配るといった形式で進めている。

(2) 各授業の概要

第1時では、事前に集めた生徒の記録を元に、学年全クラス(7学級)の記録をヒストグラムにして提示し(図5)、どのクラスが時間感覚が優れていそうかについて話し合った。図5の7つのヒストグラムにはどれも中に縦線が引いてあるが、20秒ジャストを示す線である。

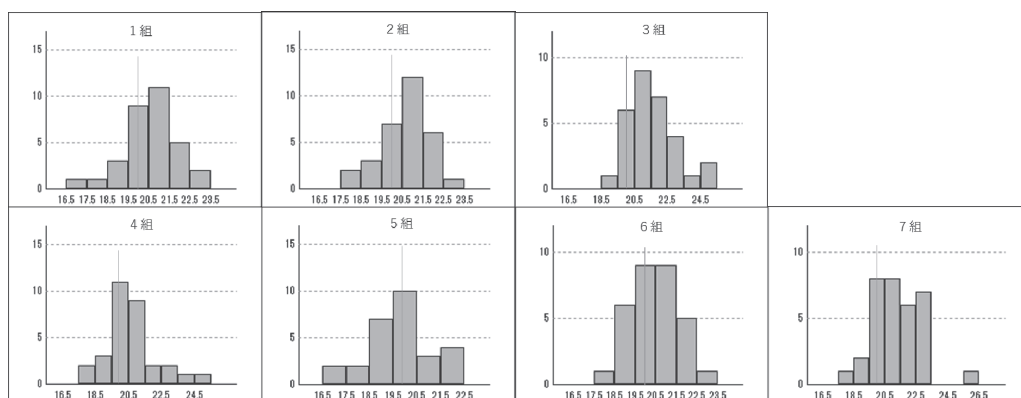


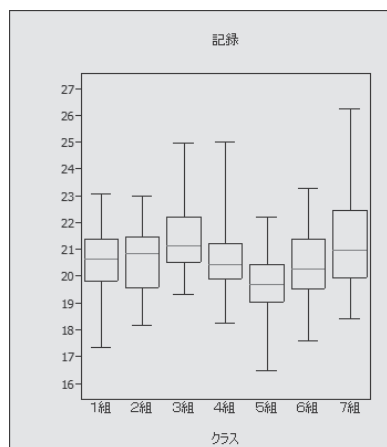
図5：7学級の記録のヒストグラム

1年での学習内容の復習とヒストグラムで多群比較を行うのは大変であるという印象を持たせて、次時の箱ひげ図の導入へとつなげるねらいである。生徒からは、「(ヒストグラムが)7つ

もあると少し見比べにくかった」という声も出ていた。

第2時ではテストの点数を例に箱ひげ図を導入部で教師から紹介した。続けて、記録順に並べ替えた1組の記録と2組の記録を載せたワークシートを配布し、箱ひげ図を作る練習をさせた。その際に合わせて5数要約や四分位範囲などの用語も伝えた。

その後、全クラスの箱ひげ図(図6)を提示すると、「ぱっと見てとても分かりやすい」という声も出ていた。全クラスの特徴を書き出させてみると、次のような反応が得られた。



- ・5組の箱が20秒に寄っているから、1位だと思う。
- ・箱もひげも上の方に寄っていて、全体的に記録が遅いため、7位だと思う。
- ・4組と6組を比べると、4組の方が箱は小さくてぎゅゅとしているけどひげも長い。どちらのクラスの記録がいいか難しい。

図6: 7学級の記録の箱ひげ図

第3時では、この実験が得意な人にどんな特徴があるのか仮説を考えさせた。本時でのアイデアを元に学年全体でアンケートを実施し、第5時にはその分析を行うことを伝え、各自仮説を考えさせた。6人グループ話し合いをさせたが、仮説自体は1人1つずつ考えさせた。回答の種類が多くなると、分析の際の分類が多くなりすぎてしまうと考える、基本的に2択の質問を考えさせるようにした。次のような仮説が出てきた。

- ・早起きしている人の方が、記録が良いのではないかな。
- ・ほぼ毎日朝ご飯を食べる人の方が、記録が良いのではないかな。
- ・自分は心配性だと思う人の方が、記録が良いのではないかな。
- ・料理をよく作る人の方が、記録が良いのではないかな。
- ・音楽関係の部活・習い事をしている人は記録が良いのではないかな。

各自の仮説に基づき、アンケートを作成し、朝の会などを利用して学年全クラスで実施してもらった。

第4時では、生徒の箱ひげ図の読み取りの仕方に不安があったため、8月の最高気温のデータを箱ひげ図で分析させることとした。気象庁から気象データをダウンロードし、岡崎、石垣島、東京、軽井沢、多治見、札幌の中でどの都市が夏に過ごしやすいのかを分析させた。

