岐阜数学教育研究
2015, Vol. 14, 7-13

統計学における手計算の学習効果について

赤堀　克己[1]

大学生や一般社会人用の統計学のテキストでは，実際のデータを意識して煩雑な数値を用いる問題がほとんどであり，電卓等の計算機の利用を前提としている。たしかに，単純計算に時間を費やすのは意味をなさないであろう。しかし，高校数学の学習過程や大学入試において電卓を使わず手計算を前提としている現状を鑑みるに，統計学の学習の初期段階において計算機の利用が最善かどうかは検討の余地があると思われる。そこで，本実践においては，極力数値を簡略化して手計算でも容易に答えを得ることができる教材を用意し，大学１年生を対象として実践を行った。さらに，Excel を用いての実習も同時に行い，両者の比較検討を試みた。本稿では，演習後のアンケートから統計学における手計算の学習効果についての考察を述べる。

＜キーワード＞　統計学，手計算

１．はじめに

多くの情報が溢れている現代社会において，データを適切に処理し，必要な情報のみを抽出して活用することが様々な場面で求められている。その際，必要とされるのが統計的なものの見方や考え方である。それを受けて平成２０年度の中学校学習指導要領数学科の改訂([1]を参照)において，「資料の活用」が領域として追加された。また，平成２１年度の高等学校学習指導要領数学科の改訂([2]を参照)においても，数学Ⅰの「データの分析」が必須になった。さらに，大学においても文系理系問わず多くの大学生が「統計学」を履修する。

高等学校において，「統計学」は「数学」の授業の中で学習するがゆえ，「数学」と同様の学習方法でよいという印象を学習者に与える。しかし，「統計学」を学習する際，「数学」とは大きく異なる点がいくつか存在する。そのひとつとして，数値の違いである。「数学」の演習用の問題において，煩雑な数値が用いられることはほとんどありえない。しかし，「統計学」の場合は，現実のデータを意識するがゆえ，計算が煩雑になる数値を用いられることが多い。特に，大学生や一般社会人用の統計学のテキストにおいては手計算の制約がないため，煩雑な計算を必要とする演習問題がほとんどである。煩雑な単純計算自体は「統計学」の学習においては，意味をなさないであろう。しかし，「数学」の学習過程において最後まで答えを手計算で求めることに慣れてきた学習者が，「統計学」の学習において，最終的な答えを計算機で求めるとき，満足感や達成感を感じることが出来るだろか？

筆者は長年，大学で Excel を用いて実習を行う「情報処理科学」を担当してきた。その中で統計的な内容も取り上げてきた。Excel の実習の後に，その日取り上げた「統計学」の概念の理解を確認するため，数値を

[1]岐阜薬科大学

簡略化して手計算でも容易に答えを得ることができる問題を受講者に解いてもらっていた。受講者の取り組みの様子や解答後の感想から,「統計学」の概念の理解において,「手計算」には一定の学習効果があるという印象を得ていた。以下, 学習者の意識を実証するために実施したアンケート結果を報告するとともに,「統計学」と「数学」の特性の違いに留意した「統計学」の効果的な学習方法についての考察を述べる。

２．研究のねらい

２．１．高等学校での履修状況

平成２７年度の大学１年生は，高等学校において新しい学習指導要領のもとで学習しており, 必須である数学 I の「データの分析」の履修率は 100%であろうと予想していた。しかし, 今回のアンケートにおける新課程履修者の数学 I の「データの分析」の履修率は，65.2%(23 人中 15 人)に留まった。一方で,「確率分布と統計的な推測」(数学 B) の履修率は，21.7%(23 人中 5 人)と旧課程履修者の 7.1%(14 人中 1 人)に比べて高い数値であった。この結果より, 今回の学習指導要領の改訂が高等学校の「統計学」の学習に少なからず影響を及ぼしていることがわかる。数学 B の統計分野の履修率の上昇は喜ばしいことではあるが，２割程度の履修率では，高等学校と大学の間でスムーズな内容の深化が可能になったとは言い難いのではなかろうか？旧課程時と同様に，まだまだ「統計学」の学習においては, 大学での効率的な学習が求められるのが現状であると筆者は感じている。

２．２.「数学」の学習過程と「統計学」の特性

高等学校の「数学」の学習過程において演習は不可欠である。その際, 手計算で最後まで答えを求める。学習者が答えを得て, それが正解であった場合，学習者は他の教科にはない達成感を感じ，学習意欲が益々向上し，次の学習につながっていく。

