

中学校数学科における箱ひげ図の指導とその留意点

中村好則*

(2019年2月21日受理)

Yoshinori NAKAMURA

Guidance and notes of box plot in mathematics of Junior High School

1 はじめに

平成29年3月に小・中学校学習指導要領が、平成30年3月には高等学校学習指導要領が公示され、小・中・高等学校教育を通じて統計的な内容の改善・充実がなされた。特に、今回の改訂では、中学校数学科の統計領域において、従前の「資料の活用」の領域の名称が「データの活用」に改められた。これにより、小学校算数の「データの活用」領域と高等学校数学Iの「データの分析」の内容とともに、統計指導の内容構成の系統性がより明確となった。内容的には、中学校数学科では、「具体的には、第1学年で、従前どおりヒストグラムや相対度数を扱うとともに、第2学年で、四分位範囲や箱ひげ図を新たに扱うこととし、収集したデータから次第に情報を縮約することによって、大量のデータや複数の集団の比較が可能となるよう構成した。また、それぞれの学年において学んだ統計的な表現を関連付けながら統計的に問題解決することによって、より深い統計的な分析が可能となるように構成した（文部科学省2018, p.10, 下線と波線は筆者、以下同様）」ことが述べられている。また、中学校第1学年で扱われていた中央値や最頻値が小学校第6学年へ移行し、小学校第6学年ではドットプロットが新しく内容に加わり、小学校第6学年から中学校第2学年まで継続的に量的なデータの解析の学習が行われるように

なった（藤井2017）。

また、高等学校学習指導要領解説数学編（文部科学省2019, p.45）では、中学校の統計領域の指導について、以下のように述べられている。

『小学校、中学校の「データの活用」領域においては、「問題－計画－データ－分析－結論」の五つの段階からなる統計的探究プロセスを意識した、統計的な問題解決の活動が大切にされている。

問題	・問題の把握	・問題設定
計画	・データの想定	・収集計画
データ	・データの収集	・表への整理
分析	・グラフの作成	・特徴や傾向の把握
結論	・結論付け	・振り返り

中学校では、このようなプロセスを通して、複数の集団のデータの分布に着目し、その傾向を比較して読み取り批判的に考察して判断したり、不確定な事象の起こりやすさについて考察したりする力などを養っている。』

つまり、中学校の統計領域の指導においては、統計的なデータの活用の内容を学習することによって、統計的に問題解決する力を育てることが求められている。

そこで、本論では、特に、中学校数学科におい

* 岩手大学教育学部

て、新たに指導する内容となった箱ひげ図（中学校第2学年Dデータの活用D（1）データの分布）に焦点を当て、その指導と指導上の留意点について考察する。

2 箱ひげ図とは

1) 箱ひげ図と四分位数の求め方

箱ひげ図は、平成21年の学習指導要領の改訂で、高等学校の数学Iの「データの分析」において指導されることとなった。箱ひげ図と関連して、四分位数、四分位範囲、四分位偏差が扱われた（文部科学省2009, p.25）。今回の改訂で、四分位偏差（四分位範囲を2で割った値）は中学校数学では扱われない。以下では、箱ひげ図に関わって、いくつかの用語を確認する。

箱ひげ図は、図1のように、最小値、第1四分位数Q1、中央値（第2四分位数Q2）、第三四分位数Q3、最大値を箱と線（ひげ）を用いて1つの図に表したものである。つまり、四分位数とは、全てのデータを小さい順に並べて四つに等しく分けたときの三つの区切りの値を表し、小さい方から第1四分位数、第2四分位数、第3四分位数という。第2四分位数は中央値のことである（文部科学省2018, p.120, 四分位数の定義）。また、最小値、第1四分位数Q1、中央値（第2四分位数Q2）、第三四分位数Q3、最大値の5種類の値で代表する方法を5数要約という（椛元編2013, p.13）。箱ひげ図に図1のように平均値を示すときもある。

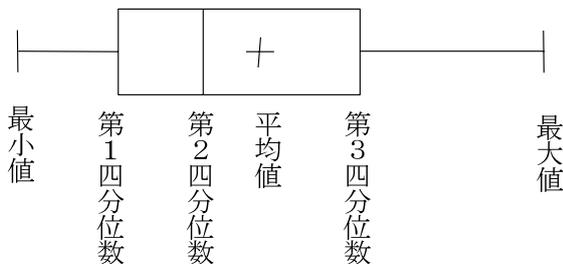


図1 箱ひげ図と5数要約

ここで注意しなければならないのは、四分位数を求める方法が複数あるということである。中学校学習指導要領解説（文部科学省2018, p.120）では、四分位数の求め方そのものは書いていないが、次の具体例で四分位数の求め方を示している。

「例えば、次の九つの値があるとき、中央値（第2四分位数）は5番目の26である。

23 24 25 26 26 29 30 34 39

この5番目の値の前後で二つに分けたときの、1番目から4番目までの値のうちの中央値24.5を第1四分位数、6番目から9番目までの値のうちの中央値32を第3四分位数とする。」



図2 解説の箱ひげ図と四分位数の例 (P.121)

つまり、全てのデータを小さい方から順に並べて中央値を取りそれを第2四分位数とする。その中央値（第2四分位数）を除いて、中央値より小さいデータの中央値を取りそれを第1四分位数とする。中央値（第2四分位数）より大きいデータ

