

# 数学 I 「データの分析」との連携

## ～データの測定から箱ひげ図まで～

神奈川県立横浜清陵総合高等学校 五十嵐 誠

数学 I に新設された単元「..データの分析」では、分布を比較表現するために、ヒストグラム（度数分布図）の他、四分位数と箱ひげ図が使われる。適切なデータ収集の実習と表計算ソフトを使うことで教科情報がこの単元をサポートし、生徒の分析力を高めることができる。生徒は自分のデータが全体の分布に参加していることから興味を持ち、その可視化と分析のために表計算ソフトの関数やグラフ機能を扱う必然性を得る。

### 1. 数学 I で新出した箱ひげ図

ほとんどの現職の数学科教員が初めて扱う「箱ひげ図」。ヒストグラムとセットで扱うことで理解が深まる。

#### 1.1 四分位数と箱ひげ図

箱ひげ図はデータの分布を度数によって 4 等分して表現した図である。データの分布を 4 分割する 3 つの値を小さい順に第 1 四分位数 ( $Q_1$ ) から第 3 四分位数 ( $Q_3$ ) とよぶ。第 2 四分位数 ( $Q_2$ ) は中央値である。

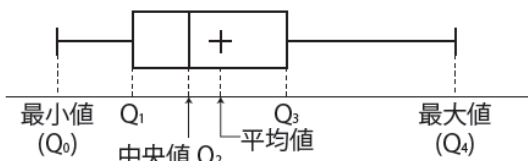


図 1 四分位数と箱ひげ図

#### 1.2 ヒストグラムと箱ひげ図

1 母集団内の分布の特性を考察するためにはヒストグラムが優れているが、複数の母集団の分布を比較する場合は、並列して比較できる箱ひげ図の方が表現しやすい。

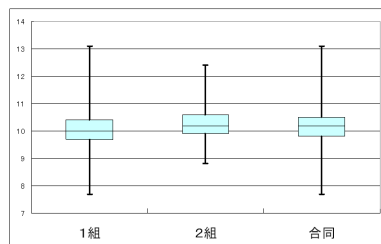
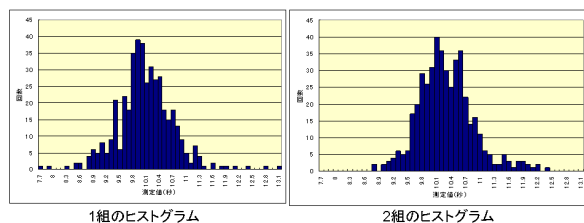


図 2 度数分布をヒストグラムと箱ひげ図で比較

### 2. 実践事例

本校では 1 年次で情報 A または情報 B を選択する。それぞれ異なる方法で得た測定値を用いてヒストグラムと箱ひげ図を作成し、考察した。

#### 2.1 2 つの方法でデータを得る

「情報 A」選択クラスでは、紙テープを目分量で 10 cm に切って測定値を分析した。並行して「情報 B」選択クラスでは、目を瞑り感覚だけで 10 秒を計時し、ストップウォッチで実際の時間を測定して分析した。表計算ソフトには Microsoft 社の Excel を用いている。

どちらも、まずは個人のデータからヒストグラムを作り、次にグループ、クラス、学年全体、と母集団を拡大していく。母集団が拡大するに従って、分布の形状に特徴が現れてくる。この変容の過程から情報に意味が生まれて分析が可能になることを理解し、特異なデータについて考察する態度を育成することができる。



図 3 テープを目分量で切り、測定値を交換する

10 秒を計測する実習では、計測を重ねるにつれ精度が上がっていくことを感じるため、さらにどのような分析をしたいか発問したところ、「2 連続で 10 秒 0 が出るようになるには、何回ぐらい測ればよいか調べたい。」などの回答を得た。仮説を立て、その実証にはどのような実験を行えば良いかなど、生徒とともに考えることも興味深い題材である。

