

# 「データの分析」学習会を開催して — 数教協ちば高校サークルの活動より —

松戸南高等学校 浅間 基秀\*

## 1 はじめに

数教協ちば高校サークル（以下、サークル）では、新課程の数学Ⅰで新しく導入された統計の内容「データの分析」についての学習会を、2回にわたって開催しました。

サークルは学期末の考査期間ごとに年3回、授業プリントや考査問題を中心に授業実践の交流をしています。少ないときには5人ほどですが、多いときは10人ほどが集まります。5年前までは固定されたメンバーで細々と集まっていたのですが、採用が急回復して若い先生方が高校現場に戻ってくるのに合わせ、新しいメンバーも増えてきました。

## 2 学習会開催に向けて

2011年12月のサークルで、メンバーのよしださんが話題提供を3つされました。1つめは2010年度のセンター試験の問題から、工業高校の生徒が数学Ⅱの代わりに受験できる「工業数理基礎（第2問）」の紹介でした。工業高校では数学の先生ではなく、工業の先生が教えることが多いようですが、実用的な問題・題材が多いとのことでした。

ある地域で、ある期間に、マグニチュード  $M$  以上の地震が  $N$  回発生したとすると、 $M$  と  $N$  の間に以下のような関係が見られることが知られている。

$$\log_{10} N = a - bM \quad (1)$$

$a$  と  $b$  は地域と期間によって異なる定数である。自然界の統計的なデータには式(1)のように対数を含む関係が多く現れるが、この場合は図3に示すような片対数グラフを用いてデータを表すとよい。  
(図の引用は省略)

2つめは2010年度の名古屋工業大学の英語の入試問題で、内容としては小学校高学年から中学校の一次方程式くらいの数学です。単科大学ということもあって、思い切った出題ができたのではという話でした。

\*千葉県数学教育協議会事務局長

3つめは新課程で出てくる「相関係数の散布図」で、グラフソフト GRAPES を利用して縦軸と横軸共に、先に標準偏差で割って standardize (基準化 or 標準化) するもので、相関の比較が容易だということでした\*1。

おまけとして、パソコンとプロジェクタがあったので、よしださんの USB 内のとっておきの教材もいくつか紹介していただきました。そのうちの1つが「箱ひげ図」で、データの並べ替えを工夫する「幹葉図」も紹介され、時間があれば参加者全員で実際に作業したいね、と参加者全員で残念がった次第です。

次の2012年3月に開かれた高校サークルを前に、県南の方から「千葉市でやっていただけると参加できるのだが」という問い合わせが私にあり、サークルでその紹介をしました。千葉市でやるならば是非、『データの分析』をよしださんをお願いしたい、ということになりました。10月28日に蘇我勤労市民プラザで会場を確保し、よしださんのツテもあって日本統計学会の統計教育委員会と統計教育分科会の後援も得て、学習会「新課程数学I『データの分析』の授業－基礎講座と模擬授業」の開催にこぎ着けました。

### 3 「データの分析」学習会報告

講師紹介 よしだはじめ

1988年より河合塾コスモコース講師

1981～86年にはビジネス系専門学校でコンピュータ関連科目、統計学などを担当。コンピュータ SE, コンサルタントとしても活動している。

著書:『数学の教室とコンピュータ』(編著),『あそんでまなんで電卓パズル』(訳)他。

今は無き三省堂『高校数学教科書』の執筆メンバー

千葉県数学教育協議会委員長

サークルのメンバー以外にも私学や県外、大学院生など計23名が参加しました。以下、よしださんの「レジュメ」を基に、講演内容に簡単に触れていきます。

#### I部 基礎講座 「データの分析」の授業の視点と要点

##### ● 序

##### (1) 数学と統計学の違い

→ 統計学は応用科学のひとつ。

##### (2) 統計の対象は集団

→ 集団同士の比較をする。

→ 集団の中の個の位置づけを知る。

##### (3) 統計で求める量は「目安」の値

→ 同じ目的で複数の指標が使われる(例・日経平均と TOPIX は株価を表すが、「どちらが正しい」というものではなく、どちらも意味があつて使われるもの。「使われる」ということが応用科学の1つの面である。)

\*1[http://www.ne.jp/asahi/math.edu/ami/myprog/kijunka\\_kara\\_soukan2.pdf](http://www.ne.jp/asahi/math.edu/ami/myprog/kijunka_kara_soukan2.pdf)

- (4) 小・中・高の統計教材を統一的視点で  
→ 教える側は一連の流れを頭に入れておくべき。

● 箱ひげ図の背景

- (5) いつ、だれが考えた?

