

IT 人財育成を見据えた数学授業の 新たな方向性

— GNU R を用いた検証授業を通じて —

小泉真也・松尾響・佐藤元彦

●要約

「ビッグデータ時代」と言われる昨今、データ分析に係る人財育成は、ビジネスの拡大や国際競争力の面から、産官それぞれが要求する喫緊の課題である。学校教育においては確率・統計の位置づけは重要なものとなり、近年の情報の活用を重視する教育政策が見据える方向性は、数学の授業体系に対しても大きな変化を示唆している。

著者らは、中学および高校の確率・統計授業の新たな方向性として、プログラミング言語という語彙・文脈の仲立ちによって数学的理解を高めることを企図し、統計解析ソフトウェア GNU R を活用した授業体系を提案している。今年度はこの提案に対する検証として、中学および高校に向けた授業をそれぞれ企画し、実際に授業を行って、生徒、授業者、そして監修者の立場からの評価及び意見を集約した。

本稿では、生徒の評価を分析することによって著者らが企画した検証授業の効果を示すとともに、授業者および監修者の評価から、確率・統計における授業体系の改善について考察した。

●キーワード

教育の情報化

確率・統計

ビッグデータ

1. はじめに

「教育とは、政治社会と個々人の双方が必要とする能力・態度の形成のために機能するものである」エミール・デュルケームの著書「教育と社会学」⁽¹⁾における解釈は、ひとつの側面として、教育を行うことの意味を、そして教育のあるべき方向性を示している。

「現代はデジタル社会」と言われて久しい。「教育の情報化」はデジタル社会における必然的な流れであるといえる。教育の情報化の重要性は、昭和60年6月の臨時教育審議会第1次答申において指摘されており、昭和61年4月の同審議会第2次答申では、情報及び情報手段を主体的に選択し活用していくための個人の基礎的な資質が、読み、書き、算盤に並ぶ基礎・基本と位置付けられた⁽²⁾。平成20年以降の教育政策に係る提言では、情報の活用について数学的思考力やデータ分析といったキーワードが明記されており、最新動向として、平成26年6月24日に閣議決定された「世界最先端 IT 国家創造宣言改定」では、IT 人財の育成・発掘・支援の土台として、初等・中等教育段階におけるプログラミングや情報セキュリティ等、情報教育の早急な推進が謳われた⁽³⁾。

デジタル社会において、あらゆる物事は情報通信技術が求めるかたちで数量や名義として数値化され、デジタル・データとなってコンピュータに取り込まれてきた。多くの家電品や電話、自動車などといった、デジタル・データを扱う道具は、さながら特定の用途に向けたコンピュータのごとく進化を続けている。いわゆる「計算機」を介する以上、情報通信技術は数学理論のフォロワーであり、両者の関係性は密接にして相補的である。演算は新たなデータを生み、我々は分析のためにデータを取得し、分類・整理・成型・取捨選択によって価値のある意味を見出そうとする。思考や判断といった無意識の営みは計算によって意識化され、情報技術は計算を高速化し、通信技術と相まって我々の意思決定にスピードをもたらした。言わば「デジタル社会」とは、我々の生活の中にデジタル・データ化できるものがあるとき、生活の諸問題は「計算」によって素早い解決が可能な社会であることを意味する。

今後も発展を続けるデジタル社会の中で、人財育成の観点からも、数学教育の位置づけはより重要になるであろう。特に、産業界が注目する「ビッグデータ」ビジネスにおいては、データサイエンティスト不足が顕在化している。2013年の情報通信白書⁽⁴⁾では、ビッグデータをフル活用した場合、現状でも年間7兆7700億円の経済効果が見込めるといっている。しかしながら現在、日本におけるデータサイエンティストは千人程度といわれ、米調査会社ガートナーの試算では、将来的に日本国内で約25万人のデータサイエンティストが不足状態となるとの見通しを示している⁽⁵⁾。

著者らは、一連の政策提言が、授業をわかりやすくしたり、校務を効率化したりするためのコンピュータの導入が主であった「教育の情報化」が新たな段階に移ることを示唆するものであると捉えている。特に、数学の問題解決授業は情報教育との融合が図られ、その文脈は大きく変わると推察する。平成25年6月6日、文部科学省は大学入試センター試験を廃止し、高校在学中に複数回受験できる「到達度テスト」を創設する検討に入っており、ICTによる問題解決がこうした試験に導入しやすくなることも推し量られる。