(ア) $N \equiv 0 \pmod{4}$ のとき、つまり、 N が4で割り切れるとき

例1 $N=8$ のとき

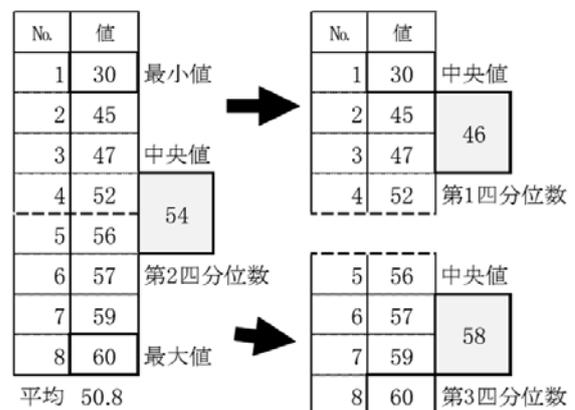


図3 四分位数の求め方の例(ア) (平山2010, p.2)

(イ) $N \equiv 1 \pmod{4}$ のとき、つまり、 N を 4 で割ると余りが 1 のとき

(エ) $N \equiv 3 \pmod{4}$ のとき、つまり、 N を 4 で割ると余りが 3 のとき

例2 $N=9$ のとき

例4 $N=11$ のとき

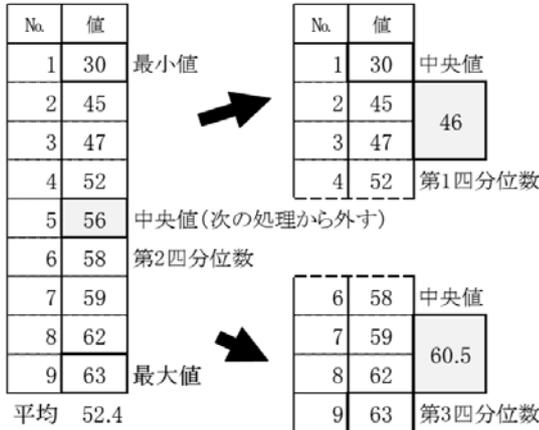


図4 四分位数の求め方の例(イ) (平山2010, p.2)

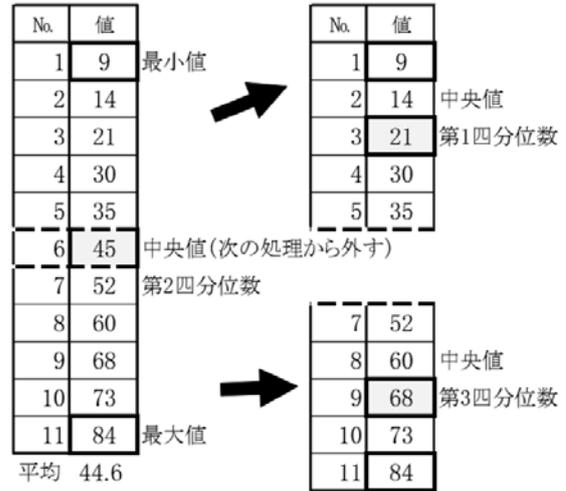


図6 四分位数の求め方の例(エ) (平山2010, p.3)

(ウ) $N \equiv 2 \pmod{4}$ のとき、つまり、 N を 4 で割ると余りが 2 のとき

例3 $N=10$ のとき

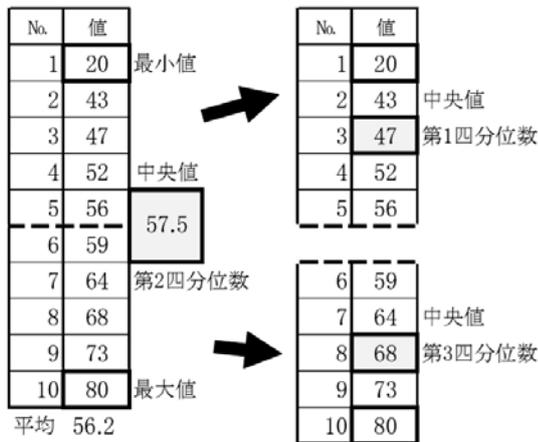


図5 四分位数の求め方の例(ウ) (平山2010, p.3)

の中央値を取ってそれを第3四分位数とする。この求め方だと、データの個数によって、4通りの求め方があることになる。例えば、平山(2010)は、このことについて、以下の図3から図6の4つの例で説明している。

2) 箱ひげ図の指導へのICT活用

学習指導要領では、「コンピュータなどの情報手段を用いるなどしてデータを整理し箱ひげ図で

表すこと」が述べられている(文部科学省2018, p.120)が、学校現場でも多く使われているExcelでの四分位数の求め方は、学習指導要領解説の求め方とは異なることは有名である。箱ひげ図は、Excel2016からグラフ作成機能に加わった。そのときに使用される四分位数の求め方は、2種類ある。どちらも、学習指導要領解説の求め方とは異なる。Excelで箱ひげ図を作成するときの四分位数計算は、「包括的な中央値」と「排他的な中央値」があり、「データ系列の書式設定」でどちらかを選択できる。第1四分位数と第3四分位数の計算時に、「包括的な中央値」は中央値を含めて計算し、「排他的な中央値」は中央値を除いて計算する。この場合、中央値(第2四分位数)はどちらの計算方法でも同じ値となる。Excelの関数では、「包括的な中央値」はQUARTILE.INC(クアタイル・インクルーシブ, INCはincludeの略)に対応し、「排他的な中央値」はQUARTILE.EXC(クアタイル・エクスクルーシブ, EXCはexcludeの略)に対応する。従来からあるExcelの関数QUARTILE(クアタイル)は、QUARTILE.INCと同じ計算結果となる(図7)。

QUARTILE.INC (配列, 戻り値)	
0%と100%を含めた範囲	
QUARTILE.EXC (配列, 戻り値)	
0%と100%を含めない範囲	
QUARTILE (配列, 戻り値)	
戻り値	
0	最小値
1	第1四分位数 (25%の位置)
2	第2四分位数 (50%の位置), 中央値
3	第3四分位数 (75%の位置)
4	最大値

