

新課程「データの活用」で新たに追加される内容について

～四分位範囲，箱ひげ図～

数研出版 編集部

1. はじめに

中学校では 2021 年度から実施される次期学習指導要領において，統計分野の更なる充実が図られています。とりわけ，現行学習指導要領から高校の数学 I で扱われている「データの分析」から，「四分位範囲」「箱ひげ図」といった内容が，中学 2 年に移ることになったのが特徴です。

この四分位範囲や箱ひげ図は，ヒストグラムなどと違って日常で目にする機会がほとんどなく，なじみの薄いものかもしれません。しかしこれらは，データの分布を少ない手間でおおまかに読み取ったり，複数のデータの分布を比較したりする場面において，その力を発揮します。

そこで本稿では，新たに中学校で学ぶことになる四分位範囲と箱ひげ図について，その特徴と現在の高校での扱い，次期学習指導要領における中学校での指導でご留意いただきたい点などを紹介します。

2. 四分位範囲

四分位範囲は，四分位数とよばれる値から求められます。四分位数とは，データの分布を 4 等分したときの 3 つの区切りとなる値のことです。現在高校で学習している「データの分析」において，四分位数は次のように定義されるのが一般的です。

「データを値の大きさの順に並べたとき，4 等分する位置にくる値を四分位数という。四分位数は，小さい方から第 1 四分位数，第 2 四分位数，第 3 四分位数といい，これらを順に Q_1 ， Q_2 ， Q_3 で表す。第 2 四分位数は中央値である。

データを値の小さい方から順に左から並べたとき，左半分のデータを下位のデータ，右半分のデータを上位のデータと呼ぶことにする。ただし，データの大きさが奇数のとき，中央の位置にくる値は，下位のデータにも上位のデータにも含めないものとする。

下位のデータ						上位のデータ			
2	3	5	7	11		13	17	19	23
下位のデータ						上位のデータ			
2	3	5	7	11		13	17	19	23 29

このとき，第 1 四分位数 Q_1 ，第 3 四分位数 Q_3 を次で定める。

第 1 四分位数 Q_1 は 下位のデータの中央値

第 3 四分位数 Q_3 は 上位のデータの中央値

例えば、次のようなデータ①があるとします。

18, 20, 21, 23, 23, 25, 27, 29, 31, 31, 32, 39, 41 …… ①

このデータについて、第2四分位数すなわち中央値は $Q_2=27$

$$\text{第1四分位数は } Q_1 = \frac{21+23}{2} = 22$$

$$\text{第3四分位数は } Q_3 = \frac{31+32}{2} = 31.5$$

つまり四分位数は、データの分布を4等分したときに各部の境目となる値、と考えることができます。そのため、四分位数を求める際には、データを値の大きさの順に並べかえておくことで求めやすくなります。なお、四分位数の定義は他にもいくつかありますが、データの大きさが比較的大きい場合には、どの定義を用いても実用上意味のある差が出ることはありません。

四分位範囲は四分位数を用いて次のように定義されます。

$$\text{四分位範囲 } Q_3 - Q_1$$

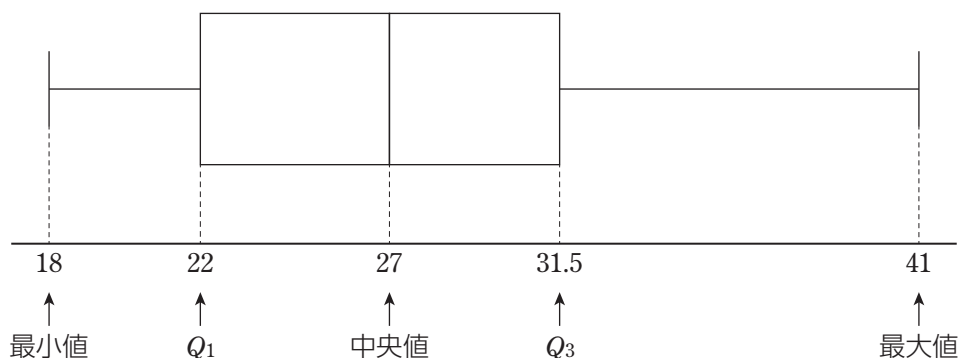
四分位範囲は、代表値を中央値(Q_2)としたときのデータの散らばりを表す値として使われます。すなわち、中心(中央値)付近のデータがどの程度散らばっているかの目安となる値で、 $Q_1 \sim Q_3$ の区間に全データの約50%が含まれることになります。このように、データの散らばりを表す値は「散布度」と呼ばれ、高校で学習する「分散」「標準偏差」も散布度です。

3. 箱ひげ図

四分位数などを使って、分布の広がり表現したものが箱ひげ図です。

箱ひげ図は、最小値・第1四分位数・中央値・第3四分位数・最大値の5つの値を図示することによって、データの分布の広がりをわかりやすく表現するためのグラフです。箱と線(ひげ)で表現する図で、箱の長さは四分位範囲を表します。