一連の動向に対して、著者らは昨年から今年にかけて、中学および高校の確率・統計授業にフリーソフトウェア GNU R を導入し、数学理論をプログラミング言語 R の手続きとして可視化する授業を提案した。本稿では、今年度実施した検証授業について、著者それぞれによる多面的な評価や顕在

化した問題点を報告し、IT 人材育成を見据えた数学授業の留意点を示す。

以下に本稿の構成を示す。2 章では検証授業の概要と受講者アンケートの分析結果を示す。3 章では授業者らと監修者それぞれの立場から検証授業を評価する。4 章では教育現場の視点で R 言語の習熟について言及し、GNU R を導入した確率・統計の指導案を示す。

なお、本稿では、特に断りがない限り、以降 GNU R を「R」、R で用いられるプログラミング言語を「R 言語」、R 言語において入出力を伴う処理体系を「関数」、そして命令構文と関数の総称を「コマンド」と呼ぶ。また、情報通信技術を意味する用語として「IT」を用いるが、教育現場での利用を示す場合および専門用語については「ICT」の表現を用いることとする。

2. 検証授業

2.1. 概要

表 1 に著者らが企画した検証授業について、中学校向けおよび高等学校向けのカリキュラムを示す。ここで、中学校向けカリキュラムについては、自然現象における確率・統計の性質の理解のために、中学校の範囲外である正規分布を扱った。また高校向けカリキュラムについては、標本分散の理解のために、高校の範囲外である不偏分散を扱った。

授業の進行は、いずれも 1 コマ 50 分の授業を 1 日あたり 2 コマとして、著者から、数学理論を松尾が、情報理論を小泉が担当する二名体制とし、佐藤が監修した。

表 1. 検証授業のカリキュラム

a) 稚内南中学校

コマ	テーマ	概要
1	確率・統計の定義と認識	「同様に確からしい」とは 乱数を用いたコイン・サイコロのシミュレーション
2	代表値（平均、中央値、最頻値）	定義の確認、中学生の身長を例にした分布データの作成、ヒストグラム・度数分布表の作成、標本調査
3	「平均」の意味づけ	標本平均と母平均の関係 宝くじの当籤確率と期待値の関係
4	モンテカルロ法	乱数を用いた円周率の近似

b) 稚内大谷高等学校

コマ	テーマ	概要
1	確率・統計の定義と認識、代表値、標本平均と母平均の関係、標本分散と母分散の関係	稚内南中学校の 1～3 コマの内容に相当 代表値では、平均、中央値、最頻値に加えて、分散、標準偏差について確認
2	四分位数、箱ひげ図	四分位数の定義、1 コマ目の代表値を求めたデータから四分位数を求め、箱ひげ図を書く
3	相関係数	散布図、相関係数の定義
4	打率を用いた野球の打順の戦略	グループワーク、ゲーム形式の実践

検証授業の実施には、稚内南中学校と稚内大谷高等学校の協力を仰ぎ、稚内南中学校は平成26年8月24日と9月3日に同校コンピュータ室において、中学3年生7名の参加と、教師2名の参観で実施した。稚内大谷高等学校は同年9月24～25日に稚内北星学園大学の演習室において、高校3年生38名の参加と、教師1名の参観で実施した。なお、稚内南中学校については、新たに3時間目から参加した生徒がいたため、Rの使い方など1回目と2回目の復習を行った。

2.2. 受講者アンケート

検証授業の効果を解析するために、中学・高校それぞれの検証授業終了後に、無記名による5段階評価のアンケート調査を行った。図1は、十分なデータ数を得た稚内大谷高校の回答について主成分分析を行い、横軸に第一主成分、縦軸に第二主成分を取ってバイプロットで示したものである。