## 2.2 分布図を読み解く

生徒が実測したデータを扱うことで、「数学」の教科書に用意されるデータでは得られないリアリティがある。特異な値が含まれたり、不自然な分布形状になったときこそ、「情報科」の教材としての価値が高くなる。

### 2.2.1 はこひげ図からの考察

図は、「情報 A」で目分量の 10cm を実測したデータの分布をクラス別に比較したものである。

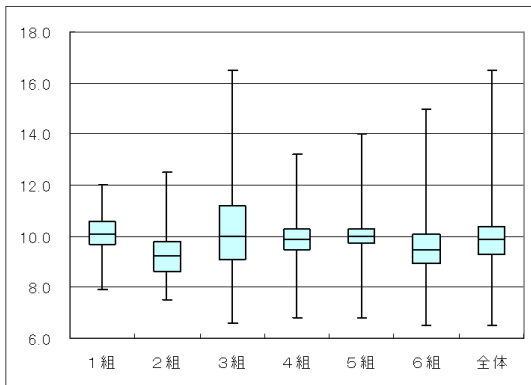


図 4 6クラスの分布を箱ひげ図で比較

この実習では、はがきを一時的に配布して、その短辺を 10cm として覚えさせた。授業を最初に行った 3 組では、分布の幅が最も大きかった。教員側の指示が甘かったためか、はがきの長辺を 10cm として覚えた生徒がいたようだ。

実際にデータを調べたところ、席が隣の特定の生徒が 15cm 付近の測定値を記録していた。

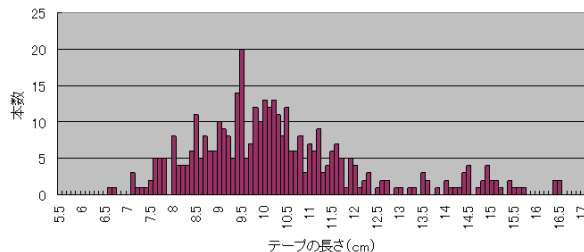


図 5 15cm 付近に測定値のピークが見られる分布

また、6 クラス中の 2 クラスは「情報 B」の選択者が約 30 名いるため、「情報 A」の受講生が 10 名ほどの少人数展開である。具体的には 1 組と 2 組であり、分布から指示がよく伝わっていることが読み取れる。

さらに、4 組・5 組・6 組の分布形状には各クラスの雰囲気が見れている。

### 2.2.2 ヒストグラムの特異な兆候の考察

次の二つの図は生徒と教員が行った実習データ

の分布である。

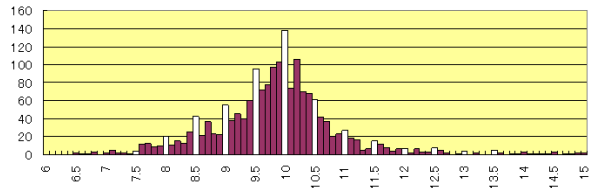


図 6 情報 A の授業より(1800 サンプル)

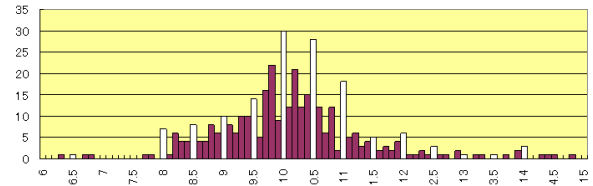


図 7 教員のワークショップより(410 サンプル)

棒グラフを白抜きにしてあるが、区切りのよい値 (5mm ごと) にピークがみられる。測定の厳密さも影響するだろうが、担当教員の間では「どちらとも読めるような場合は区切りのよい数値で読む傾向があるのでは」という仮説を立てている。

参考までに、情報 B の授業ではストップウォッチを使って感覚の 10 秒を実測したデータを分析したが、こちらの測定には主観が入らない。実際に、ヒストグラムでは上記のようなピークの特徴は表れなかった。