一方,「統計学」はデータを身近なテーマから取り上げることができることから，学習者の興味・関心を引きやすい。しかし，現実のデータを扱うと計算は煩雑になり，最後まで手計算で行うことに面倒くささを感じることは否めない。その面倒くささは, 学習者の意欲を減退させる。今回の実践では,「統計学」の学習におけるその欠点の排除を目指し,「数学」の学習過程と同様の学習過程が「統計学」でも可能かどうかを検証するものである。

２．３．学習者の意識の調査

我々は，手計算でも容易に最後まで計算可能な問題を用意し，受講者に解いてもらった。手計算の後，同じ問題で Excel の実習を各自行ってもらった。手計算と Excel, どちらが統計量の理解に効果があったかを調査するため，受講者にアンケートを実施した。

３．実践内容

講義名：「情報処理科学」

実践日：平成 27 年 10 月 5 日(月)

場所：岐阜薬科大学三田洞学舎村山情報処理センター

対象：大学１年生（有効回答数３７名）

３．１．受講者について

今回の受講者のほとんどが高等学校の「情報」の授業で Excel を経験しており，さらに前期の「情報基礎実習」でも Excel の操作方法を復習済みである。

３．２．教材における工夫

今回用意したのは平均，分散，共分散，相関係数に関する問題であり，各統計量の定義の理解を目指すものである。用いるデータを各統計量が意味をなす範囲で極力簡略化し，手計算の煩雑さを無くした。また，答えも受講者が学習後の満足感を感じられるように，出来る限り簡潔な数値になるように工夫した。

３．３．実践の流れ

まず，統計量についての説明を行った。その後，実際に手計算で平均，分散，共分散，相関係数に関する問題を受講者に解いてもらった。手計算の演習終了後，今度はＥｘｃｅｌを用いて手計算の答えの確認作業を各自でしてもらった。

[受講者への配布プリント]

ｎ個のデータ x_i $(i=1,2,\ \cdots,\ n)$について，平均 $\bar{x}$，分散 σ^2，標準偏差 σ はそれぞれ

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,\quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2,\quad \sigma = \sqrt{\sigma^2}$$

である。

問１（１）次の５個のデータについて，平均，分散，標準偏差を手計算で求めよ。

５，７，３，１，４

（２）同じデータについて，平均，分散，標準偏差をそれぞれＥｘｃｅｌの関数　ＡＶＥＲＡＧＥ，ＶＡＲＰ，ＳＴＤＥＶＰを用いて求めよ。

２変量データ$(x_i,\ y_i)$ $(i=1,2,\ \cdots,\ n)$について，

(1) x の分散$\sigma_x{}^2$，y の分散$\sigma_y{}^2$

$$\sigma_x{}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\ \text{（ただし，}\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\text{）}$$
$$\sigma_y{}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2\ \text{（ただし，}\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i\text{）}$$

(2) x と y の共分散σ_{xy}

$$\sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

(3) 相関係数ρ_{xy}

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$$

問２（１）（x，ｙ）についての３個のデータ（０，１），（１，２），（２，３）について，共分散と相関係数を手計算で求めた後，Ｅｘｃｅｌで確認せよ。ただし，Ｅｘｃｅｌの実習において，共分散は COVAR，相関係数は CORREL を用いよ。

（２）（１）と同様のことを３個のデータ（０，３），（１，２），（２，１）についても行え。

（３）（１）と同様のことを４個のデータ（１，１），（１，２），（２，１），（２，２）についても行え。

(解答)

問１　平均：$(5+7+3+1+4)/5 = \frac{20}{5} = 4$

分散：$\{(5-4)^2+(7-4)^2+(3-4)^2+(1-4)^2+(4-4)^2\}/5$

$= (1+9+1+9+0)/5 = \frac{20}{5} = 4$

標準偏差：$\sqrt{4} = 2$

問２（１）$\bar{x} = (0+1+2)/3=1$，$\bar{y} = (1+2+3)/3 = 2$

$\sigma_x{}^2 = \{(0-1)^2+(1-1)^2+(2-1)^2\}/3 = \frac{2}{3}$，$\sigma_y{}^2 = \{(1-2)^2+(2-2)^2+(3-2)^2\}/3 = \frac{2}{3}$