アンケート項目について、Q.3およびQ.4は数学に対する普段の興味と学力をそれぞれ示す。Q.5からQ.7まではコンピュータの利用に対する普段の興味関心を示し、Q.5は情報機器の活用状況、Q.6はコンピュータ操作に対する興味、そしてQ.7はMicrosoft Excelの理解の自己申告である。Q.8は今回の授業に対する期待であり、Q.8-1は数学の復習、Q.8-2はコンピュータの利用、そしてQ.8-3はICTを用いた数学授業の体系についての回答である。Q.9は検証授業を通じて生徒が体感した効果であり、Q.9-1は確率・統計の理解、Q.9-2は授業の楽しさ、Q.9-3はRの使い方の理解、Q.9-4はRの操作・動作の楽しさ、そしてQ.9-5はRの継続的な学習への興味についての回答である。図中1～34はアンケート回答者それぞれの第一および第二主成分得点の合成成分をプロットしたものである。矢印はアンケート項目ごとの主成分得点に対する寄与の大きさを示し、矢印が示す方向に配置された回答者ほど、アンケートの評価スコアは高い。

図1では、数学に対する興味および学力が高いと回答した生徒は第二象限にプロットされ、情報に対する興味およびスキルが高いと回答した生徒は第三象限にプロットされた。したがって、第一象限

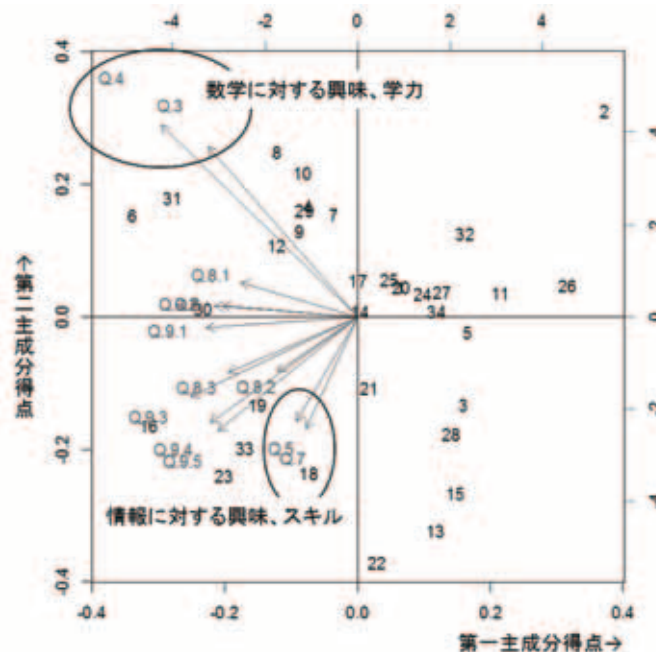


図1. 稚内大谷高校検証授業のアンケート分析（第一～第二主成分）

は情報に対する興味およびスキルが低いと回答した生徒、そして第四象限は数学に対する興味およびスキルが低いと回答した生徒であることを意味する。

図1から、数学に対する興味および関心と情報に対する興味およびスキルは、異なる性質を持つことが確認できる。授業に対する期待、および授業を受けての授業に対する評価の高さは、数学と情報の興味それぞれのベクトルが成す劣角の範囲に位置付けられている。特に授業の評価に係る Q.9のベクトルは、ほぼ第三象限に集中した。これは、今回の検証授業が、情報が得意な生徒に対してプログラミングの理解と数学に対する理解の面で両方の教育効果があったことを示すものである。

数学が得意な生徒は、検証授業に対して数学の復習を期待したという影響が強く、学習効果は大きくなかったと判断できるが、授業の楽しさを提供できたという点では学習効果の向上に可能性を示したと考える。

数学および情報に苦手意識を持つ生徒については、学習効果は確認できなかった。しかしながら、情報を苦手とする第一象限のプロットが原点付近に集中し、数学を苦手とする第四象限のプロットが第三象限寄りに位置していることから、授業体系の改善によってこれらの学生にも学習効果を得る可能性は決して小さくはないと考える。また、Q.9-4と Q.9-5のベクトルが極めて近い向きにあり、ソフトウェアの動作に楽しみを見出すことが、自主的かつ継続的な情報分野の学習につながると考えられる。