→ John Wilder Tukey (1660年代後半に考案, 1977年の著書 Exploratory Data Analysis に記載)

→ ベル研究所で、通信に関する大量のデータ処理に関わる中で編み出す。従来の平均・標準偏差だけでは捉えきれない・処理しきれないデータに対応。

- (6) なぜ、こんな方法をとる?

→ 標準的な確率分布ありきではなく、データそれ自身の中にあるパターンを検証。

→ データを表示するツールを「箱ひげ図」だけではなく「幹葉図」とともに開発。

→ 「箱ひげ図の本当の名称は?」との疑問もあるようだ (Tukey の狙い通り=ショックを与える名称)。

- (7) 海外の教育では? 遅れた日本!

→ 急速に広まる。国際会議でも、統計の授業で子どもたちの作った作品として見かけるように。

→ 日本では前回の指導要領で中学校から統計が消える。高校は残るも、選択のため実質消滅。60年代後半に編み出され、80~90年代に海外では広まって、標準的なものになっている間、日本では統計教育の空白期間に。

- (8) 実際に使われているのか?

→ アメリカの役所の報告書など

→ 日本では医薬分野で普及。一般には、全国学力調査の記事 (朝日新聞 2007.10.25) が最初で最後?

→ 中学校数学教科書 (大日本図書 1年 p.289, 発展のページ) でも「箱ひげ図」が。

● 教科書を比較して

- (9) 統一されていること (しまったこと)

→ 用語「四分位数」とその求め方。

Quantile の訳語

- i. 四分位数 (今回の教科書はこれで統一)

中位数の語尾と一致

- ii. 四分位値

中央値, 最大値などの語尾と一致

- iii. 四分位点

図形的に捉え, 内分点の語尾と一致

- iv. 四分位

語尾が統一されていないので, 共通する「語幹」だけで標記。

→ (よしださんの強い意見) 教科書の同じ文中に, 中央値と四分位数が並ぶのは私の言語感覚では不可解, 悪文としかいいようがない。私は「四分位値」を使います。

- (10) 統一されていないこと

- 「ひげ」の両端の形状は教科書では3パターン。字体の違いのようなもの。
- 箱ひげ図に平均値を載せるか載せないか。平均値は本来、箱ひげ図と関係はない。集団の中での自分の位置を点で示す方法もあり、自由である。学習指導要領解説で例示されたがために教科書会社が縛られてしまっている。

## (11) 「四分位数の定義は他にもある」とは

→ 第1, 3四分位値の定義はいろいろ

- i. 教科書の定義は文部科学省によって(?)統一され、中央値を除いたデータの上下のさらに真ん中の値。データの個数が偶数の場合は中央値をとるデータがないので(ii)と同じに。
- ii. 中央値を含めてデータの上下のさらに中央の値。中央値を求めるのに、最大値と最小値を使うのだから、二分した上下の真ん中を求めるにも中央値を使うのが自然でしょう、との考え。
- iii. Excelの関数では例で示すと、データが8個の場合、中央値は4番目と5番目の平均を取り、これを4.5番目とみなす。第1四分位値はこれと1番目の真ん中の2.75番目のデータとみなし、実際には2番目のデータと3番目のデータを1:3に内分した値とする。
  - データが多くなれば違いはほとんどない。箱ひげ図のかき方の違いのようなもの。(内分は高2の「図形と方程式」で扱うので、今の高校では無理)

## (12) 指導要領の「四分位偏差」は結局?

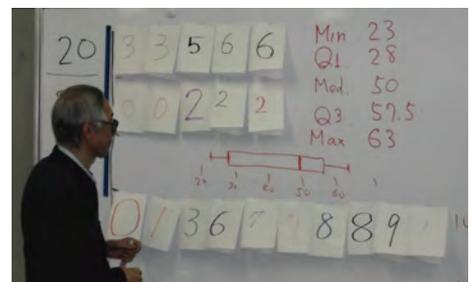
→ 四分位偏差=四分位範囲÷2だけど意味なし。「偏差」とは標準偏差のように平均からの程度離れているかを表すものであり、「四分位偏差」は中央値からの偏差を表すものではない。