$\sigma_{xy} = \{(0-1)(1-2)+(1-1)(2-2)+(2-1)(3-2)\}/3 = \frac{2}{3}$

$\rho_{xy} = \frac{\frac{2}{3}}{\sqrt{\frac{2}{3}}\sqrt{\frac{2}{3}}} = 1$

（２）$\bar{x} = (0+1+2)/3=1$，$\bar{y} = (3+2+1)/3 = 2$

$\sigma_x{}^2 = \{(0-1)^2+(1-1)^2+(2-1)^2\}/3 = \frac{2}{3}$，$\sigma_y{}^2 = \{(3-2)^2+(2-2)^2+(1-2)^2\}/3 = \frac{2}{3}$

$\sigma_{xy} = \{(0-1)(3-2)+(1-1)(2-2)+(2-1)(1-2)\}/3 = \frac{-2}{3}$

$\rho_{xy} = \frac{\frac{-2}{3}}{\sqrt{\frac{2}{3}}\sqrt{\frac{2}{3}}} = -1$

(3) $\bar{x} = (1+2+1+2)/4 = \frac{3}{2}$，$\bar{y} = (1+1+2+2)/4 = \frac{3}{2}$

$\sigma_x{}^2 = \{(1-\frac{3}{2})^2 + (2-\frac{3}{2})^2+ (1-\frac{3}{2})^2+ (2-\frac{3}{2})^2\}/4 = \frac{1}{4}$

$\sigma_y{}^2 = \{(1-\frac{3}{2})^2 + (1-\frac{3}{2})^2+ (2-\frac{3}{2})^2+ (2-\frac{3}{2})^2\}/4 = \frac{1}{4}$

$\sigma_{xy} = \{(1-\frac{3}{2})(1-\frac{3}{2})+ (2-\frac{3}{2})(1-\frac{3}{2})+ (1-\frac{3}{2})(2-\frac{3}{2})+ (2-\frac{3}{2})(2-\frac{3}{2})\}/4 = 0$

$\rho_{xy} = \frac{0}{\sqrt{\frac{1}{4}}\sqrt{\frac{1}{4}}} = 0$

４．実践における受講者の活動の様子

４．１．手計算と Excel による実習

旧課程履修者の中には，分散，共分散，相関係数を初めて学習する受講者もいたが，集中して課題に取り組んでいた。また，Excel の操作自体に戸惑っている受講者はおらず，操作方法についての質問はでなかった。

５．アンケート結果と考察

実習後,「手計算」と「Excel」に対する受講者の感想を比較するため，以下のアンケートを実施した。

Ｑ１．今日の演習において,「手計算」と「Excel」，どちらが統計量（分散，共分散，相関係数）の定義の理解により役立ちましたか？

1. 「手計算」
2. 「Excel」

Ｑ２.「手計算」についての感想を書いてください。

問１に対する回答の集計結果が[表１]である。

回答番号	新課程履修者	旧課程履修者	合計
1	19 人	10 人	29 人
2	4 人	4 人	8 人
合計	23 人	14 人	37 人

[表１]

[表 1]より,「手計算」の方が統計量の定義の理解に有益だったと回答した受講者は全体の 78.4%にのぼり,「手計算」が統計学の学習において一定の意味を成すことが確認出来た。

また，Q2 の「手計算」の感想から，統計学の学習における問題点が見えてきた。以下，受講者の感想の中から際立ったものを紹介する。

[受講者の感想]

(肯定派)

① 実際に手を動かすことで定義を覚えることが出来ました。
② 手計算をすると時間はかかるが, どのように算出するかが思い出されたのでいい復習になりました。
③ Excel よりも理解しながら計算が出来たので楽しかったし，Excel に頼り過ぎてしまうと自分の力が全くつかないように感じました。
④ センター試験予想問題より数字が綺麗だったのでやりやすかった。
⑤ 数値が簡単でデータ量も少なければ手計算の方が圧倒的に楽だと思った。
⑥ きちんと頭を使ってできた。センター試験のことを思い出して嫌な気持ちになった。

(一長一短派)

⑦ 計算は大変だが, 解き終わって答えがあっているときはうれしい。
⑧ 単純な数値なら手計算の方が定義を理解がしやすいですが, 実際に計算するのは面倒でした。
⑨ 大変だけど，理解しやすい。
⑩ 面倒だけど，達成感がある。

(否定派)