3. 検証授業の評価

3.1. 授業者による評価

本節では、初めてコンピュータを用いて授業を受ける学生や、ICTを導入して授業を行う教員の立場から今回の検証授業を評価する。

3.1.1. 授業の準備および進行の効率化

平成20年告示の新学習指導要領では、中学・高校の数学教育の情報化の充実にあたり、イメージしづらいものの可視化、および情報収集とその分析・活用が謳われており、確率・統計の授業におけるRの導入は一連の要求を満たすひとつの手立てである。

たとえば、データ作成では、ホワイトスペースやカンマなどの区切り記号を用いたテキストファイルに対するデータ入力処理を自動化し、乱数の活用は任意のデータ生成においてタイピング等の冗長な作業を低減する。表およびグラフの作成では、度数分布表、ヒストグラム、散布図それぞれに対応する関数があり、効率よく資料を整理し視覚化を実現する。さらには、コインやサイコロのシミュレーションのような多数の繰り返しを再現する場合は、一度入力したコマンドを呼び出す履歴機能やR言語の繰り返し構文、そして関数作成によって、実際にコインやサイコロを振るよりも大幅に時間を短縮して事象を提示することに有効である。また、授業の実施でもたらされる作業の効率化は、そのまま教材作成の効率化にもつながる。

3.1.2. 数学と情報の相補性

数学理論とプログラミング言語の相補性が、数学的理解を高めることに効果的であったかについ

では、今回の検証授業では確証的な事象は得られなかった。「ある数学理論に対応する関数が、定義や定理に従って、どのような他の関数の組み立てで成立するか」を理解するためには、生徒にR言語の構文の語彙が十分に身につけていることが必要である。例えば、Rにおける要約統計量の計算は、主に「配列」というデータ構造に対して行うものであり、配列を扱うことができなければ要約統計量を求めるに至らない。標本調査では、母集団となる配列（ここでは X と置く）に対して、無作為抽出を行うための関数を理解し、抽出データを別の配列（ここでは x_i と置く）に格納して標本データを作成する必要がある。さらに、標本平均の意味づけを深く理解するにあたって、複数の標本 x_1, x_2, \dots, x_n それぞれの平均の平均が母集団の平均に近づくことを確認するために、別の配列が必要である。

元来、プログラミングスキルの習得には、問題に対する数学的な理解が前提にあり、手続きに加えて、データを管理することが求められるものである。今回、極めて短期間の検証授業では、数学とプログラミング言語とを同時進行で学習するという目的の達成に、授業者らが想定した以上の負担を生徒に強いていたと察する。現に中学・高校いずれの検証においても、生徒たちがコマンドを入力することに終始し、全体像をつかむことが難しかったように見える場面があり、実際に生徒からも「入力作業が難しかった」という反応があった。

3.1.3. 生徒の負担

生徒の負担の問題は、数学的理解の面では往々にして授業者らの説明不足が引き起こしたものである。視覚的な提示に対して生徒の関心が得られ、プログラミングの手続きと自動的に得られる正解を事実として認識させることはできたが、半面、手続き（プログラミングソース）の意味づけの理解に至るまでには、時間の不足を実感した。

情報学的側面からは、生徒の情報リテラシーの差が授業の進行に大きく影響することを実感した。特にタイピングスピードの影響は顕著である。中学生の参加者は平均的にタイピングが遅く、当初、作業の完了を待って授業を進行したところ授業の消化に大きく遅れを取ってしまった。高校生の参加者は比較的タイピングが速く、南中学校で実験授業を行った時よりもスムーズに授業を進めることができた。また、Rはタイプミスや構文エラーを補う機能はなく、初めから入力し直しとなるので、その点に苛立ちを抱く生徒も少なくなかった。検証授業では極力回避したディレクトリ概念も、実際に扱うこととなった場合は生徒の負担を大きくすると考える。こうした情報リテラシーや、プログラミングの概念を理解させるために、別途時間を用意すべきというのは授業者らの偽らざる実感である。

ただし、タイピングが速い生徒が、遅い生徒に教える場面があったり、一度理解した手続きを他の問題に応用したりといった場面が確認できており、コンピュータを使った授業の中でも生徒同士での学び合いや主体的な学習は成立すると考える。実際の現場では、教師が授業の進行に従ってテキスト化を行うことで、生徒が作業し、教師が説明をするべき場面を明確化でき、授業の進行の想定はしやすくなると思う。また、テキスト化に際しては関数入力を少なくして、補足を多くする工夫が必要である。