上のデータ①の箱ひげ図をかくと、下のようになります。



四分位範囲で述べたことと同様に、箱の中にはデータを4分割したとき、中央の約50%のデータが入ることになります。また、箱ひげ図に平均値を記入し、中央値との差を考えることもあります。なお、ここでは箱ひげ図を横向きにかきましたが、縦向きにかくこともよくあります。

4. 四分位範囲・箱ひげ図の特徴

四分位範囲には、次のような特徴があります。

- ・分散や標準偏差と比べて、計算の手間が少なくて済む場合が多い。
- ・四分位範囲が大きいほど、中央付近のデータの散らばりが大きい。
- ・極端に大きい値や極端に小さい値は $Q_1 \sim Q_3$ の区間から外れるので、極端に離れた値の影響を受けない。逆に、中央付近以外のデータについて、調べたり比べたりすることはしづらい。

データの大きさが大きくなると、分散や標準偏差は手計算で求めるのが大変です。四分位範囲であれば、それほど手間をかけずに中央付近のデータの散らばりを調べることができます。また、分散や標準偏差は平均値付近のデータの散らばりを表す値であり、データの中に極端な値があると、その影響を少なからず受けてしまいます。中央付近のデータの傾向を調べる場合は、四分位範囲を用いる方がデータの傾向を正しく読み取りやすくなる場合があります。

箱ひげ図には、次のような特徴があります。

- ・2つ以上のデータの分布を比較する際に、並べてかけるので視覚的に比較がしやすい。
- ・箱ひげ図は、ヒストグラムなどと比べて情報の精度は低い。そのため、一度データから箱ひげ図を作った場合、それをヒストグラムなどに戻すことはできない。

箱ひげ図は、2つ以上のデータについて並べてかくことができます。そのため、2つ以上のデータの分布を視覚的にわかりやすく表し、比較しやすくするという特徴があります。一方、箱ひげ図をかく元となる四分位数はデータを4等分した区切りとなる値であるため、個々のデータの値、データの大きさなどを箱ひげ図から読み取ることはできません。そのため、箱ひげ図をかいた後にヒストグラムや度数分布表をかく場合には、一旦元のデータに立ち戻る必要があります。

これらの特徴も含めて、『中学校学習指導要領解説 数学編（平成29年7月）』には、次のように記載されています（下線は編集部によるものです）。

D データの活用

D (1) データの分布

(1) データの分布について、数学的活動を通して、次の事項を身に付けることができるよう指導する。

ア 次のような知識及び技能を身に付けること。

(ア) 四分位範囲や箱ひげ図の必要性和意味を理解すること。

(イ) 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。

イ 次のような思考力、判断力、表現力等を身に付けること。

(ア) 四分位範囲や箱ひげ図を用いてデータの分布の傾向を比較して読み取り、批判的に考察し判断すること。

第1学年では、ヒストグラムや相対度数などについて学習している。第2学年では、これに加えて四分位範囲や箱ひげ図を学習することで、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する力を養う。

四分位範囲や箱ひげ図の必要性和意味(アの(ア)、アの(イ))

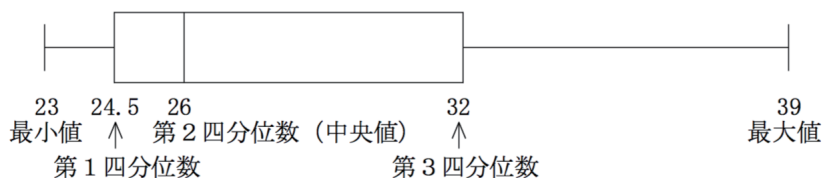
第1学年では、量的データの分布を捉える方法として、ヒストグラムや相対度数などについて学習している。ヒストグラムは分布の形はわかりやすい一方で、中央値などの指標が分かりづらい。二つ以上のデータを比較する際に、視覚的に比較がしやすい統計的な表現として、箱ひげ図がある。

箱ひげ図とは、次のように、最小値、第1四分位数、中央値(第2四分位数)、第3四分位数、最大値を箱と線(ひげ)を用いて一つの図で表したものである。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。なお、四分位数を求める方法として幾つかの方法が提案されているが、ここでは四分位数の意味を把握しやすい方法を用いる。

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。



箱ひげ図の箱で示された区間に、全てのデータのうち、真ん中に集まる約半数のデータが含まれる。この箱の横の長さを四分位範囲といい、第3四分位数から第1四分位数を引いた値で求められる。上の例では四分位範囲は $32 - 24.5 = 7.5$ である。四分位範囲はデータの散らばりの度合いを表す指標として用いられる。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。

