

統計ソフトウェア Fathom を用いた統計教育 — 統計教育用の開発された数学ソフトウェアの検討 —

東京理科大学大学・理学部 清水 克彦 (Simizu katuhiko)
Department of Science,
Tokyo University of Science
東京理科大学大学院・科学教育研究科 兼子 航 (Wataru Kaneko)
Graduate School of Mathematics & Science Education,
Tokyo University of Science

1 はじめに

高等学校における新学習指導要領が来年から本格的に実施されることとなり、特に数学におけるデータ分析の指導をどのように行うかという議論が活発に行われている。今回の改定では、グラフをまとめたり、代表値を求めるといった記述統計の色合いが濃かった統計教育のみだけでなく、データからその集団の傾向を読み取ったり考察を行う所までが含まれている。本稿においては、これらを短時間で行える可能性のある統計教育用ソフトウェア Fathom を紹介し、またそれを利用した教材化について述べていく。

2 これからの統計教育と統計ソフトウェア

急速に発展しつつある情報化社会において、あらゆる場面で氾濫するデータのなかから、必要なデータを選択し、データを正しく読み取ったり、整理したりする能力はこれまで以上に、学生にとってますます必要になってきている。統計教育の重要性は、これまで以上に高くなってきている時代に入っていると言える。その現れの1つが次期学習指導要領の数学における「資料の活用」の単元の導入であろう。これからの統計教育においては、統計的な見方・考え方、統計的な数値に対する感覚やセンスがいっそう大切になってくる。

それに対して、日本の算数・数学科における統計教育はかならずしもうまくいったとはいえない。これまでの算数・数学科における統計教育の反省点・問題点を挙げると、(1) 理論重視の指導：まず、統計の数学的理論を教えることを目的としてしまい。児童・生徒の統計的な見方・考え方を育てることができなかったこと、(2) 技術先行の指導：グラフの描き方や平均の求め方などの技能を優先し、統計的な見方・考え方を軽視したことを指摘することができる。同じく、確率の指導においても数学的確率を優先し、統計的な確率を軽視してきており、平均の指導でもその意味よりも計算方法が指導されてきていた。

先進的な試みを見てみると、NCTM（全米数学教師協議会）では統計の分野を Data Analysis と呼んでおり、そのスタンダードでは、「データの分析がいろいろな面、たとえ

ばビジネス・政治（世論調査）・経済（消費者調査）・研究などで使われている。最近ではデータの分析に、有力で身近になってきたテクノロジーが使われており、この影響により我々の能力に大きな変化が起こっている。」と指摘している。さらに、NCTMでは、「データの分析」という分野を「変化」や「不確実さ」を知的に扱うことができる能力を育成するための分野と位置づけ、「もし、カリキュラムにこの分野が入らなければ、このような能力は発達しなくなってしまうだろう。」と数学教育で統計教育を行うことを主張している。

このような背景には、統計の知識や能力の育成は、Literacyつまり基礎能力として位置を持っていると捉えるようになってきている点を挙げることができる。それを現在では、Quantitative Literacy（数量的なリテラシー）と呼び、現代における読み・書き・そろばんとして、統計的な見方・考え方の育成が位置づけられている。このような目的のもと、児童・生徒の統計的なプロセスでは、データを大切にして、理論や技術だけを優先しないこととし、「実際のデータは、統計の学習にその正当性と刺激の両方の面を与えてくれる」という立場にこれからの統計教育は立っている。すなわち、データを前にして、児童・生徒がこれから取る統計的なプロセスについて議論や評価をして、様々な角度から見るようにし、統計的な見方・考え方の基礎をなすことを目指しているのである。

そのとき、統計教育における統計ソフトウェアの活用は次のような利点を持っている。

- すばやく統計的な処理ができる
- 調べる資料を検索できる（Webなど）
- 統計データを集める場となる（Webなど）
- 多様な統計表現を可能にする
- 数の遊び場を提供する（シミュレーション）
- 手順を代行するだけの使用ではなく、理解を支援するために使用をすることが効果的にできる