3.1.4. 授業に対する生徒の関心

ICTの導入にかかわらず、遊戯性が強く、グループワークを伴う授業企画は、生徒の関心を高める

のに効果的であった。遊戯性の強い内容は、プログラミングによって作成したツールとの親和性が高く、ICT を導入した授業の強い特徴づけとなる。

検証授業では、中学生向けの3時間目「宝くじの買い方を考える」や、高校生向けの4時間目「打率を用いた野球の打順の戦略」がこれにあたる。「宝くじの買い方を考える」では、生徒達に宝くじを連番またはバラで何枚買うかを考えさせ、宝くじの抽籤機をプログラムで再現して、当籤金額や手元の残金の多さを競い、宝くじの当籤金の期待値はくじ1枚の値段よりも低く、できるだけくじを買わないことが理論上損をしない戦略であることを学習した。「打率を用いた野球の打順の戦略」では、実際の野球選手の打撃成績から作成した単打や本塁打などの確率的なデータを資料とし、生徒達を4つのグループに分けて、ドラフト形式で選手を各グループに分配し、打率と打順がどのように野球の得点に結びつくかを考えながら打順を編成して、各グループが編成したチームを試合形式のプログラム上で対戦させた。このような授業形式は、生徒ごとの戦略の比較や、議論の発露に効果的であり、学習意欲を高めることができると考える。

3.2. 監修者による評価

本節では、第三者的な立場から、稚内南中学校と稚内大谷高等学校のそれぞれ4回の講義に対して、生徒がRのコマンドのタイピングに要する時間や、現行の確率統計の単元においてRのコマンドの意味と確率統計の数学的理解が有機的な結合がスムーズに行うことができるか、また各単元に割り当てられている授業時間数の中でRを用いた授業を行うことが可能であるかなどを中心に評価する。

3.2.1 時間的な制約

実験授業では4回の講義数という制約上、分散の説明に時間を割くことができなかつた。正規分布の関数の意味は高等学校の数学Bの中で扱う際でも理解することは困難が伴うことが予想される。分散の大きい場合と小さい場合のヒストグラムの形状を比較させるなどして、離散的な説明での解説を行い、分散の意味を理解させる方法がある。例えば数学と英語の試験で平均点が同一のケースの場合に、あらかじめ教員側で分散を変えて乱数で作りに出したデータを学生に提示し、学生がそれらのデータを基にヒストグラムを作成する。分散が異なれば平均値がたとえ同じでも、ヒストグラムの形状が異なり、その形状の特徴を視覚的に伝えるのが効果的であると思われる。Rでは平均値と分散の値を定めて乱数によるデータを作り出す場面がある。分散の意味を認識させる準備にはどうしても数回の講義での解説が必要になる。

3.2.2. 情報リテラシーによる制約

実験授業に参加した中学校の学生は、Rのコマンドのタイピングの作業に関しては、我々が当初予想したほど困難を感じてはいないようだった。授業者らの評価にあるように実験授業では、確かに学生がコマンドのタイピングに終始してしまう場面はあったが、4回の実験授業に無理矢理にトピックを入れざるを得なかつたのが影響していると思われる。コマンド文は一見すれば長く難しいものと感じるかもしれないが、Rのコマンドをタイピングする作業だけなら、電卓で計算する作業とあまり変わりはないものも多い。R言語のプログラムの背景説明を一切省いて、要所でRを用いて授業を

展開するだけなら極めて容易なことである。

学習の到達目標が、コンピュータを導入して可視化し、確率統計の理解を促進させるというものなら、すぐにでもRの授業への導入は可能であると思われる。Rは、プログラムとしても言語構造がシンプルであり、かつ数学的思考を反映したものであり、プログラムの学習には最適な言語の一つである。プログラム言語構造の解説を省き、コンピュータを導入した授業を行い、成果が得られたとして満足したとしても、ICT教育の底上げに貢献しているとはとても言い難い。初等・中等教育段階におけるプログラミング教育の導入を現行の数学の授業で行う場合は、R言語の構文の解説や、R言語に関する練習問題などの課題を、確率統計の授業の中で積極的に取り入れる必要がある。また初等・中等教育段階におけるプログラミング教育の授業が新たに新設科目として導入される場合も、プログラミングの実践的応用課題として、確率統計が重視されると思われ、いずれのケースでも確率統計の単元にプログラミングの内容を盛り込む必要が生じることは確実であると考えられる。