指導要領に載っているのですが、教科書に記載せざるを得ないが、軽く扱われている。文部科学省(に限らず役所)は間違っても「ごめんなさい」とは言わない。

## (13) 教科書にない、でも必要なこと…それはデータの並べ換え

ここで並べ替えに伴って有効な幹葉図作りの演習を、参加者の年齢で実際に試してみることに。3月のサークルでも並べ替えるにあたって幹葉図を作る手順をよしさんから教えていただきましたが、限られた時間で手順とその結果だけだったので、煙に巻かれたようでした。それがきっかけで、じっくりと教えてもらう学習会を企画するに至ったのですが。

- i. (A4の紙が配られ)紙を半分に折る。
- ii. 自分の年齢の十の位を左側に、一の位を右側に。
- iii. 他の参加者にわからないように用意された袋に入れる。
- iv. 黒板の左端に20, 30, 40, 50, 60の「幹」を立てる。
- v. 袋から1枚ずつ取り出し、該当する「幹」の所に、半分に折った紙の一の位が見えるように「葉」を並べていく。
- vi. それぞれの「幹」の中で「葉」を順番に並べ換えて幹葉図の完成



結果は写真の通り紙の枚数が人数（20代が5名、30代が5名、50代が10名、60代が1名で、見事に40代が不在ということが明らかになり、教職員集団の年齢構成の縮図が現れました。20代、30代の先生方に期待するところ大だと、改めて感じた次第です）で、そのままヒストグラムにもなります。ヒストグラムを作るために「正」の字を書いていただけだと、その中でのデータの並べ替えができないので、幹葉図は一石二鳥の働きをするとのことでした。

● その他

(14) 散布図と相関について

（時間が押して省略）

(15) データの収集、結果の評価と解釈

→ 従来の手法と反するものではない。簡便法でもない。その特徴を理解して使うことが大事。

(16) 成績評価は？ センター試験は

→ 「統計検定」が始まり、3級程度がセンター試験の目安となるか？

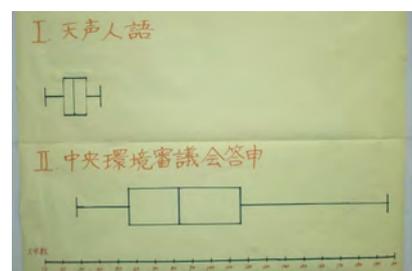
## II部 模擬授業「箱ひげ図でデータを比べてみよう」

2～3人ずつのグループ分けをして、よしださんが予め用意された次のデータを基にして、箱ひげ図を模造紙に書いて張りだし、発表することになりました。

- (1) 経済＝都道府県別最低賃金
- (2) 生活＝ やさいの価格
- (3) スポーツ＝ サッカーの試合の入場者数
- (4) 文章＝ 文章の1文あたりの文字数
- (5) 芸能＝ AKBメンバーの年齢
- (6) 確率＝ 変形サイコロで1の目が出る確率

できたグループから模造紙を張りだし、代表者が発表していくことに。それに対するよしださんのコメントを紹介します。

- (1) 最低賃金は実際に支払われている賃金ではないことに留意。今回は相関は触れなかったが、最低賃金と物価の相関を調べるのは単純な話ではない。物価が高いので最低賃金も高くなるのは「相関関係」というよりも、「関数関係かもしれない」、のように考察する必要がある。
- (2) 月ごとの変化を強調するのであれば折れ線グラフに。「価格の分布を見るのは箱ひげ図で」と役割の違いを再確認。
- (4) 文系の分野からデータを持ってきた。文献統計学では書かれた文章から作者を推理。源氏物語の中で最後の宇治十帖だけ別人が書いたという説があるが、それを統計学で確認する。



- (5) 高校生に興味を持ってもらえる題材。HKT の右側が伸びているのは、東京から移籍したメンバーが加入したからで、彼女がいなければもっと低年齢でまとまっていた。
- (6) 私も以前は班ごとにデータを並べて折れ線で結んだりしたが、折れ線グラフは変化を見るためのもので不適切だった。データの分析を先に教えておいて、確率の収束を実験して箱ひげ図で確かめる授業プランも。