そのための統計教育用ソフトウェアとして全米科学財団の資金援助を受けて開発されたものが本稿で紹介する Fathom である。

3 統計教育用ソフトウェア Fathom

この Fathom は高度なデータグラフ化ソフトウェアであり、データ分析を行い、数学と統計学的な概念を分析し、学習するための詳細な機能を学生に提供するものである。開発したのは、Key Curriculum Press 社（以下 KCP 社）で、この会社は 1971 年に設立、数学や理科教育における書籍や教材の開発、教員に対する指導支援を行っている。

windows や Machintosh 版を販売しており、つい最近では、TI-nspire CX のカラーグラフ電卓にも搭載されている。

KCP 社によれば、Fathom は具体的に以下のような分野の内容で使うことが考えられている。([1])

代数学 1

グラフのデータ, 傾きの意味を 2 変量データを使用して理解させる.

代数学 2

データを変換したり, 関数あてはめ (回帰分析), 対数の使用や予測を行う.

幾何学

幾何学的確率の学習.

微分積分学の前学習と微積分の学習

複雑な機能を持つデータモデリングや微分係数, 定積分の計算方法を学ぶ.

本格的な統計と AP¹®統計

データの探索, 確率とサンプリングシミュレーション, 統計的推測の実行 (カイニ乗, t 検定, 一元配置分散分析 (ANOVA の方法), および複数の線形回帰を含む) を行うことができる.

自然科学や社会科学

データを扱う自然科学や社会科学, その他の科目でも Fathom を使うことができる.

またこれらの考えをもとに, Fathom には標準で 300 以上ものサンプルファイルが含まれており, 教師もすぐに具体例を用いて Fathom の操作に慣れることができる.

sectionFathom の特徴 Fathom はドラッグ&ドロップを中心とした直観的な操作が基本となっている. 難しいプログラミングも必要なく, 扱うことができる. また様々なところに表れるアニメーションがより生徒の理解をうながしている.(例としてはデータからサンプルをとる場面など)

以下の図のように Fathom は表やグラフ, データの集合などを自由に配置でき, 1 つの枠のなかで複数のものを表示することができる.

3.1 データの取得

Fathom のデータ取得は複数ある. 用途に応じて様々な形でデータを集めることができるのもこの Fathom の特徴である. データを Fathom に取り込むには主に以下の 4 つの方法があるが, ここでは 2 つ目のインターネット特に米国国勢調査のデータを取り込んでみる.

1. 直接テーブルに入力する.
2. インターネット上のから取り込む (URL を指定したり, ファイルを指定, 米国国勢調査によるデータをダウンロードできる.)

¹アメリカには, アドバンスド・プレースメント (AP) と呼ばれる, ハイスchool の特に学力優秀な生徒を対象に, 在学中に大学レベルの学習機会を与え, 所定の試験に合格すれば大学の単位として認定する取組がある (指導は, ハイスchool の教員が行う) (文部科学省)

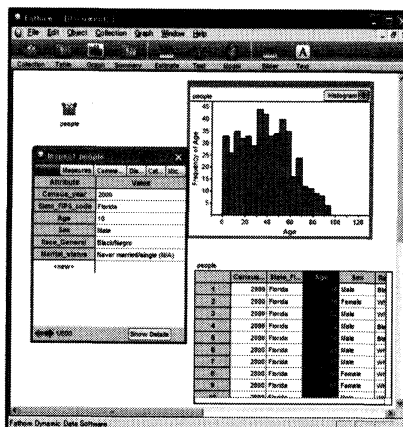


図 1: Fathom ドキュメントウィンドウ

3. エクセルやテキストデータから読み込む
4. バーニアセンサー² から直接 Fathom へデータを取り込む。

[手順]