3.2.3. それぞれの学年における教育効果

中学2年生の確率の最初の授業では、学生がサイコロを振って、一つの目の出る割合が徐々に1/6に近づくことを実際に数百回とサイコロを振って確かめることが行われている。手作業で50分の授業の中で学生が千回以上サイコロを振って実験することは不可能であるが、Rであれば、乱数でサイコロ出る目を何千回、何万回でも再現することができ、手作業では不可能な回数の試行の観察も可能になる。サイコロの1の目の出る確率が1/6に近づく様子が体験でき、授業を受ける学生の理解への貢献が期待できる。

中学3年生で扱う標本調査では、要約統計量の考え方をあえて導入することを提案する。Rを用いて標本平均の平均を求めることを繰り返して、母平均に値が近づいていくことを確かめることで標本調査の考え方を理解させることができる。教科書で扱ってはいない内容ではあるが、上記のRの一連の作業を通して母集団と標本との関係を理解し、R言語の構造を理解する上でも良い例題であると言える。

高等学校の数学Iのデータの分析では、四分位範囲、四分位偏差および相関係数を新たに学ぶ。四分位範囲、四分位偏差は、データの数が多い場合、他の計算ソフト同様、Rでも3つの四分位数はデータを下から25%ずつに区切る数と考える。データの数が小さな場合には、データ数が偶数、奇数で四分位範囲、四分位偏差の定義を明確にした上で箱ひげ図などを描画させることは重要な作業である。しかし同時にデータの数が非常に多い場合に、25%ずつに区切った定義の仕方があることや、その定義での計算ソフトの結果による箱ひげ図からデータを読み取る学習もまた重要である。データ分析の場合は当然、データ数が非常に多いと考えるのが普通だからである。Rの関数には、四分位範囲、四分位偏差や箱ひげ図を求めるのに対応する関数があり、データ分析において、その特徴を瞬時に理解することが可能となる。このような力量を養うことが統計ではより重要なことである。

高等学校数Bでは、確率分布や統計学推測の単元の中で、二項分布、正規分布、母平均の推定など本格的な統計を学ぶことになる。しかし全国のほとんどの高等学校では、この単元が大学入試の出題から外れることから、数Bの確率分布や統計学推測の単元を学習する学生は皆無に等しい。確率統計学の学問的な重要性から、中学・高等学校の単元に確率・統計が新たに加わったはずだが、高

等学校の修了段階で、代表値、分散、相関係数のみの学習に留まる学生が大半というのが現状である。したがって大学の確率統計学の講義でより踏み込んだ内容が展開できるようになったとは全く言えていないのが実情である。確率統計学の重要性を考えるなら、今後は数学 B の確率分布や統計学推測の単元は必修にすべきであろうと考える。

3.2.4. 参観者の評価

稚内南中学校の検証授業を参観した数学教諭からは、数学の授業にも ICT 技術を取り入れることは非常に重要であり、今回の実験授業の内容はソフトウェア選定も良く、面白い授業だったとの感想を得た。R を用いた授業では、生徒がコンピュータを使う前にプロジェクターやコンピュータを通して教師が一度例を示してから、生徒がタイピングをする方が良いなどの意見も頂いた。例えば、2 時間通しの授業として企画し、1 時間目は教師が例を示し、2 時間目に生徒たちが実践するという授業形式である。稚内南中学校からは今回の取り組みに高い関心を寄せていただき、今後、検証授業と同様の実践を継続的に行っていきたいとの意向があったことを記す。一方で、稚内北星学園大学の坪内晃教授は、未知のソフトウェアの操作方法を習得することに時間を多く割くとなれば、検証授業の体系は数学で行うべきか、技術・情報で行うべきか議論が分かれるであろうとの見解を示している。