図7 四分位数と求める Excel の関数

図8は、学習指導要領解説の例のデータを使って、Excelで、四分位数を求めたものである。「包括的な中央値」では第1四分位数と第3四分位数が解説のものとは異なる。「排他的な中央値」では同じ値になっているが、求め方は違っても同じ値になることはある。初期設定では「排他的な中央値」になる。

No.	データ	五数要約	文科省	包括的な中央値	排他的な中央値
				QUARTILE.INC	QUARTILE.EXC
1	23	最小値	23	23	#NUM!
2	24	第1四分位数	24.5	25	24.5
3	25	第2四分位数 (中央値)	26	26	26
4	26	第3四分位数	32	30	32
5	26	最大値	39	39	#NUM!
6	29			0%と100%を含めた範囲	0%と100%を含めない範囲
7	30			中央値を含める	中央値を除く
8	34				
9	39				

図8 Excel関数による四分位数の計算結果

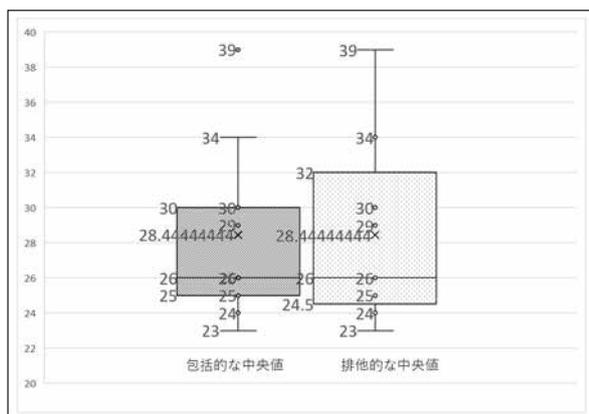


図9 Excelで作成した箱ひげ図

図9は、学習指導要領解説の例のデータを使って、Excelで、箱ひげ図を作成したものである。初期設定では、外れ値を検出し、外れ値のある箱

ひげ図を作成する（「特異ポイントを表示する」がオン）。外れ値とは、他の値から大きく外れた値であり、箱ひげ図では、箱ひげ図の箱の幅（四分位範囲=第3四分位数-第1四分位範囲）の1.5倍以上、第1四分位数から小さい場合、或いは第3四分位数から大きい場合を外れ値とする。しかし、中学校では外れ値は考慮されない場合が多いので注意する必要がある。つまり、中学校では、外れ値は考慮されずに、ひげの長さは、最小値から第1四分位数と第3四分位数から最大値までである。データが多い場合は、「包括的な中央値」と「排他的な中央値」での箱ひげ図の違いは出にくい。データ数が少ない場合は、「包括的な中央値」の箱ひげ図は、箱の大きさが短く、ひげの長さが短く、外れ値が表示されやすくなる。

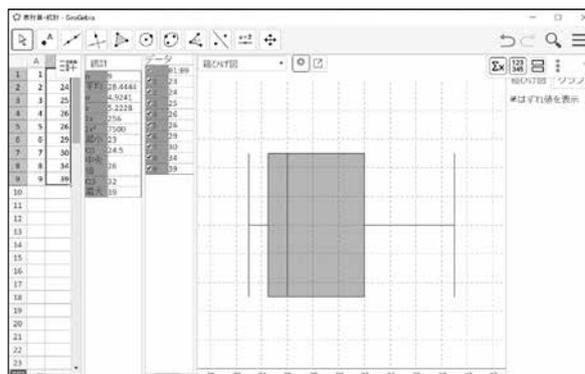


図10 GeoGebraで作成した箱ひげ図

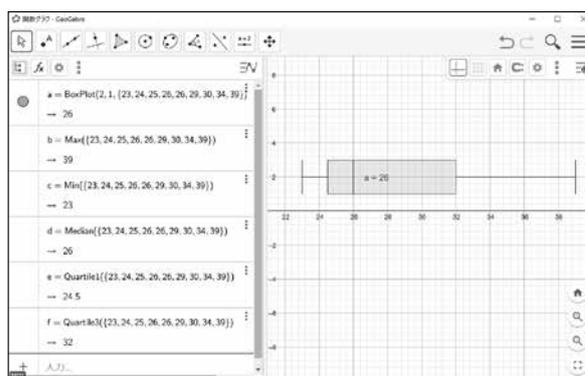


図11 GeoGebraのグラフ関数機能による箱ひげ図

一方、GeoGebraでは、四分位数の計算は、学習指導要領解説と同じ計算方法が使われる。図10は、GeoGebraの表計算・統計の機能を使って、箱ひげ図を作成したものである。

また、GeoGebraの関数グラフ機能でも、箱ひげ図の作成と、四分位数の計算ができる。箱ひげ図はBoxPlot関数、最大値はMax関数、最小値はMin関数、中央値はMedian関数、第1四分位数はQuartile 1関数、第3四分位数はQuartile 3関数で求められる(図11)。

「四分位範囲や箱ひげ図の必要性と意味を理解すること(アの(ア),アの(イ))」では、「極端にかけ離れた値が一つでもあると、最大値や最小値が大きく変化し、範囲はその影響を受けやすいが、四分位範囲はその影響をほとんど受けないという性質がある。また、この図中に、平均値を記入して中央値との差を考えたり、第1四分位数や第3四分位数と中央値との差を考えたりすることにより、データの散らばり具合が把握しやすくなるので、複数のデータの分布を比較する場合などに使われる」ことが述べられている(p.121)。