四分位範囲や箱ひげ図を用いて批判的に考察し判断すること（イの(7)）

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。

指導に当たっては、日常の事象を題材とした問題などを取り上げ、それを解決するために必要なデータを収集し、コンピュータなどを利用してデータを整理し、四分位範囲を求めたり箱ひげ図で表したりして複数の集団のデータの傾向を比較して読み取り、その結果を基に説明するという一連の活動を経験できるようにすることが重要である。例えば、中学生の体力は以前に比べて落ちているといえるかどうかについて考える。データとしては、生徒にとっての考察のしやすさから、同じ学校の中学校2年生男子の体力テストの結果を用いることができるであろう。そこで、ハンドボール投げに焦点化し、2000年、2005年、2010年、2015年のデータから箱ひげ図（図1）を作成するなどして分布の傾向を比較して読み取り、これを基に、「中学生の体力は前に比べて落ちているといえるかどうか」について考察する。具体的には、「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

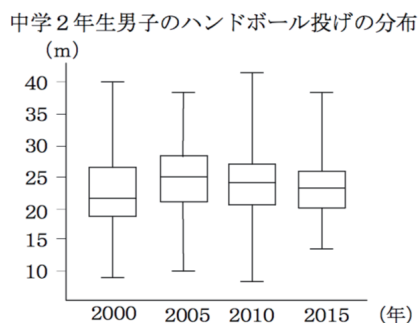


図 1

ることが考えられる。ここで、より詳しく検討するために、範囲の違いに着目して、2010 年と 2015 年に絞ってヒストグラムを作成し、これらの分布を詳しく比較することが考えられる。それにより、かけ離れた値があるかどうかなど、範囲が大きく異なる理由について検討することができる。また、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 2）を作成することで、経年変化の様子を調べることができる。さらに、「ハンドボール投げのデータだけで十分か」と批判的に考え、握力など他の体力テストのデータから箱ひげ図（図 3）やヒストグラムを作成するなどして、一層詳しい考察を加えることも考えられる。

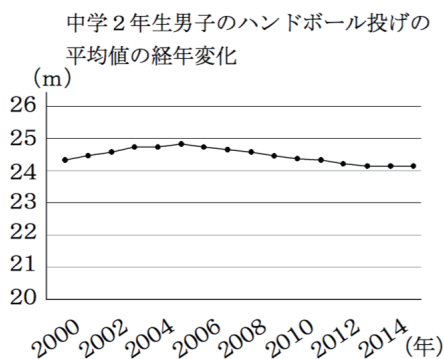


図 2

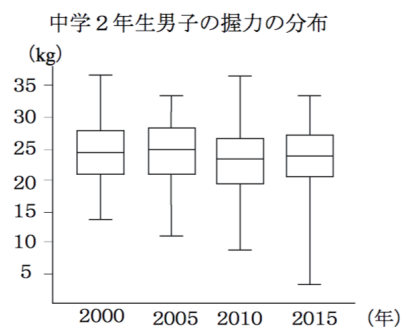


図 3

このように、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。また、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。なお、同じ学校のデータについての判断の結果を、中学生一般について考察することは、第 3 学年で学習する標本調査に関わることである。

5. 箱ひげ図と度数分布表・ヒストグラム

箱ひげ図や、中学 1 年までに学習する度数分布表・ヒストグラムにはそれぞれに特徴がある一方、すべての場合において万能な表、図ではありません。よって、調べたい事柄に合わせて、データの加工や用いる表、図を適切に判断することが大切です。

例えば、「ある中学校の、2018 年度 1 年生 100 人のハンドボール投げの記録」というデータがあるとして、このデータを加工して調べる方法として、何があるか考えてみます。

(1) 度数分布表やヒストグラムにまとめる

データを度数分布表やヒストグラムにまとめることで、その分布を調べやすくなります。傾向を正確に読み取るには、階級の幅を適切に設定することが大切です。

(2) 箱ひげ図をかくて比べる

箱ひげ図は、2つ以上のデータの分布を比較するのに適しています。よって、過去の生徒のデータと比較してみたい場合には、箱ひげ図をかくて比べるとよいでしょう。ただし、箱ひげ図は主に中央の50%のデータについての分布をみるためのものですので、情報量としては度数分布表やヒストグラムよりも少なくなることに注意が必要です。

なお、データをひとたび度数分布表やヒストグラム、箱ひげ図に加工すると、データを完全に復元することはできません。よって、(1)や(2)のような加工をした後で、再びデータの詳細について調べる場合には、最初の元となるデータに戻る必要があります。このように、データの処理では、何かを得るために加工することで何かを失うということが起こります。

6. 終わりに

四分位範囲・箱ひげ図は、その計算や表し方自体は難しいものではないため、中学生でも問題なく理解できるものです。そのため重要になってくるのは、四分位範囲・箱ひげ図を学ぶ意味をいかに理解させるかです。既習の度数分布表やヒストグラムと絡めながら、四分位範囲・箱ひげ図の特徴を捉えることで、その理解は深まるはずです。

引用・参考文献

文部科学省 (2017) 『中学校学習指導要領解説 数学編』

大島利雄 他 (2016) 『数Ⅰ／327 改訂版 数学Ⅰ』(平成28年検定) 数研出版

大島利雄 他 (2017) 『改訂版 数学Ⅰ 教授資料』 数研出版

岡部恒治 他 (2017) 『改訂版 高等学校 数学Ⅰ 教授資料』 数研出版