4. ICT による数学教育の変化

数学授業に ICT を導入することは、授業内容に大きな影響を与える。たとえば、代表値は中学 1 年生の範囲であるが、自分でデータをつくり、整理するといった作業が ICT によって容易になったことで、対象となるデータの分布の特徴は不可分となると考える。中学生向けの検証授業では対象データとして身長を扱ったことで、自然現象の説明として正規分布の理解は避けて通れないと考え、正規分布の性質と代表値の関係を図示した資料を提示し、R の正規乱数によって身長のデータを作成して、代表値をそれぞれ求めた。さらに、この身長のデータから抽出したデータと、最初に作り出したデータを比較し、標本調査と関連付けて指導した。正規分布は中学数学の範囲外ではあるが、生徒達は興味を持ちながら学んでいたように見えた。よって、正規分布は発展的な学習として中学生で学ぶことによって、良い知識を得られる可能性があると感じた。一方で、高校生向けの検証授業で不偏分散を扱ったのは、R が母集団の分散や標準偏差を関数として備えていないことが背景にあり、不偏分散から母分散の説明を組み立てねばならなかったことで、十分な説明ができなかった。関数を実装していない場合は関数を定義すればよいが、数学理論を優先し内容が濃密であったことも相まって、プログラミング言語の学習に踏み込むに至らなかった。このようにソフトウェア側の制約が授業内容に影響を与えることも踏まえておく必要がある。

ICT を導入した数学の授業は、イメージしづらいものの可視化や、情報収集とその分析・活用において、授業時間内の作業を短縮することに貢献する。効率化によって得られた余裕は、生徒とのコミュニケーションに活用でき、生徒の興味・関心を喚起することにつながる。稚内南中学校では、はじめに確率・統計に対するイメージや、どのような場面で役に立つのかといった発問を行って、確率・統計が生活に密接に関わっていることを伝え、生徒たちは学習意欲を持って協力し合いながら発問に答えていた。稚内大谷高校の授業でも、生徒たちは授業者らの発問に活発に反応し、興味・関心を持

って主体的に学習できている姿を見ることができた。

中学・高校のコンピュータ室の規模で、実際に授業を行うにあたり、各端末に R を導入する作業はスムーズに進めることができた。ただし、当然ながら生徒あたり 1 台の端末を確保することは必要であり、今後、生徒 1 人に 1 台の端末を提供する政策が実現したとしても、R を利用するためには Android や iOS のタブレットではなく、Windows や Mac OS が動作することが必須である。

4.1. 留意すべき点

授業準備に際しては、数学理論と R 言語との関連づけを考慮し、数学理論と情報理論をバランスよく説明する資料を作成したうえで、中学および高校であれば 45~50 分の授業に適切な配分が求められる。図 2 は、数学理論と R 言語の相補的理解の事例を示す。図 2 の a) において、配列 x の平均を求める場合①および②の処理は、どちらも正しく平均値を返す。それぞれの意味づけとして、②は要素の総和を要素数で割るという平均の定義に適い、①は平均の関数であることを理解させるべき文脈かつ②の正答を示すものである。一般的な留意点として、新出の関数をプログラミング語彙の理解につなげ、定義や手続きの説明が R 言語を媒介として成り立つように資料が作成されなければならない。図 2 の b) は、配列 x の母分散を求める事例である。R には母分散の関数は実装されていないが、①および②の処理によって、どちらも正しく分散を返す。「無いものをつくる」という実践は「『公式の暗記』と『変数の代入』」にとどまらない理解の醸成に役立つと考える。また、数学の公式を、計算の優先順を理解して、手続きとして記述し直す創造力を養うことは、制約を分析して、問題を解決する筋道を見つけることに役立つと考える。