## 3. Excel のグラフ機能の活用

代表的な表計算ソフト Excel については例題を使って基本的な関数やグラフ機能を教えてあるが、与えられたデータを扱うだけでは活用能力が身につかない。その意味でも、自らが測定した値を可視化する作業は大切である。今回の実習事例を元にヒストグラムと箱ひげ図の作り方を紹介する。

### 3.1 ヒストグラムの作り方

Excel を使ってヒストグラムを作るために、まず、測定値が入力された範囲から、COUNTIF 関数を使って度数分布表を作成する。次に、グラフ機能を利用して、横軸が計測値の縦棒グラフを作る。この作業において、横軸の項目値の設定の工夫と、確率分布の意味を持たせるために、棒グラフ間の隙間を無くす設定が必要である。

次ページの図 8 のように、「計測値」と「回数」という項目の度数分布表から、グラフウィザードを利用してヒストグラムを作成する手順を示す。

- ①項目を含めてデータの範囲を選択
- ②グラフウィザードをポイント

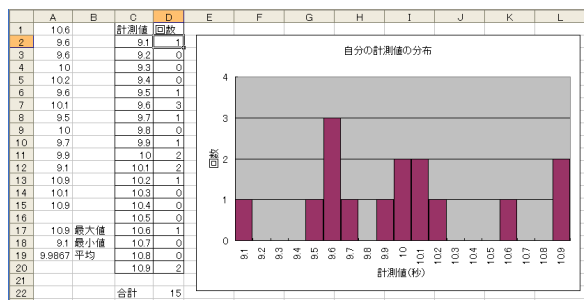


図8 度数分布表からヒストグラムを作成

- ③縦棒グラフを選び, [次へ]をポイント
- ④[系列]をポイント
- ⑤[系列]ウィンドウの[計測値(系列1)]を選択して[削除]をポイント
- ⑥[項目軸ラベルに使用]ウィンドウで計測値の値の範囲を選択して, [次へ]をポイント
- ⑦棒グラフを右クリックして[データ要素の書式設定]ウィンドウの[オプション]タブで, [棒の間隔]を0にする
- ⑧適宜, 軸の項目やグラフのタイトル等を設定

### 3.2 箱ひげ図の作り方

計測値のデータ範囲を指定して, QUARTILE 関数で  $Q_0$  から  $Q_4$  の値を求める。

表1 四分位数と求める関数

値	値の意味	求める関数
$Q_4$	最大値	=QUARTILE (範囲, 4)
$Q_3$	第3四分位数	=QUARTILE (範囲, 3)
$Q_2$	第2四分位数	=QUARTILE (範囲, 2)
$Q_1$	第1四分位数	=QUARTILE (範囲, 1)
$Q_0$	最小値	=QUARTILE (範囲, 0)

説明のため, 箱ひげ図の4つの要素を便宜的に「上ひげ」, 「上箱」, 「下箱」, 「下ひげ」, 下箱から下の空白部分を「土台」と呼ぶこととする。

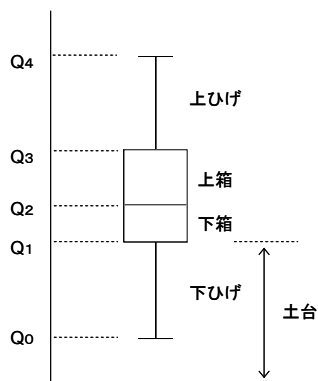


図9 箱ひげ図の各要素の便宜的名称

$Q_4$  から  $Q_0$  の値の表から, 「上ひげ」「上箱」「下箱」「下ひげ」「土台」の要素にあたる量を計算し, 箱ひげ図の形に整形していく。考え方は, 「土台」, 「下箱」「上箱」の積み上げ棒グラフを作って「土台」の部分为非表示にし, 「上ひげ」と「下ひげ」の部分をそれぞれ「上箱」の正方向と「土台」の負方向の誤差表示機能で表現する。