1. Fathom ドキュメントウィンドウを開き、左上の *File* をクリックする。
2. *import* → *U.S. Census Data* を選ぶ。

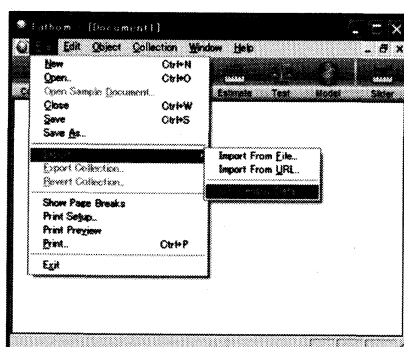


図 2: U.S. Census Data の場所

3. *Choosing cases* から欲しいデータの年や州、都市、人口か世帯か、サンプルサイズを決める。
4. *Attributes* データの性質（人々のデータから年齢や性別、人種や収入、など）の一部を取り出すことも可能

²溶存酸素濃度計、pH センサー、ステンレス温度計、伝導計、気圧計と湿度計他さまざまな機能を備えた測定器 (© 2011 Vernier Software & Technology, LLC)

5. 選択し終わったら, 右下の *Download Data* をクリック.

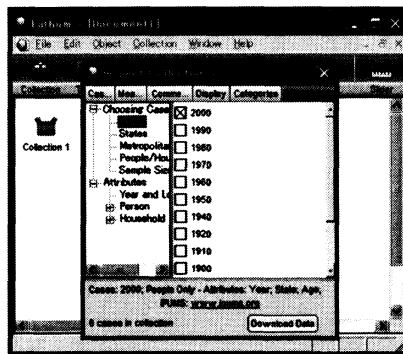


図 3: データをダウンロードする画面

成功すると下の図3のように, 箱に黄色い球 (データを表している) が入っているようになる.

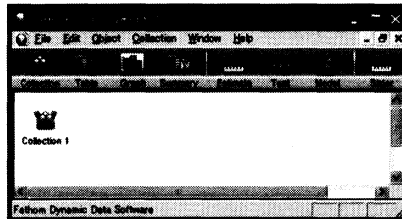


図 4: Fathom ドキュメントウインドウ

3.2 グラフ描画

Fathom のグラフ描画の特徴は以下3つある.

1. 最大3次元データまで扱えるが, 空間表示にせず, 色の違いでその3次元を表すことである.
2. ヒストグラム・棒グラフ・箱ひげ図・ドットプロット・散布図・パーセントイルドットといった豊富なグラフ表示である. しかも複数の異なるグラフを同時に同一データから表示させることができ, 簡単に個々のグラフの特徴を利用してデータ分析を扱うことができる.
3. 動的ソフトウェアの所以ともいえるが, グラフを動かしたり, データの所在を判明させることができる点である.

[手順]

1. *Table* を選択して、ドラッグ&ドロップして、中にデータを入れておく。今回は以下のようなデータを入力しておいた³

value	1	2	3	4	5	6	7	8	9
度数	1	2	3	4	5	4	3	2	1

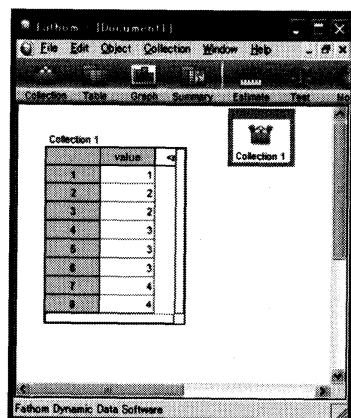


図 5: グラフの作成 1

2. *Graph* をドラッグ&ドロップで下に持ってくる。データの *value* という名前の部分にマウスを当てると手の形になるので、そのままそれをグラフの *x* 軸の部分にドラッグ&ドロップでもって送る。