- (3) については JEF 市原・千葉のホーム（フクダ電子アリーナ）の試合ごとの入場者数分布の推移だったのですが、若い先生が早速スマホで検索し、

2008 年 J1 15 位（入場者数が全体的に多い）

2009 年 J1 18 位

2010 年 J2 4 位（降格して全体的に減ったが、3 位までの昇格争いで最大の入場者を記録した試合もあった）

2011 年 J2 6 位

と、しっかり分析・解説を加えてくれました。このデータを出したのは、よしださんが市原市民で、これらの入場者数のうちの何人かはよしださんだからとのこと（笑）。また、天声人語は読み取り重視で、短い文章が主体。お役所の文書は一つ一つの文章が無駄にだらだらと長い（笑）との発表も共感を呼んでいました。

各グループの発表が終わったところで、よしださんから「データの分析を早めに教えて、夏休み・冬休みに課題を出してもらうのがいい」と提案がありました。参加者から「40代は人がいないのに、箱ひげ図にすると、箱の中に40代がいるような感覚になってしまう。間違えないように気をつけたい」と感想が出され、すかさず「そこが箱ひげ図の限界。ヒストグラムをつけることが大事。ヒストグラムも箱ひげ図も、ともに幹葉図から作られるので一石二鳥」とよしださんがコメントしてくれました。

アンケートから感想をいくつか紹介します。

- ・初めて参加させていただきました。大変興味深く、有意義な午後のひとときが過ごせました。ありがとうございます。今後生かすチャンスがあれば、さらに研究していきたいと思えます。
- ・データの分析は私たちのような平成生まれの層は、中学校・高校で学習する機会がありませんでした。そのため授業の流れのイメージがわからなかったため、今回の勉強会に参加させていただき、大変勉強になりました。自分自身、あまり勉強していない内容でしたから、今回の参加者の中で最も生徒の感覚に近かったのではないかと思います。
- ・今回初めて、こういった会へ参加しました。改めて、数学の楽しさを知ることができました。ありがとうございます。
- ・実際に作業してみて、よい記憶（印象）に残る勉強会になりました。是非続きをお願いします。
- ・生徒に実習（実際にかかせるなど）をさせる場合には多大の時間を要することがわかった。
- ・非常に参考になった。これから授業をするので、どうしたらよいか迷っていました。ありがとうございます。

- ・ 導入の背景や用語の細かな点が確認できてよかった。また幹葉図や箱ひげ図を実際を書くことで理解が深まった。データの分析は課題学習と相性が良さそうだと感じた。とても有意義だった。
- ・ 実際にデータを調べて箱ひげ図を作る，参加型の授業を取り入れることで，楽しいものにできそうだと感じられた。
- ・ 箱ひげ図を作った後の分析や考察を課題にしないと，箱ひげ図を作って終わりになってしまう，と思いました。
- ・ 箱ひげ図の限界などもわかり，よかった。

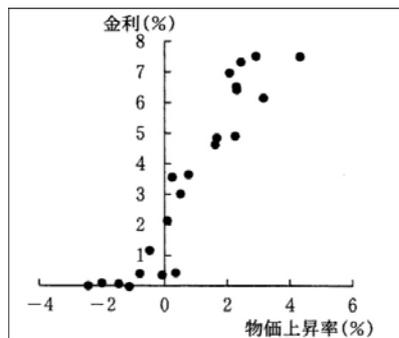
## 4 「データの分析」学習会 Part 2

第1回統計学習会と触れられなかった相関係数について，2011年12月のサークルで紹介していただいた内容の学習会もやりましょう，ということになりました。会場は前回と同じ蘇我勤労市民プラザを確保し，「データの分析」学習会 Part 2 として2013年10月20日に開催しました。

県外から2名，私学から3名のともに20代の先生が参加され数教協の輪が広がったのが嬉しかったです。講師のよしださんからは引き続き詳細な講義をしてくださいました。1958年の学習指導要領告示以降の中学校・高等学校数学科での相関の指導の変遷，統計分野の内容と学年配当の変遷をわかりやすく提示され，相関係数と関わりの深い散布図についても数学科以外にも理科・社会で実際に使われているとして，次のように示してくれました。