3 箱ひげ図の指導とその留意点

1) 箱ひげ図の指導

学習指導要領解説では「四分位範囲や箱ひげ図を用いて、複数の集団のデータの分布の傾向を比較して読み取り、批判的に考察したり判断したことを説明したりすることができるようにする」ことが述べられている(文部科学省2018, p.121)。具体例として、「中学生の体力」について、2000年、2005年、2010年、2015年のハンドボール投げのデータをもとに考察したものが挙げられている。これ

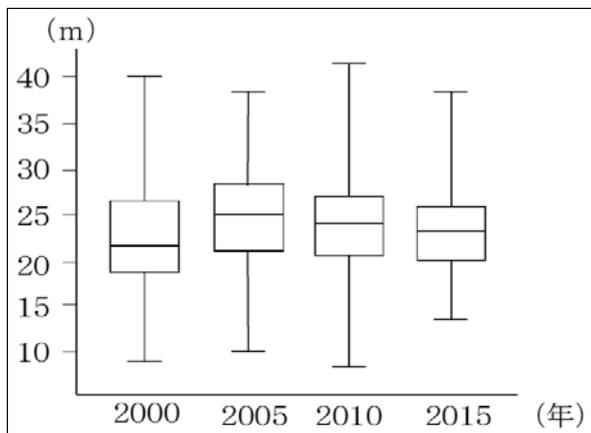


図12 ハンドボール投げの分布 (p.122の図1)

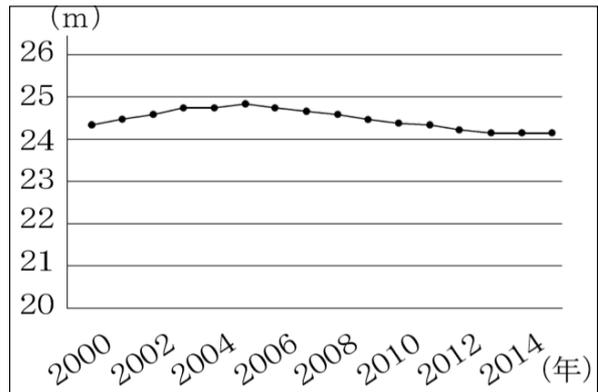


図13 ハンドボール投げの平均値の経年変化 (p.122の図2)

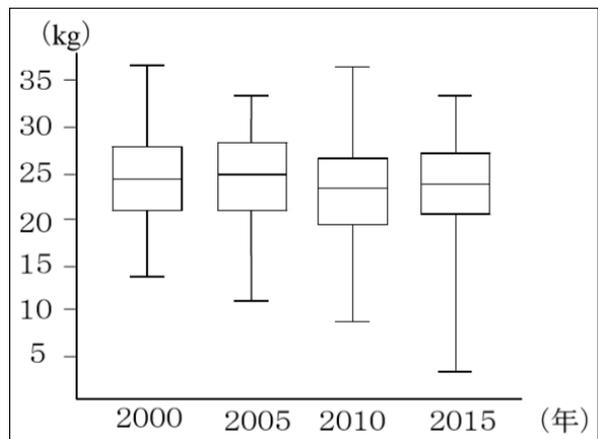


図14 握力の分布 (p.122の図3)

らの箱ひげ図をもとに「四分位範囲を表す箱は、2005年からそれほど大きく下がっておらず、中央値を中心とする全体の約半数のデータはそれほど下がっているわけではないので、体力が落ちているとは言えない」と判断することが考えられることが述べられている。

学習指導要領解説では、さらに「体力に対して多様なデータや統計的な表現を用いて多面的に吟味することで、批判的に考察することの必要性に気付くことが大切である」ことが述べられている(p.121)。5年ごとではなく毎年の中央値や平均値などに着目して折れ線グラフ(図13)を作成したり、ハンドボール投げのデータだけでなく、握力などの他の体力テストのデータから箱ひげ図(図14)やヒストグラムを作成するなど、さらに詳しい考察を加えることも考えられることが述べられている。

留意点としては、「データの傾向を捉える場合、日常生活では、簡潔さの観点から箱ひげ図のみを用いて説明することが予想される。しかし、そのことによって分布の形など、失われる情報もあるので、必要に応じてヒストグラムなどと合わせて用いることが必要な場面もあることに留意する」ことが述べられている (p.121)。

2) 箱ひげ図の指導上の留意点

藤井 (2009) では、中学校での分布の見方として、

- ① 分布自身に興味がある場合、
- ② 基準グループと分布を比較する場合、
- ③ 2つの集団の分布の比較をする場合

があることを挙げている。さらに、藤井(2017)は、これに箱ひげ図が加わることで、

- ④ 3つ以上の集団の分布を比較する場合
- ⑤ さまざまな視点からデータの分布を比較する場合

の2つを加えている。その理由として、「箱ひげ図の特徴は、分布の特徴を最大値、最小値と3つの四分位数という5つの数値だけで表現している点になる。もちろん5つの数値だけに要約しているため、ヒストグラムに比べて情報がかなり失われていることになる。そのため、2つのデータの分布の比較をする場合には、ヒストグラムや度数折れ線などを用いた方がさまざまな観点から分析することができるだろう。その意味で、箱ひげ図を活用するためには、上の3つの分布の見方だけでなく、…略…2つを加えておきたい」ことを述べている。

3) ICT を活用した箱ひげ図の指導

小林 (2013) は、箱ひげ図は同じでも、ヒストグラムが異なる例を挙げている。以下は、100点満点のテストの結果で、クラスAは13人のデータ、クラスBは12人のデータである。