第5時では、アンケート結果に基づき、各自の仮説に沿った箱ひげ図を準備し、印刷して配布

した。質問項目の回答が2択となっているため、さらにそれを男女で層別し、どの生徒も4群比較をする形で箱ひげ図は作成した。

個別の分析の後、グループで各自の分析結果を発表させた。「四分位範囲が狭いため、中央50%の人の記録が寄っている。ひげも短めだから、記録が安定している」など、用語を用いて発表をする生徒もいた。

(3) 授業者による反省

実践を終えた後の榊原先生の反省事項としては3つ挙げられていた。

- ・時数はもう2, 3時間ほしい
- ・生徒の箱ひげ図の読み取りに怪しい部分があるため、さらなるケアを必要とする
- ・批判的な考察を取り入れたい

生徒の箱ひげ図の読み取りに関しては、「この箱ひげ図は箱とひげのバランスがいいから記録が良いといえるのではないか」という類の意見が何度も出ているのだが、20秒に近いかどうかの位置を考慮せずに上下の長さのバランスのみに注目しているのではないかという印象がある。

また、お互いの分析結果を発表するだけで終わってしまったため、相手の発表内容に対し批判的に検討し、お互いが意見を交流するような活動も取り入れたいということである。読み取りに関する怪しいところも批判的に交流することで改善が見込まれる。

(4) 指導事項との対応と考察

榊原実践において扱った指導事項について分析すると次のようになる。○はほぼ扱えた内容、△は部分的、×は扱えていない内容となる。なお、指導後に生徒の習得状況を調査した結果に基づくものではなく、あくまで授業の活動として取り入れることができたかどうかということと、授業内での生徒の活動を観察した結果に基づくものである。

表2：本実践での箱ひげ図の指導事項の扱いについて

指導事項	扱い
・四分位数・四分位範囲（5数要約）	○
・箱ひげ図の書き方・読み方	△
・箱ひげ図の利点	○
・箱ひげ図の欠点	×
・箱ひげ図とヒストグラムの比較	△
・はずれ値の見極めと扱い	△
・箱ひげ図を用いた問題解決活動	△

箱ひげ図を読むことについては、第2時、4時、5時と多くの時間を使ったが、箱ひげ図の誤認に対するケアは扱えておらず、また、榊原先生の反省にもあるように、バランスがいいことと記録がいいことを混同するような様子も観察された。

ヒストグラムでの多群比較を経て箱ひげ図を提示することで利点は伝わりやすかったと思われるが、ヒストグラムの利点を改めておさえることなどはできなかった。箱ひげ図の欠点についてもほとんど触れられていない。はずれ値については7組の男子1人が学年全体で最も長くなっており、その生徒がいなかったら7組の箱ひげ図の様子が変わるなど若干話題に出ていた程度である。

問題解決活動を5時間の単元構成の中でなんとか取り入れることができた。仮説を生徒に考えさせるなどできたが、教師が印刷した箱ひげ図のみで分析していたため、もっと充実させたい。例えば仮説や調査項目についてもっと生徒のアイデアを取り入れることや、分析に際して箱ひげ図だけでなく他の統計グラフや統計量も用いて総合的に判断させるなどである。

5. まとめと今後の課題

本稿においては、中学校第2学年で箱ひげ図が指導されるにあたり、指導時数や必要な指導事項について検討・提案を行った。また知立中学校での指導事例について指導事項の観点で検討を加えた。本稿で提案した指導事項は箱ひげ図の指導に関わって指導すべきと思われる事項を先行研究も参考にまとめたものだが、実際の指導においてすべてを取り入れなくてはならないという趣旨での提案ではない。あくまで実践に際して、あるいは実践後の検討の目安としての提案である。指導時数がそれほど多く見込めないのは確かであり、榊原実践の5時間は妥当な時間設定だと思われる。指導事項に関して扱いきれていなかった理由も、指導者の力量や単元設計の問題ではなく、5時間では土台無理であるともいえる。現実的に指導できる時間の中で効率を上げることや、指導事項についての取捨選択も考える必要があるだろう。今後の課題としては、指導事項についてのさらなる検討と効果的な指導法や教材の開発、指導事項を網羅するために必要な時数の設定や単元構成を検討することが挙げられる。

引用・参考文献

- 岡本和夫他. 2015. 『未来へひろがる数学2』. 啓林館.
- 小口祐一. 2010. 「データ分布の読み取りにおける学習者の誤った判断－様のバイアスと箱ひげ図の誤った読み取り－」, 『科学教育研究 Vol.35, No.2』. pp.128-138.
- 高橋陽一郎他. 2011. 『数学I』. 啓林館.
- BellCurve. 『統計WEB』 (<https://bellcurve.jp/statistics/>) (2019年11月20日参照)
- 松元新一郎・梅田英之・富田真永. 2012. 「高等学校・数学I教科書における「データの分析」の現状と今後のあり方」. 『日本数学教育学会数学教育論文発表会論文集, 45』. pp.719-724.
- 文部科学省. (2017). 『中学校学習指導要領解説数学編』. 文部科学省.