表2 箱ひげ図の要素を表す量と表示方法

要素	量	表示方法
上ひげ	$Q_4 - Q_3$	上箱の誤差(正)として表示
上箱	$Q_3 - Q_2$	積み上げ棒グラフ
下箱	$Q_2 - Q_1$	積み上げ棒グラフ
下ひげ	$Q_1 - Q_0$	土台の誤差(負)として表示
土台	$Q_1$	積み上げ棒グラフで非表示

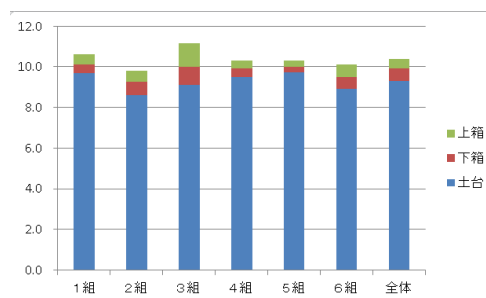
次の図では, 6クラスの計測値から  $Q_4 \sim Q_0$  を求めた表を作り, これを元に箱ひげ図の要素の量を下の表に集約している。積み上げ棒グラフを作りやすくするために要素の順を工夫してある。

	A	B	C	D	E	F	G	H	I
18									
19			1組	2組	3組	4組	5組	6組	全体
20		$Q_4$	12.0	12.5	16.5	13.2	14.0	15.0	16.5
21		$Q_3$	10.6	9.8	11.2	10.3	10.3	10.1	10.4
22		$Q_2$	10.1	9.3	10.0	9.9	10.0	9.5	9.9
23		$Q_1$	9.7	8.6	9.1	9.5	9.7	8.9	9.3
24		$Q_0$	7.9	7.5	6.6	6.8	6.8	6.5	6.5
25									
26			1組	2組	3組	4組	5組	6組	全体
27		土台	9.7	8.6	9.1	9.5	9.7	8.9	9.3
28		下箱	0.4	0.7	0.9	0.4	0.3	0.6	0.6
29		上箱	0.5	0.6	1.2	0.4	0.3	0.6	0.5
30		上ひげ	1.4	2.7	5.3	2.9	3.7	4.9	6.1
31		下ひげ	1.8	1.1	2.5	2.7	2.9	2.4	2.8

図10 6クラスの四分位数と箱ひげ図の各要素

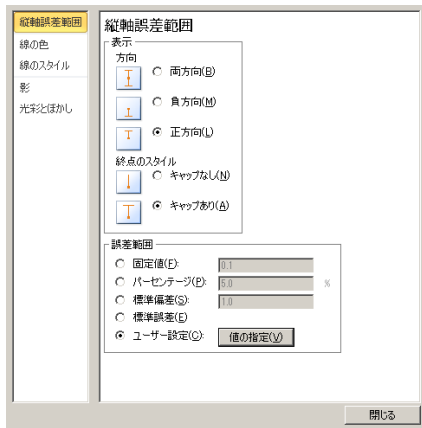
操作の例として, この図の下の表(26行~31行)をもとに箱ひげ図を作成する手順を示す。