A : 40, 10, 70, 20, 60, 50, 20, 80, 70, 70, 20, 40, 40

B : 30, 10, 80, 30, 10, 10, 40, 40, 80, 80, 60, 60

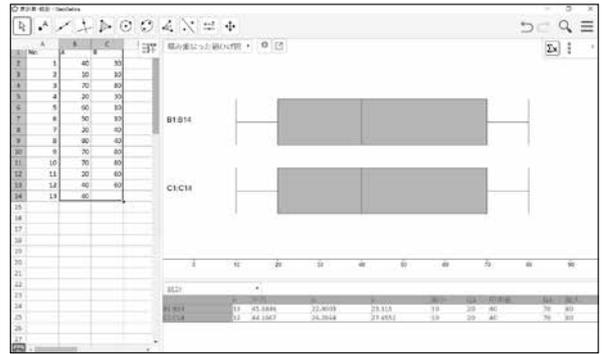


図15 GeoGebra で作成した箱ひげ図の比較

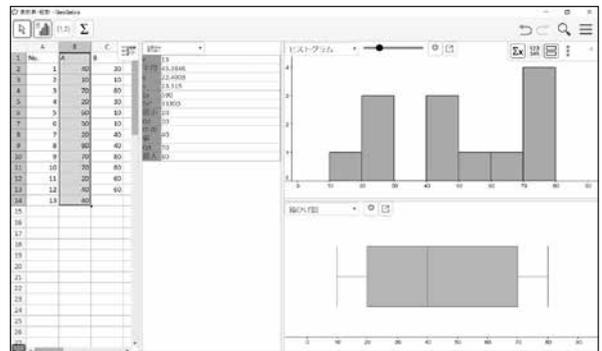


図16 GeoGebra によるヒストグラム (クラス A)

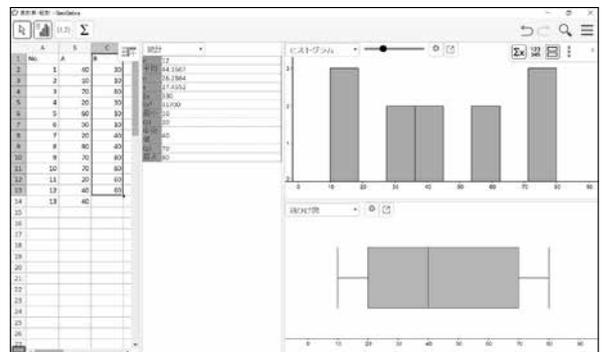


図17 GeoGebra によるヒストグラム (クラス B)

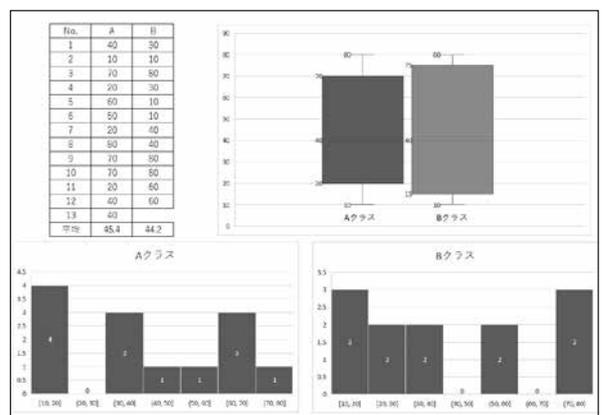


図18 Excel による箱ひげ図とヒストグラム

GeoGebra で箱ひげ図を作成すると、四分位数はどちらも同じ値であり、箱ひげ図も同じグラフとなる（図15）。しかし、それぞれのヒストグラムを作成してみると、散らばり具合の違いがよくわかる（図16と図17）。様々な視点からデータの分布を比較することの重要性がよくわかる。

Excel で箱ひげ図を作成すると、クラス A とクラス B では、第 1 四分位数と第 3 四分位数が異なる値となり、箱ひげ図も同じにはならないので注意が必要である（図18）。Excel2016からは、ヒストグラムも簡単に作成できるようになった（図18）。

4) ICT を活用した指導事例

課題「盛岡市は仙台市より本当に暑いと言えるのか」を題材に、ICT を活用した指導事例を紹介する。統計的探究プロセスの①問題、②計画、③データ、④分析、⑤結論に沿って述べる。

① 問題（問題の把握、問題設定）

盛岡市は、仙台市よりよりも暑いとよく言われる。しかし、盛岡市は仙台市よりも約180km も北に位置し、本当に暑いのだろうか。確かに、盛岡市は内陸に位置し、仙台市は海に面しているため、盛岡市の方が暑いかもしれない。このような疑問から、探究課題を「盛岡市は仙台市より本当に暑いと言えるのか」と設定することにした。

国立大学教育学部理数教育コース第3学年の学生34名を対象に「盛岡市は仙台市より暑いか」を質問し、「そう思う、そう思わない、どちらとも言えない」の三択で回答を依頼した（2019年5月20日）。その結果は、図19の通りである。回答した学生の出身県は図20の通りである。56%もの学生が「そう思う」と回答しており、感覚的には盛岡市の方が仙台市よりも暑いと考えられるが、実際にはどうなのだろうか。

