

数学① 「データサイエンス入門」

三 橋 一 行

今回、授業内容の説明文として次の文章を参加者募集ホームページに掲載してもらった。「データを数学的に処理して判断してみよう。実験の結果やアンケートなど集められた数字からいったいどんなことが言えるのか？そして、そのための数学的処理というのはどんなものだろうか？統計的推定や仮説検定などを中心にコンピュータを用いてデータサイエンスの初歩の初歩に挑戦してみましょう。」

コンピュータを利用して、データ分析を体験してもらうのが授業の目的である。

1. <はじめに>

昨年度と異なる点は、テーマをしぼり、数値データからヒストグラム、平均値、分散、標準偏差をもとめて分析するという基本的な統計学の学習とした点である。最近の統計学の流行により、小学校や中学校でも統計学に関する授業に力が入れている。参加者にとって、この課題は簡単すぎるのではないかという不安を事前に抱いていたが、実際は、予想よりもはるかに授業が進まなかった。まず、予定されていた授業の流れを説明する。次に授業がどのような状態で終わったかについて述べて、その原因を探る。

今回は、中学校1～3年生までが参加しており、所属の中学校も様々である。したがって、通常ありえない極端な結果が、この授業の結果として出ていると考えられるが、今後、統計教育が高等学校でも盛んになっていくことを踏まえると、今回の状況は、今後の高校数学の授業で多くみられる生徒や学校の状態であると考えることができる。

2. <授業の内容>

授業の流れは、次の①～⑧のとおりである。

- ① 説明の後、コンピュータの表計算ソフトを起動。右のデータシート（各パソコンにサーバーからダウンロードできるようにした）をダウンロードする。
- ② ダウンロードしたデータシートを用いて
 - ・コンピュータで平均値を求める方法を教えて、求めさせる。
 - ・次に分散の意味を説明し、コンピュータでの求め方を教える。
 - ・標準偏差の意味を説明し、コンピュータでの求め方を教える。

年齢	血圧
56	147
42	125
72	160
36	118
63	149
47	128
55	150
49	145
38	115
42	140
68	152
60	155

- ③ 次に下のデータシートを用いて、架空の成績データを用いて、平均値、分散、標準偏差を求める。そして、問いに答える。
- ④ それだけでは、はっきりしない点があるので、ヒストグラムや度数分布を作成する。
- ⑤ ④の結果も踏まえて、再び問いについて考察（分析）して見る。
- ⑥ 相関係数に触れ、具体的データを用いてコンピュータで求めてみる。
- ⑦ 数学的背景などは置いておき、区間推定や仮説検定もコンピュータで行ってみる。
- ⑧ まとめ

3クラスの数学のテストの成績です。100点満点。
 問1 以下のデータを見て、次の問に答えなさい。
 (1) 授業が最もやりにくいクラスはどのクラスですか？理由も含めて答えましょう。
 (2) 各クラスにどのような授業をしてあげれば、数学の成績があがるとおもいますか。理由も含めて答えましょう。

※このデータはフィクションです。

出席番号	Aクラス	Bクラス	Cクラス
1	42	100	60
2	52	99	65
3	35	86	62
4	42	82	70
5	56	12	54
6	78	92	57
7	66	45	66
8	54	32	89
9	25	13	50
10	35	24	61
11	42	25	58
12	66	31	59
13	63	100	56
14	67	10	66
15	100	54	61
16	22	99	54
17	46	98	78
18	77	78	52
19	24	68	77
20	33	93	48
21	38	100	50
22	92	28	69
23	85	87	70
24	46	32	71
25	13	15	59
26	16	42	58
27	48	3	55
28	57	99	59
29	36	10	62
30	89	24	75

3. <授業の結果>

授業としては、ハード面、思考力の面などで格差が大きく出てしまい。データ分析の面白さがある程度つかめた生徒、コンピュータとの格闘で苦戦してしまいデータ分析までたどり着けなかった生徒、コンピュータの処理はしたものの数値から判断するには及ばなかった生徒といった具合に大きく3タイプに分かれたようである。地元や近県の中学校1～3年生が混ざったこの授業では致し方ないことではある。

もっとも進んだ生徒1名が⑤まで約2時間でたどり着けた。ほとんどの生徒は、2番目のデータシートで度数分布作成段階で苦戦していた。度数分布をパソコン上で作成するのが難しかったようである。しかし、それ以前に①の段階についてこられない生徒が多数いた。かなりゆっくり説明をしたつもりだが、説明を受けてもついていけないという状態であった。また、ハードウェアがフリーズしてしまうなど予想外の障壁にぶつかってしまった生徒もいた。この様な状態のため授業は思いのほか進まず、不本意ながら、足並みをそろえることができずに時間切れとなって