データの相関をみる散布図

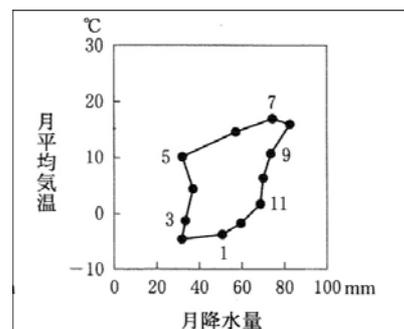
【2007年度センター試験・政経第5問】



正解選択肢「物価上昇率が高いときほど，金利が高いという傾向がある」

循環順序のあるデータをみる散布図

【2009年度センター試験・地理B第1問】



講義の中でよしださんが最も強調していたのは，一般的な教科書では，共分散を求め，これを標準偏差の積で割ることで相関係数を導入しているのに対し，今はなき「三省堂の教科書」では，対になっているデータから散布図をかく→図から相関のようすを見る→2組のデータをそれぞれ standardize する→散布図と standardize したデータから相関係数を定義する，ということでした。これは，あとで行われた模擬授業のグループ作業でも，計算の手間が多少緩和されるということで，実感することができました。

模擬授業では 47 都道府県の

1. 道路舗装率
2. 道路交通事故件数
3. 乗用車保有台数

の 3 つのデータ（総務省統計局統計学習サイトより）について 3 つのグループに分かれ、3 通りの組み合わせのうちの 1 つの組み合わせについて相関係数を求める作業をし、最後に散布図もつけて発表しました。



終了後のアンケートで○感想と◎昨年度の「データの分析」の実施状況が次のように寄せられました。

- よしだ先生、今日は貴重な授業をありがとうございました。
- ◎ 昨年度は未実施。今年度の 2 年生が 9 月以降、時間をたくさん（15～20 時間）とっています。
- はじめて相関係数を求めることをやりました。手で計算するのは中々大変なことがわかりました。データを少なくして何とかやりたいものです。楽しくやれました。
- ◎ 2 学期に 10 時間程度。
- 実践的な生データを用いてやってみることは、やはり楽しいと思いました。
- ◎ 文部科学省で定められた内容をこなす程度。終了後にフィードバック等を行っていない（中 3 = 中高一貫、10 時間程度）。
- 勉強になりました。GRAPES の基準化の見方が特に参考になりました。まだまだ勉強が不十分なのを実感しました。
- ◎ 本校ではまだ実施していません。数学 II を終えた後 1, 2 週間（週 4 時間）で、2 クラス合同でやる予定です。限られた時間で何をどう教えるか、何を作業させるか悩んでいます。今回の学習会で得たことを含めて、また考えていきたいと思います。
- 実際のデータで相関係数を出す経験は初めてだったので良い経験でした。今後データの正確さにもよりますが、もし何か良い例があれば相関係数を見てみたいと思います。基準化の意味がわかりました。
- 細かいところに触れながらの学習会で、とても勉強になりました。
- 「データの分析」は学生時代にしか習わず、全く白紙の状態これから教えていくことになるので、今回の学習会は本当に参考になりました。
- ◎ 教員 1 年目なのでこれから実施です。
- ◎ 中 3 の資料の整理で散布図を扱う。中 3 の 3 学期に相関係数を予定。中学生なので意味を教えながらグループワークを取り入れている。
- ◎ データの分析はカット。時間が足りなくてカットだらけ。

## 5 学習会を終えて

今回は授業実践の交流を中心にした学習会にしたいと考えています。新課程の 1 期生が大学入試センター試験を受ける 2015 年 1 月の問題を見て、早ければその年の 2 月、3 月にも分析・

検討会と合わせてできればいいですね（問い合わせは motosan61@yahoo.co.jp まで）。最後に、授業実践を2つ紹介して稿を終えます。

## 6 松戸南高校での授業（箱ひげ図の作成）

2013年2月5日1・2限、午前部1年次の標準クラスで、箱ひげ図を作るグループ作業に取り組みました（1年生は96人で、発展1クラス20人、標準28人×2クラス、基礎1クラス20人の4展開）。一応、前時（1月31日）の授業で、箱ひげ図に関わるデータの並び替えと四分位数等の説明・教科書の練習問題をやった上でのことです。