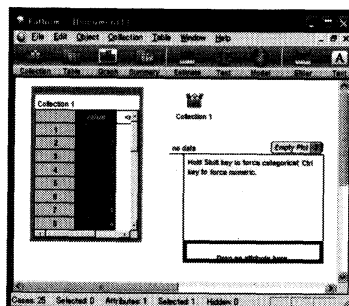


図 6: グラフの作成 2

³実際はデータを一つ一つ入力し、度数で入力していないがデータの数が多いので度数で表示している。

これでグラフをつくることができる。デフォルト表示が *Dot Plot* なので下の図のように表示されている。

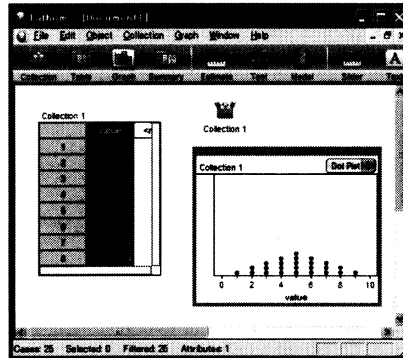


図 7: グラフの作成 3

3.3 データの分析

Fathom に内蔵されているプログラムを用いて、標本分布の作成や無作為抽出によるデータの取り出しが容易にできたり、100 以上の定義された関数を用いてさまざまなデータ抽出をすることが可能である。ここでは、サンプリングを実際に行う。

[手順]

1. データのある *Collection* を右クリックすると、*Sample Cases* があるのでそれを選択する。これだけでサンプリングが可能。成功すると今度は箱の中に青い球が入っているのがみえる。

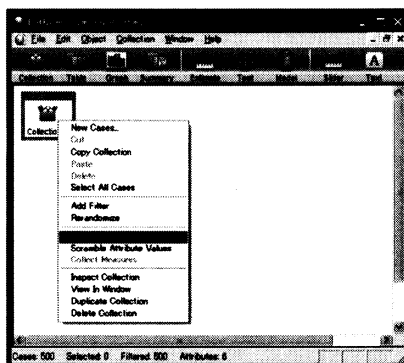


図 8: *Sample Cases* を選択

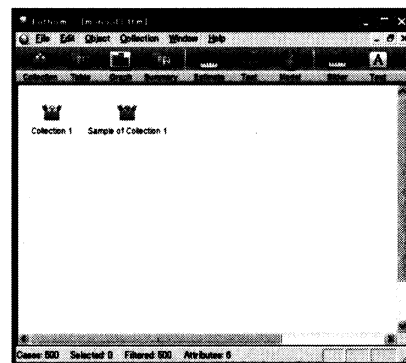


図 9: サンプリング成功

2. このデータの *Table* を作成すると、先ほどの *Collection* の一部のデータが入っているのがわかる。

Case	Year	State	Age	Sex	Race	Marital
1	2000	Pennsylv.	35	Male	White	Married
2	2000	Iowa	85	Male	White	Widowed
3	2000	Florida	26	Male	Other ra.	Never m.
4	2000	Texas	50	Female	Other ra.	Married
5	2000	New Jer.	58	Male	White	Married
6	2000	Massachus.	29	Male	White	Never m.
7	2000	California	43	Female	White	Separated
8	2000	North Ca.	38	Male	White	Divorced
9	2000	Michigan	63	Female	White	Widowed
10	2000	New Jer.	61	Male	White	Married

図 10: サンプルングデータ

2. また、このサンプルをダブルクリックすると、*Inspect Sample of Collection* のウィンドウが出てくる。ここではサンプルングに関する設定を行うことができ、サンプル数の調整もここで行うことができる。*Sample more Cases* ボタンを押せば設定した条件でサンプルングを行うことができる。