⑪ 計算が面倒くさかった。
⑫ 計算ミスをして時間がかかり, 大変でした。
⑬ Excel がないと大変だと思いました。

受講者の感想を分類すると, (肯定派), (一長一短派)，(否定派)の３つに大きく分かれ

る。まず，(肯定派)の①から③の感想は筆者が当初予想したとおりのものだった。④の感想も数値の工夫の効果を実証するものである。

次に，(一長一短派)の感想も集約すると，「手計算は（Excel にくらべて）大変だが，学習効果を感じることが出来る」と解釈してよいだろう。特に，⑦や⑩の感想は数学の学習過程に慣れている受講者にとって，今回の試みは学習意欲向上の効果があることを示すものである。

３番目に，(否定派)の感想も，④，⑤，⑫の意見もあることから推測するに「Excel と比べると」と注釈を加えてもよいだろう。Excel を用いると瞬時に答えが出てしまうので，相対的に手計算が面倒に感じたと思われる。

最後に，④や⑥のようにセンター試験の問題に関連した感想もいくつか見受けられた。国公立大学の２次試験で統計の問題は極めて少なく，高校生にとってセンター試験の数学Ⅰ・数学Ａの必須問題のみが直接的な影響を及ぼす。以下は，今年度のセンター試験における相関係数に関する問題である。

[平成２７年度センター試験　数学Ⅰ・数学A]

第 3 問 [2]　ある高校２年生 40 人のクラスで一人２回ずつハンドボール投げの飛距離のデータを取ることにした。次の図２(省略)は，１回目のデータを横軸に，２回目のデータを縦軸にとった散布図である。なお，一人の生徒が欠席したため，39 人のデータとなっている。

	平均値	中央値	分散	標準偏差
1 回目のデータ	24.70	24.30	67.40	8.21
2 回目のデータ	26.90	26.40	48.72	6.98

1 回目のデータと 2 回目のデータの共分散	54.30

(共分散とは 1 回目のデータの偏差と 2 回目のデータの偏差の積の平均である)

次の $\boxed{\text{ク}}$ にあてはまるものを，下の⓪～⑨のうちから一つ選べ。

1 回目のデータと 2 回目のデータの相関係数に最も近い値は，$\boxed{\text{ク}}$ である。

⓪0.67　①0.71　②0.75　③0.79　④0.83

⑤0.87　⑥0.91　⑦0.95　⑧0.99　⑨1.03

この問題は，相関係数の定義($\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$) を理解していれば，容易に解答できる問題である。しかし，データは少数点以下２位まで与えられており，計算はかなり労力を要する。この点が，④や⑥の感想になったと思われる。

６．課題と提言

６．１．学習手順の問題

今回の実践では，統計量の理解に重点をおいて展開した。簡略化したデータは，計算の煩雑さを排除してくれるが，同時にリアリティーも消失させてしまう。授業の開始時は，やはり学習者が興味を引くテーマで，生のデータを活用したい。では，今回の実践のような演習を復習問題として行うのがよいのか？適切な展開順序は今後の課題と捉えている。

６．２．統計学を理解するための数学の活用

問２の相関係数ρ_{xy}の説明の際に，－１≦ρ_{xy}≦１であることを以下のようにベクトルを用いて解説した。

[－１≦ρ_{xy}≦１であることの証明]

２変量データ$(x_i,\ y_i)$ $(i = 1, 2, \cdots, n)$について，

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,\quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

とする。さらに，２変量の誤差ベクトルをそれぞれ

$$\overrightarrow{\varDelta x} = (x_1 - \bar{x},\ x_2 - \bar{x},\ \cdots,\ x_n - \bar{x}),\quad \overrightarrow{\varDelta y} = (y_1 - \bar{y},\ y_2 - \bar{y},\ \cdots,\ y_n - \bar{y})$$

とすると，

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

$$= \frac{\overrightarrow{\varDelta x}\cdot\overrightarrow{\varDelta y}}{\left|\overrightarrow{\varDelta x}\right|\left|\overrightarrow{\varDelta y}\right|}$$

$$= \cos\theta$$ （ただし，θは$\overrightarrow{\varDelta x}$と$\overrightarrow{\varDelta y}$のなす角）

ゆえに，－１≦ρ_{xy}≦１である。

この証明は，ベクトルや内積の知識を利用している。数学の知識が統計学で効果的に利用できる例である。数学Ｂ履修後ならば，ｎ＝３の場合は高校生でも十分理解可能である。統計量の定義をただ天下り的に理解するのではなく，この例のように多面的な理解を試みれば，統計学に対する学習者の意識も変わってくるのではないか？と筆者は感じている。

引用・参考文献

[1] 文部科学省, 2008, 中学校学習指導要領解説　数学編

[2] 文部科学省, 2009, 高等学校学習指導要領解説　数学編