② 計画（データの想定、収集計画）

盛岡市と仙台市の8月の気温のデータが必要である。

気温は、1日の平均気温、最高気温、最低気温が考えられるが、暑さが課題のなので、最高気温

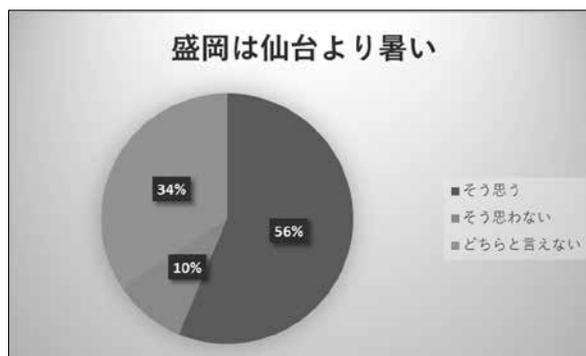


図19 「盛岡市は仙台市より暑いか」の調査結果

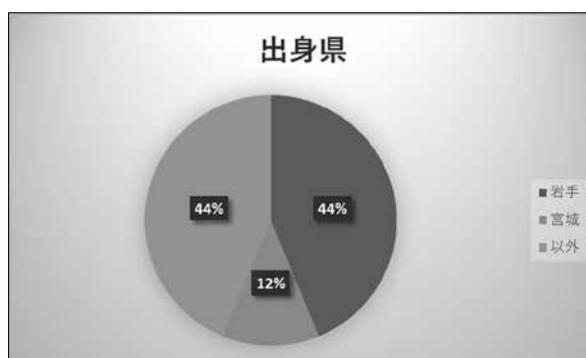


図20 調査対象学生の出身県

がわかればよさそうである。暑さと関係があるデータとして、場合によっては、湿度のデータも収集することにした。

これらのデータは、気象庁のホームページ (<https://www.jma.go.jp/jma/index.html>) から得ることができる。本来は、複数年を比較する方がいいが、最初はまず最新の2018年8月のデータで考察することとした。

③ データ（データの収集、表への整理）

データの収集計画に基づき気象庁のホームページ (<https://www.jma.go.jp/jma/index.html>) から、必要なデータをダウンロードする。データは、CSV ファイルなので、Excel ファイルに変換し、保存する（表1）。気象庁のホームページでは、対象の市町村の月ごとの折れ線グラフが作成可能であったので、仙台市（図21）と盛岡市（図22）の折れ線グラフを入手した。これらの比較だけではよくわからない。

④ 分析（グラフの作成、特徴や傾向の把握）

仙台市と盛岡市の2018年8月の最高気温をもと

表1 盛岡市と仙台市の気温データ (2018年8月)

日	平均気温		最高気温		最低気温	
	仙台	盛岡	仙台	盛岡	仙台	盛岡
1	30.7	27.3	37.3	32.3	25	23.3
2	28.8	26.6	33.5	32.7	25.5	21
3	28.2	24.8	33.6	31.5	23.9	18.2
4	27	22.9	32.9	29.4	23.1	17.9
5	26.6	20.5	34.4	23.4	22	18
6	20.3	18.5	22.8	21.2	17.4	17
7	18.5	19.3	19.9	21.8	17.3	16.6
8	22	22.2	25.4	27.8	18.5	18.6
9	24.1	23.3	25.5	31.2	22.8	18.6
10	25.4	25.1	31.8	31.9	22.8	19.7
11	26.7	24.9	32	31	22.7	19.9
12	25.3	24.5	28.6	30.7	22.8	17.8
13	26.8	26	30.8	32.3	24.2	21.6
14	27.3	25.8	31.3	32	24.4	19.6
15	28.8	27.4	34.3	33.7	25.3	23.5
16	24.2	21.2	26.9	23.7	22	19
17	21.7	17	24.9	19.5	17.5	12.5
18	21	17.7	26.2	24.5	15.7	10.6
19	22.2	19.4	26.7	24.5	17.7	16
20	23.6	22.3	27.9	29	19.2	15.9
21	26.1	25	30	30.4	22.6	20.2
22	28.8	27.8	34.8	33.9	24.6	23.7
23	28.3	29	34.3	36.3	24.7	22.1
24	25.5	26	27.4	28.1	24.5	24.7
25	28.5	26.9	34.8	31.3	24.3	23.6
26	25.6	23.7	27.8	26.1	23	21.3
27	24.2	22.9	26.2	25.6	22	19.9
28	20.9	20.8	22	22.7	19.4	19.3
29	20.3	20.9	21.3	23.3	19.2	18.6
30	22.7	21.3	28.2	23.7	20.4	20.3
31	23.2	21.9	25.5	24.9	21.5	20.5
平均	24.94516	23.31935	29	28.07742	21.80645	19.33871

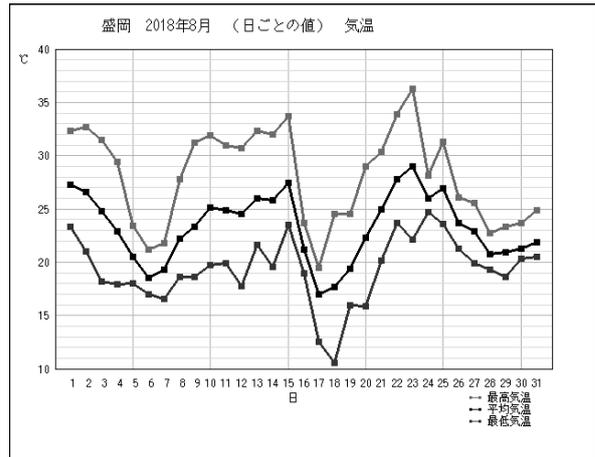


図22 盛岡市の気温の折れ線グラフ

方が小さく、その範囲に4分の1のデータがあるということである。そこで、仙台市と盛岡市の8月の最高気温のヒストグラムを作成することにした(図24と図25)。27.5度から33.5度の範囲の日数は盛岡市の方が多い。

仙台市と盛岡市の最高気温を日ごとに比べることとした。折れ線グラフが図26で、仙台市と盛岡市の最高気温の差を棒グラフにしたものが図27である。図27からは、盛岡市の方が最高気温が高かった日が13日(41.9%)あることがわかる。