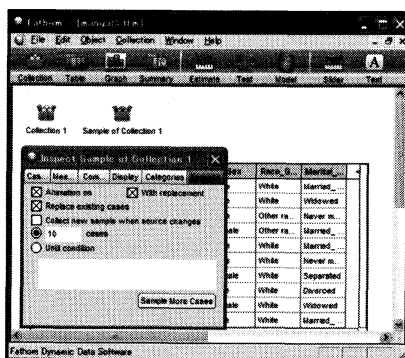


図 11: サンプルングの設定

4 Fathom の実践例

4.1 相関係数の考察

以下のようなデータ集合を用意する。

x	3	6	1	9	2	7	3	8	5	6
y	9	1	7	2	8	3	4	9	5	8

表 1: データ集合

このデータを Fathom でグラフ化する。グラフは相関係数を調べたいので、*Scatter Plot* を使用。以下がその図。

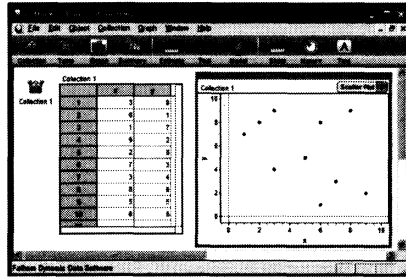


図 12: データと相関図

ここで、Graphの右隣りにある、Summaryを使用。Summaryをドラッグしてくると、空の表がある。この上に x 、左に y のデータ集合(逆でもよい)をドラッグで、持ってくると、自動的に2つのデータの相関係数が計算される。

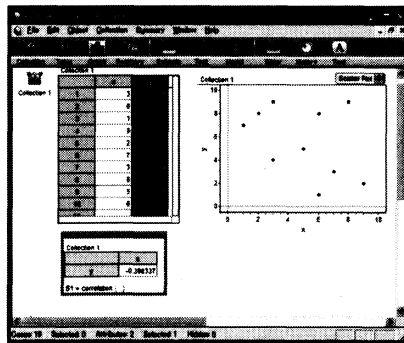


図 13: 相関係数を表示

Fathomはグラフ上のデータを動かせるので、グラフ上のドット(データ)を動かして、グラフの分布と相関係数の関係を動的にみることができる。(下図はさきほどのデータから右上がりにデータを移した時のグラフと相関係数の値)

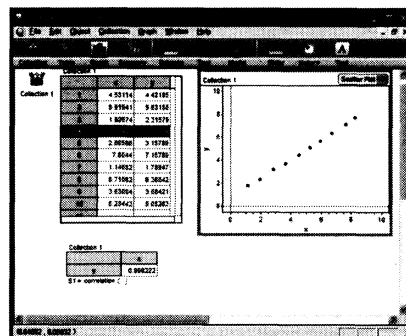


図 14: 相関係数を変化させる

4.2 スライダーを用いたサンプリングの例

今回は Sample Documents → Learning Guide Starters → Planets.ftm のファイルを利用. このファイルには太陽系のデータが含まれている.

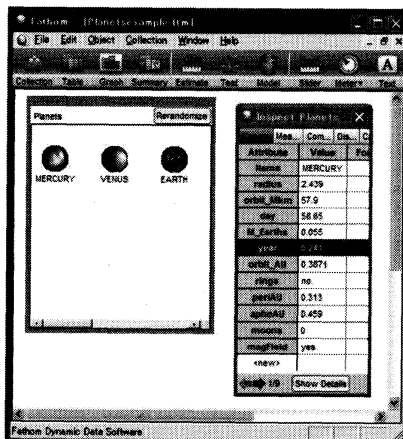


図 15: 惑星のデータ

このデータから, $orbit_AU^4$ を横軸, $year$ を縦軸としたグラフ (ScatterPlot) で作成.

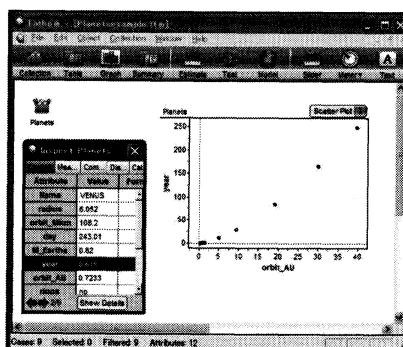


図 16: グラフ化

ここで, 上部から *Slider* をドラッグし, *Slider* の数値の変数名を **exponention** とする. また, グラフで右クリック → *Plotfunction* で $orbit_AU^{\text{exponention}}$ と記入する.