データ数が多いと並び替えが大変で、よしださんに教わった「幹葉図」が有効なので、その使い方もプリントで練習しておきました。当日出席したのは28人中19人だったので4～5人の4つのグループに分かれ、(1) AKB48の71人 (2) SKE48の66人 (3) NMB48の38人 (4) HKT48の17人の4チームについて、年齢順の並び替え→5値要約→箱ひげ図の作成と、黒板への書き込みを分担させました。

数学1プリント71 (データの並び替えと箱ひげ図) 1月31日  
1年 組 番 氏名

あるクラスの生徒40人について、1年間に図書室で借りた本の冊数を調べたところ、次のようなデータを得た。これを箱ひげ図で表すには(教科書156ページ後半問題5～2参照)

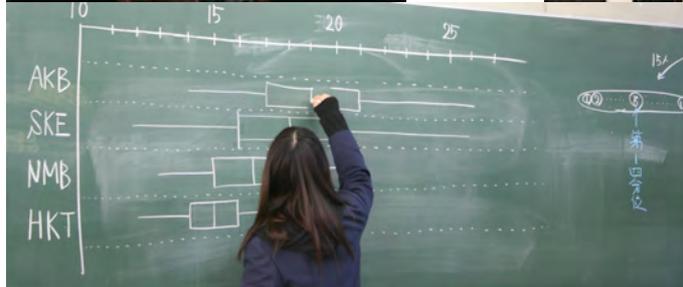
13	8	34	30	6	15	19	12	階級	各データの一位を左詰めで記入
1	17	32	10	9	21	10	14	0~	
19	20	4	4	11	20	22	8	5~	
18	2	28	11	19	8	17	14	10~	
15	30	22	12	18	24	6	12	15~	
								20~	
								25~	
								30~	

階級 左側から小さい順に並び替える

データが40個なので第2四分位数は( )番目の( )と( )番目の( )の平均をとって( )  
第1四分位数は( )番目の( )と( )番目の( )の平均をとって( )  
第3四分位数は( )番目の( )と( )番目の( )の平均をとって( )

箱ひげ図にまとめる

1年間に図書室で借りた本の冊数



## 7 青木洋子さん（国分高校）の実践発表

2014年3月のサークルに、メンバーの青木さんが1年生の2クラスで取り組ませた箱ひげ図のポスターを持ち込んでくれました。

データは気象庁のもので、過去60年分の10年ごとの全国各都市の月別平均気温をまとめたものを用意し、クラスで班を作らせ、分担して箱ひげ図にまとめたものです。各班の箱ひげ図を並べるだけではもったいないので、日本地図を大きな模造紙にかき、その上に各班がまとめた箱ひげ図の用紙を貼らせたとのことです。

60年分の箱ひげ図を並べると、東京など都市部で、60年前より平均気温が上がっており、最高気温と最低気温の差が大きくなっていること、沖縄など殆ど変化の見られない場所もあったとのことでした。

サークルの帰りに、荷物になるので私が車で持ち帰り、機会があれば国分高校まで持っていくことになったのですが…。そのままになっていたのが7月のサークルで青木さんとも相談し、8月の数教協全国研究大会のポスター展に持って行くことにしました。

ちば高校サークルとして「データの分析」学習会の様子を展示するのに合わせ、実物も展示しましょう、という訳です。合わせて8月23日に予定されていた千葉県数学教育協議会の夏の学習会の校種別実践発表も青木さんに引き受けていただきました。

当日は福井大学教育地域科学部教授で数教協副委員長でもある伊禮三之さんが顔を出され(馬橋に住んでいる娘さんのところに寄るついでだったそうです)、小・中・高の教員・OB・学生他22名が参加しました。開催地の市川市のデータは気象庁の観測地点でないため、代わりに船橋市のデータを用意され、箱ひげ図の作り方を教わり、全員で手分けして箱ひげ図を作り、それを並べるというものでした。1年ごとの推移(参加者数の見当がつかず、何年分も用意したので)では、即温暖化とまでは読み切れず、国分高校の生徒が発表してくれたように10年ごとのまとめたデータで推移を見ていくのがいいとの感想も出されました。