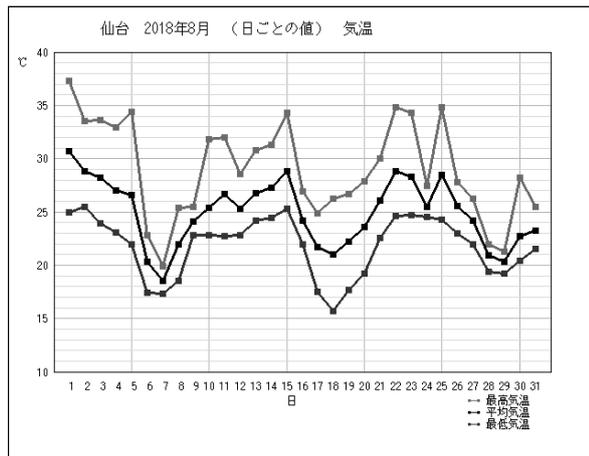


図21 仙台市の気温の折れ線グラフ

に、箱ひげ図を作成した(図23)。中央値は、盛岡の方が高いが、それ以外の5数要約は、仙台の方が高い。第1四分位数と中央値の差は、盛岡の

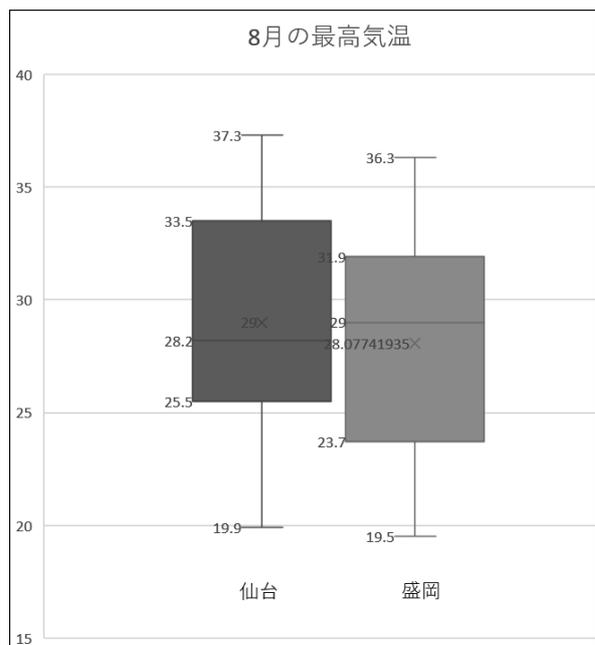


図23 8月の最高気温の箱ひげ図(仙台と盛岡)

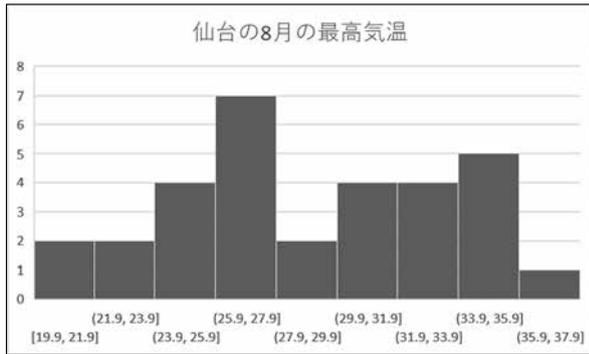


図24 仙台市の8月の最高気温のヒストグラム

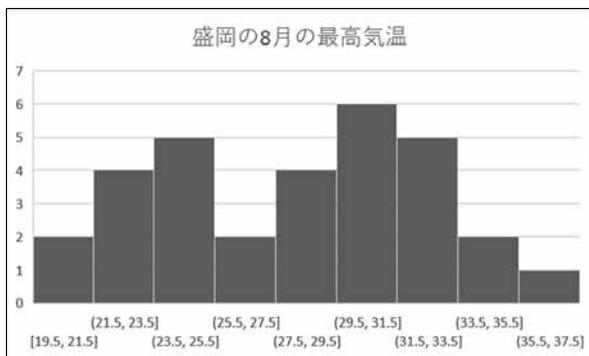


図25 盛岡市の8月の最高気温のヒストグラム

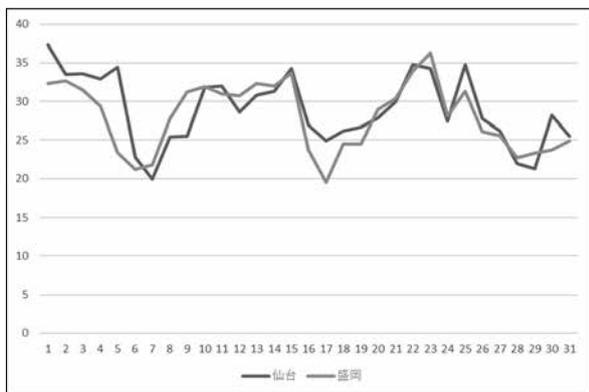


図26 仙台市と盛岡市の8月の最高気温

⑤ 結論（結論付け、振り返り）

今回の分析結果からは、「2018年の8月は仙台市の方が盛岡市よりも暑い」と言えそうである。しかし、盛岡市の方が200kmも北に位置していても約4割も盛岡市の方が暑い日があるなど、仙台市と同じ程度の暑さと言える。ここでは最高気温しか調べていないので、湿度も比較することが必要と考える。また、最高気温も2018年の8月だけ