⁴軌道 (単位: 天文単位)

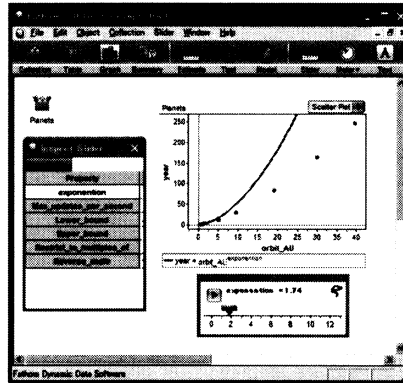


図 17: 関数をあてはめる

データの規則を既存の関数を用いて探することができる。

4.3 検定の例 -データの平均値の検定-

異なる二つの母集団の平均値の検定を行うには、普通t検定を用いる。ここでは、Fathom内のサンプルファイルを用いて、検定がどう可視化されているのかをみていく。下のデータはすでに Fathom で用意されていたものである。

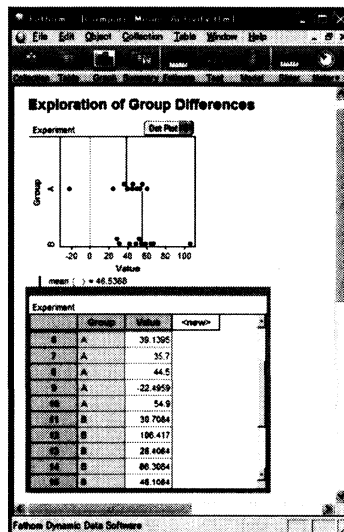


図 18: データ集合

上の *Test* をドラッグしてくる。そして右上の欄を、今回は”平均値の検定”なので *CompareMeans* を選択する。そして、*FirstAttribute* には *Value* を、*SecondAttribute* には *Group* をそれぞれグラフからドラッグしてくる。

これで検定結果がでてくる。

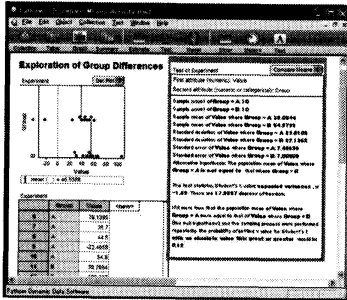


図 19: 検定結果その 1

表 2: 検定の結果

データの数	A	10
	B	10
平均値	A	38.0944
	B	54.9793
標準偏差	A	23.6108
	B	22.1362
標準誤差	A	7.46639
	B	7.00009
不偏分散	-1.65	
自由度	17.9257	

右クリックで、*Verbose* を選択すると、検定結果が見やすい形で現れる、ここには p 値もでてくる。

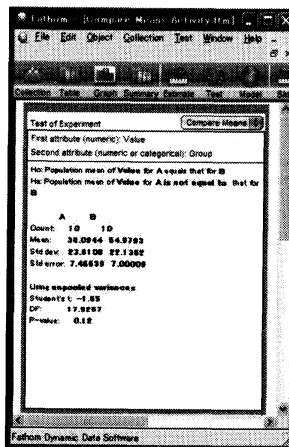


図 20: 検定結果その 2

さらに右クリックで、*ShowTestStatisticDistribution* を選択すると、検定結果を目で確認することができる。さらにこのとき、グラフの値を変化させると、p 値が変わり、グラフの青色の領域も変化させることができる。