しまった。この授業の失敗の原因を列挙してみると次の通りである。

① パソコンの使用に関して習熟度の差が大きい

パソコンの起動の仕方、マウスの使い方、表計算ソフトの使い方などパソコンを頻繁に使用している生徒とはじめて使用するに等しい生徒がいた。

② パソコンの不具合

前日までに作動点検など行っていたにも関わらず、使用中に何らかのトラブルが発生し、コンピュータ、あるいはアプリケーションが動かなくなった。

③ 統計処理の数学的習熟が不十分である

説明を受けたものの用語や式の意味がつかめない。データ分析の経験が少ないためか数学との関連性がつかめない。

④ 統計量の値を求めてもその意味を考えようとしない

③のこともあって数学的に考えることが難しい。答えだけを出す習慣が強いいためか、答えの意味を考えることを不得意としている。

⑤ コンピュータの使用は授業や統計学習の手助け（時間短縮）になるとは限らない。

コンピュータの利用で、「多くの面倒な計算が短時間に処理できて考察に時間をかけられる。」「統計学の数式は複雑であり、計算が面倒であるがパソコンを使用すればその苦労なく結果にたどり着け、統計の良さがわかる。」などということは必ずしも言えない。生徒によってはコンピュータを操作するという新たな負担が増加する。

生徒の側の問題点だけではなく、機器に関係する問題や教える側の問題も大きく関わっている。まずは、機器の問題を何とかしなければならないだろう。これに関しては、このタイプの授業の根幹であるので、学習が円滑に進む環境づくりが必要である。教える側にとってコンピュータの利用は生徒にとって学習活動の支援になる前提で考えられているが、数学的なしくみとコンピュータの関わりあい、およびその操作に関する知識・技能がなければ、それがかえって疑問を生んでしまったり、学習の障壁となってしまう。これらのことを解消しないと本来の授業目的達成は困難であろう。次に、これらについての対策を考えてみたい。

4. <対策とまとめ>

①に対しては、今回のような単発の授業においては募集の段階で必要なコンピュータの知識・技能を明記すべきであった。通常の高등학교の授業においては情報科との連携をきちんと行い、コンピュータに関する知識技能を身に付けることが必要だろう。

②に対しては、前もって点検、準備をしておくことが重要であるが、それでも発生を0%にすることは難しいであろう。操作の間違いなどで、実際の操作が指示通り行われないうこともあり、それによって想定外の状態に陥ることも考えられる。よって、

予備の機器の準備を行う、または情報科教員など操作に詳しい教員とティーム・ティーチングで授業にあたるのが良い。

③に対しては、実際に分析する前に統計学の理論的側面をしっかり勉強し、簡単な例で構わないので手計算や電卓をもちいて行うことが大切である。その際、数学を活用して統計学がつけられていることにできるだけ多く触れさせておくことが大切である。さらに言うならば、中学校段階、小学校段階での統計に関する指導がしっかりと行われていることが望ましい。

④に対しては、統計学分野に限らず、日ごろの数学の指導が問題の答えを求める作業だけで終わらないような指導が重要である。そのためには、全体の展望を持たせ、現在の単元がどの位置にあり、これがどこへつながるのかを知らせる指導も必要である。また、解くという作業を俯瞰して見つめなおし、どのような発想、方針、技術によって解かれているか知ること、そして、求められた答えが問いに対してどのような結果をもたらしているのかを味わうことも必要である。

⑤に対しては、主に指導者が注意するべきである。われわれがコンピュータの使用によって作業が楽になる便利な装置だと考えるのは、主にコンピュータを使用しなかった場合の苦勞を知っているからだ。あるいは、想像できるからである。その経験をもとにコンピュータの使用によって作業が簡単になると考えがちであるが、生徒はコンピュータのなしの場合の苦勞も分からない上に、コンピュータの操作という作業が上乘せされてしまう。コンピュータの操作に慣れていなければ、生徒の負担が増加するだけである。したがって、手計算などのコンピュータなしの場合を一度経験しておくこと、さらにコンピュータの操作に十分慣れさせておくことが重要である。統計学は多くのレディネスを必要とする分野なのである。

以上のことから、今回の理数体験の授業では、授業自体は上手く進行しなかったが、その原因を探ることで、パソコンを用いた統計の授業を行う際に注意しておくべきことをあらためて知ることができた。この記録をもとに、次回以降、同タイプの授業を行う際に、授業の停滞を未然に防ぎ、最大限の効果が得られるよう努めたいと考えている。