- ①B26:I29 を選択し, グラフ機能を利用して「積み上げ縦棒グラフ」を作成する

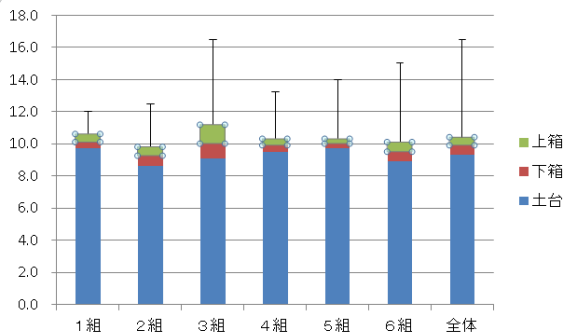


②棒グラフの「上箱」の部分を選択し、[レイアウト]→[誤差範囲]→[その他の誤差範囲オプション]をポイント

③[誤差範囲の書式設定]ウィンドウの[誤差範囲選択]タブで、[方向]を[正方向]に、[誤差範囲]を[ユーザ設定]にして[値の設定]をポイント



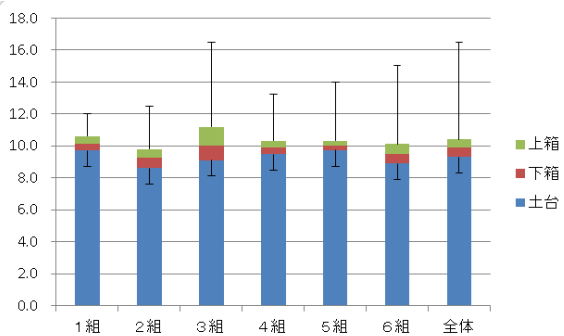
④[ユーザ設定の誤差範囲]ウィンドウで[正の誤差範囲の値]に、「上ひげ」の範囲「C30 : I30」を設定する



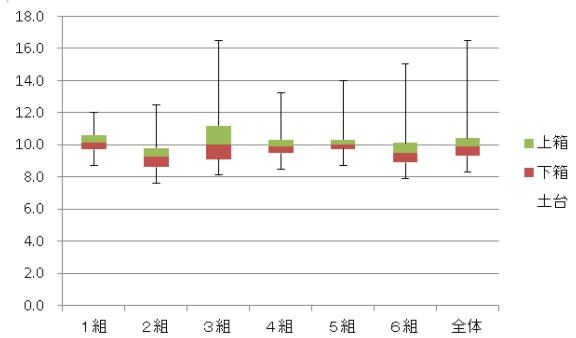
⑤棒グラフの「土台」の部分を選択し、[レイアウト]→[誤差範囲]→[その他の誤差範囲オプション]をポイント

⑥[誤差範囲の書式設定]ウィンドウの[誤差範囲選択]タブで、[方向]を[負方向]に、[誤差範囲]を[ユーザ設定]にして[値の設定]をポイント

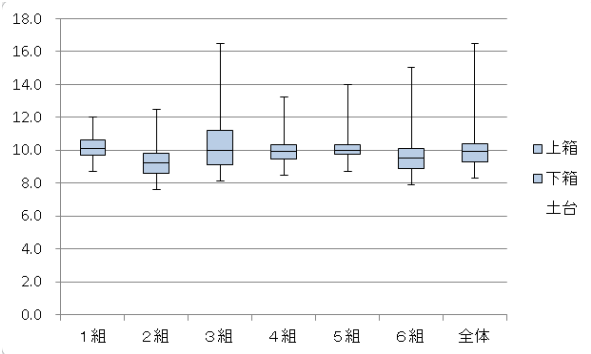
⑦[ユーザ設定の誤差範囲]ウィンドウで[負の誤差範囲の値]に、「下ひげ」の範囲「C31 : I31」を設定する



⑧「土台」の部分の塗りつぶしの色を「塗りつぶしなし」、枠線の色を「線なし」とする



⑨「上箱」と「下箱」の部分の塗りつぶしの色を揃える



⑩縦軸の[軸のオプション]で、[最小値]の値を調整して、分布を比較しやすくする

⑪[凡例]を削除し、表のタイトルを設定する

#### 4. まとめ

未だに「操作ばかりのパソコン教室」と評される授業を行っている学校があると聞く。数学に統計が導入されたことを機に、自らが計測したデータを用いた学習活動、教科を超えたアクティブラーニングに脱皮することが可能であると考えられる。

一例として実践事例を報告させていただいたが、来年度には、全国で多くの事例が報告されることと期待している。

#### 参考文献

Miyake, N., & Shirouzu, H. (2005, June). Design and use of smart tasks in collaborative classrooms. Poster session presented at the meeting of the Computer Supported Collaborative Learning, Taipei, Taiwan.

<http://coref.u-tokyo.ac.jp/nmiyake/list/material/pdf/050601csclSmartTask.pdf>