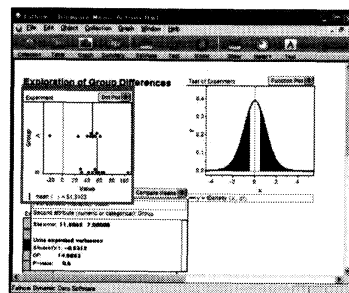


図 21: 検定結果その 3

5 Fathomを使った教材例

5.1 教材化の必要性

高校の内容で新たに必修となった統計内容、数学のデータの分析であるが、今まで統計が教えられてこなかった経緯もあり、参考となる教材が他の分野よりも少ない。また、来年から新学習指導要領が高校生は実施され、検定教科書も作られてきたが、紙の上のみでデータをいじる、というのは生徒の理解を得るといふ点ではなかなか難しい。身近なデータ(気温や身長など)を利用して、要約統計量や箱ひげ図の概念を理解させようと工夫がなされているものの、やはり計算量やグラフをかくということに時間を多く割いてしまい、本来の目標である”データを分析してそこからデータの傾向を読み取る”という所までいくのは、時間数も限られていることから難しい。そこで、新学習指導要領やその解説に「適度にコンピュータを用いて⁵⁾」と書いてあるように、今回紹介したFathomを利用して、主にヒストグラムと箱ひげ図のグラフの読み取りを行わせていく。ただし、ここでは具体的なデータを用いた読み取りではなく、ある程度の代表的な分布になるよう操作されたデータを用いて、分布の読み取り(要約統計量を調べたり、それらと分布の傾向、箱ひげ図との関係など)を行う。教材例の条件としては、コンピュータやFathomの特徴を生かし、「紙ではできない、しかも時間をかけない内容」を挙げた。

5.2 教材の目的と準備

教材の目的

ヒストグラムの分布の形状の特色と箱ひげ図の概念を、要約統計量と比較しながら理解させる。

以下のデータ集合を与えておく。

表 3: モデルデータ (各データを度数で表示)

データの値	V1	V2	V3	V4
1	1	4	1	1
2	1	3	2	2
3	2	2	3	4
4	3	1	4	8
5	4	2	3	16
6	5	3	2	25
7	4	4	1	36

⁵⁾中学校：第1学年、D資料の活用(1)(学習指導要領) 高等学校：第1部 数学 第2章 各科目 第1節 数学 3. 内容と内容の取扱い(4) データの分析(学習指導要領解説)

5.3 教材例 1: ヒストグラムと要約統計量の特徴

1. 各データを Fathom に入力し、ヒストグラムをかかせる。
2. *Graph* ウィンドウ上で、右クリックして、*Plot value* を選択。=のあとに関数を入力して平均値、中央値、第1四分位数、第3四分位数をグラフ上に表示させる。

	関数	意味
平均値	$\text{mean}(X)$	データ集合 X の平均を求める。
中央値	$\text{median}(X)$	データ集合 X の中央値を求める。
第1四分位数	$Q1(X)$	データ集合 X の第1四分位数を求める。
第3四分位数	$Q3(X)$	データ集合 X の第3四分位数を求める。

表 4: 関数とその意味

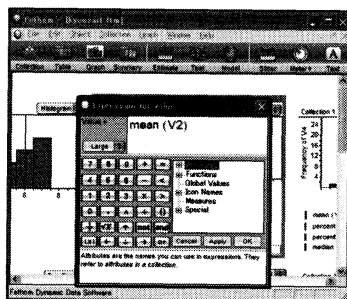


図 22: グラフに平均値を表示させる

3. 「中央値や平均値はどういったグラフの分布のときに重なる傾向にあるか。」
(グラフが対称なとき中央値と平均値は重なることを理解させたい。)
4. 「重ならない場合はどういった場所にでてくるか。各々のデータに9というデータを追加して増やしていったときに平均値や中央値がどう変化しているかに着目して述べなさい。」
(グラフが対称でない場合、最頻値のあたりに平均値や中央値がでてくること。またデータ数が極端に多くなると、中央値よりも平均値がそれにひきずられて移動しやすいということを理解させたい。)