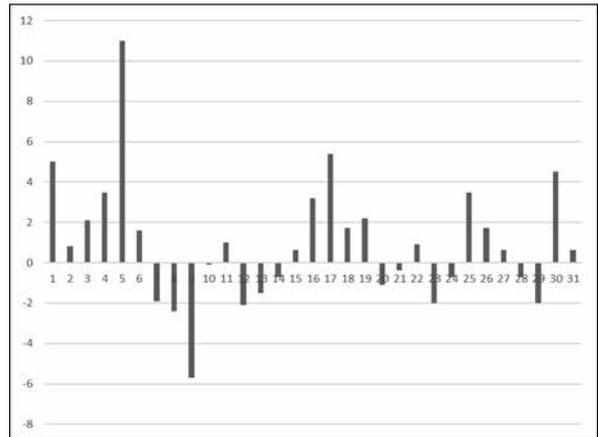


図27 仙台市と盛岡市の8月の最高気温の差

のデータの比較であったので、複数年の比較も必要である。

そこで、1988年から2018年までの8月の平均最高気温を収集し、比較してみることにした。データをもとに箱ひげ図（図28）とヒストグラフ（図29と図30）を作成すると、盛岡市の方が仙台市よりも全体的にやや上にある（図28）。仙台市と盛岡市の最高気温の差を取ると、盛岡市の方が暑かった日は961日中556日（57.9%）あり、長期的に見ると盛岡市の方が仙台市よりも暑かったと言える。

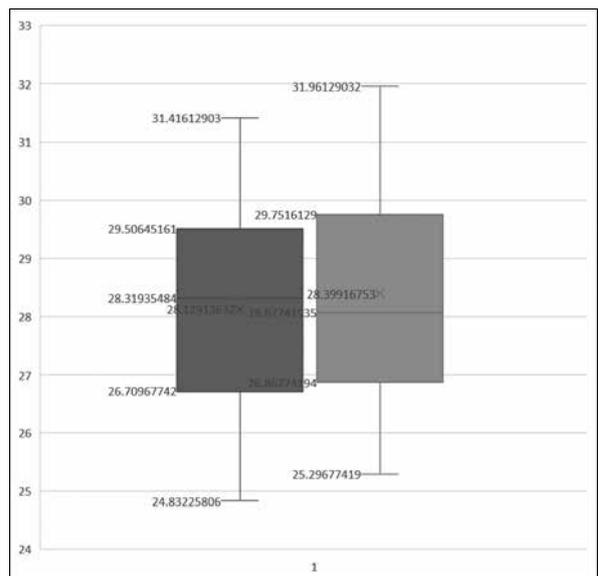


図28 10年間の最高気温の箱ひげ図（仙台と盛岡）

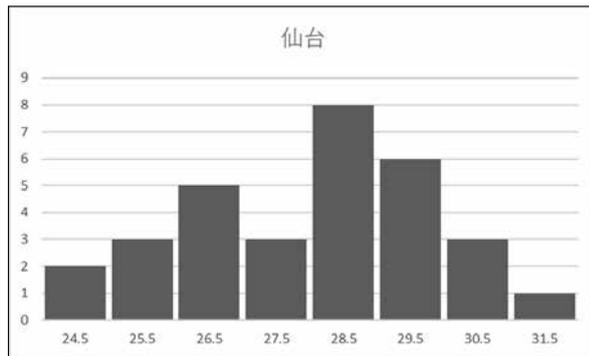


図29 仙台市の10年間の最高気温
(8月の最高気温の平均気温) のヒストグラム

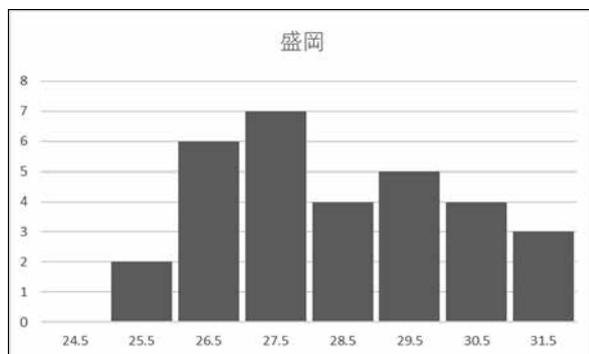


図30 盛岡市の10年間の最高気温
(8月の最高気温の平均気温) のヒストグラム

4 まとめと課題

本論では、中学校数学科における箱ひげ図の指導とその留意点について考察してきた。四分位数の求め方が複数存在し、ICTを活用する場面において、学習指導要領解説にある求め方(教科書の求め方)と異なることがあるので、指導上留意する必要があることが分かった。ICTの積極的活用を考えると、四分位数の求め方に留意しながらも、データの分布の傾向の比較や読み取りに積極的に活用し、統計的探究プロセスによる問題解決に取り組めるようにすることが重要である。

今後は、実際の中学校で箱ひげ図の指導を実践し、その成果を検討することが課題である。

付記

本論文は、2019年7月2日(火)に宮城県富谷市立東向陽台中学校で行われた富谷黒川地区教育

研究会数学会研修会での講演「中学校数学科における箱ひげ図の指導とその留意点」の内容を整理し、加筆修正したものである。

本研究の一部は科学研究費補助金「基盤研究(C)」課題番号JP18K02650によって行われた。

参考・引用文献

- 藤井良宜(2009)資料の活用での分布の見方,日本科学教育学会第33回年会論文集,pp.157-158.
- 藤井良宜(2017)箱ひげ図による探索的な問題解決,日本科学教育学会第41回年会論文集,pp.161-162.
- 平山貴(2010)高等学校数学科における学習指導要領改訂の趣旨を生かした事例の提案,青森県総合学校教育センター研究紀要,pp.1-10.
- 小林道正(2013)データ分析における「箱ひげ図」の誤解-高校教科書における多数の誤り-,中央大学論集第34号,pp.57-68.
- 松元新一郎編(2013)中学校数学科 統計指導を極める,明治図書.
- 文部科学省(2009)高等学校学習指導要領解説 数学編理数編 平成21年12月,東洋館出版.
- 文部科学省(2018)中学校学習指導要領(平成29年告示)解説 数学編 平成29年7月,日本文教出版.
- 文部科学省(2019)高等学校学習指導要領(平成30年告示)解説 数学編理数編 平成30年7月,学校図書.