5.4 教材例 2: 同じデータの散ばりと箱ひげ図, ヒストグラム

1. 中央値, 最小値, 最大値, 第1四分位数, 第3四分位数から箱ひげ図をかく。
2. 各 *Graph* をコピーして、右上のグラフの種類を *Box Plot* にかえて箱ひげ図にし、自分たちの書いた箱ひげ図と一致しているか確認する。

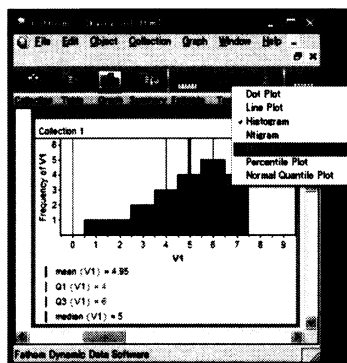


図 23: グラフの変更

3. 「データの散ばりを箱ひげ図の箱の端を動かしながら、どういふときに散ばりが大きくなるか、小さくなるか述べてよ。」
 (箱ひげ図の第1四分位数や第3四分位数を動かすと、ヒストグラムではデータの移動が起こる。データがあつまっているほど箱が小さくなり、データが散ばると箱が大きくなることを理解させたい。)

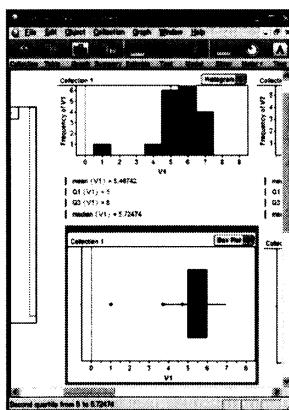


図 24: 箱ひげ図の幅を変えてみる

4. 先ほどのデータを少し変えて、以下のようにする。(太字がデータを増やしたところ)

データの値	V1	V2	V3	V4
1	5	4	5	1
2	1	3	2	2
3	2	2	3	4
4	3	1	4	8
5	4	2	3	16
6	5	3	2	36
7	4	4	5	36

表 5: モデルデータ 2(各データを度数で表示)

5. 「V4に比べて、V1,V3の箱ひげ図の箱の幅はどうなったか？」
(最頻値が2つあり、しかもそれらが離れていると、箱の幅は大きく広がる。このように、大きいデータの山が二つあるような分布では1つのデータ集合で箱ひげ図を描くと、散ばりがわかりにくくなる.)
6. 「V2とV3のデータの散ばりについてヒストグラムと比較して述べよ。」
(異なるヒストグラムがあっても、箱ひげ図がほぼ一致してしまうデータ集合が存在してしまう。散ばりをしめす箱ひげ図もよく吟味する必要がある.)

参考文献

- [1] Key Curriculum Press
<http://www.keypress.com/x5656.xml>(2011年8月8日 確認).
- [2] Vernier
<http://www.vernier.com/>(2011年8月9日 確認).
- [3] 首都大学東京産学公連携センター
http://www.tokyo-sangaku.jp/research/researchs/tresearch_view.php3?index=264(2011年8月9日 確認)
- [4] 阿部 大輔『講演会「統計教育における高大連携-米国およびUCLAにおける事例から」参加レポート』
<http://www.cret.or.jp/j/about/bumon/bumon1/report.html>(2011年8月9日 確認)
- [5] 総務省 統計局『中学校学習指導要領解説 数学 統計関係部分抜粋』
<http://www.stat.go.jp/teacher/dl/pdf/c3index/guideline/middle/math.pdf> (2011年/8月14日確認)

- [6] 総務省 統計局『高等学校学習指導要領解説 数学 統計関係部分抜粋』
<http://www.stat.go.jp/teacher/dl/pdf/c3index/guideline/high/math.pdf>
(2011年/8月14日確認)
- [7] 第一学習社 (2011)『高等学校 数学 』