数学の授業に R を導入するにあたって、最も懸念される点は R 言語の習得であろう。著者らが R を選定した背景は、本学の学生のプログラミング技術の実情が影響している。数値計算法の講義の実施にあたってはプログラミングが要求されるが、1~2 年生で学習してきた一般的なプログラミング

a) 平均の考え方

```
x <- c(170,160,150)      #{170, 160, 150} の順列を意味する配列を x に格納 (代入)
mean(x)                 # ① mean...平均値を求める関数
sum(x)/length(x)       # ② sum...総和を求める関数
                        # length...配列の要素数を求める関数
```

b) 分散の考え方

```
x <- rnorm(100, mean=50, sd=3)
                        # rnorm... 正規分布に従う乱数発生関数
                        # データ 100 個 (平均 50、標準偏差 3) を x に格納
sum((x-mean(x))^2)/length(x)
                        # ①定義に基づく母分散を求める手続き：^ はべき乗の演算子
var(x)*(length(x)-1)/length(x)
                        # ② var... 不偏分散を求める関数
```

図 2. 数学理論と R 言語の相補的理解の事例

言語では、変数宣言のような数学の解法と関連の薄い概念や、静的なデータ構造、アドレス・変数の参照といったコンピュータ寄りの概念について、情報を専攻していない学生の多くが理解できていないことを確認していた。そこで、プログラミングスキルの差異を無くすために新規の言語を用いることを考え、動的なデータ構造と数学理論に対応した関数をもつ R を採用して講義を実施したところ、教職を専攻する学生に対して一定の効果が確認できたことで、この体系を中学・高校の数学授業に広げていくという着想に至ったものである。

現場の教師が R を一から習得する場合、研修などの機会を設ける必要がある。さらに、指導内容を明確にすると同時に、学習内容に応じてどのような関数入力を行うべきかなど、具体的なテキスト化を行う必要がある。

本学の講義における試行は、R 言語では、数学授業の中で並行してプログラミング構文を学習できることを確認した。この事実は、中学・高校の授業においても同様に、生徒が数学理論とプログラミングを同時に学習できる授業の可能性を暗示するものである。しかしながら、現実問題として、R 言語の習得を数学の授業の中に組み込むことは、授業時数に影響を及ぼす。そのため、より弾力的に指導計画を編成し、数学の時間の中に R 言語を習得する時間を 1～2 時間程度、授業計画の最初の段階で組み込む必要があると考える。

4. 2. GNU R を導入した確率・統計の指導計画および指導案

4.1 節で考察した ICT 導入による数学授業の留意点をふまえ、著者らは現在、中学・高校で R を用いた確率・統計の授業を行う場合の指導計画案を検討している。参考までに、表 2 に指導計画を示す。

5. おわりに

本研究では、IT 人財育成を見据えた中学および高校の数学授業の新たな方向性として、著者らの提案による、R を導入した確率統計の授業について検証を行った。検証では、稚内南中学校および稚内大谷高校の協力のもと授業を企画して、生徒、授業者、そして監修者の評価をまとめ、提案に対する効果を確認した。

検証授業に対する生徒の評価から、著者らが企画した授業の形式は、情報が得意な生徒に対しては効果的であるという感触を得たが、数学または情報に対して苦手意識を持つ生徒からの興味関心を得るには至らなかった。しかしながら、本稿に挙げた幾つかの留意点を考慮して授業を行えば、現在、確率・統計の単元で割り当てられている授業時間数で、R を導入した授業は可能であり、それらの授業は学生の学習理解促進に寄与するであろうという感触は得られた。高等学校の数学 B では、授業開講を選択した場合、授業時間数は 45 時間にも及ぶが、R を導入することで授業時間数を大幅に短縮できる可能性があると予想している。この単元の必修化にあたっては R などの ICT を導入した授業で単元の全体の見通しをより明解に示し、授業時間数を圧縮することができるかが鍵になると思われる。

政府の世界最先端 IT 国創造宣言の中で初等・中等教育段階におけるプログラミング教育の導入が謳われ早期に実現されようとしている。世界最先端 IT 国創造宣言では初等・中等教育段階で導入されるプログラミングに関する授業内容は、プログラミングの授業が新設されるのか、それとも数学な

表2. Rを用いた確率統計学習指導計画（コンピュータ室向け）

a) 中学1年…7章 資料の散らばりと代表値（全10時間）

時数	項目	指導内容	用語	Rで使用する関数
1	章の導入	小学校で学んだいろいろなグラフについて復習する。 2つの資料を自分の考えた方法で整理して比べ、分かったことを説明する。		
2	度数分布	資料の範囲や最大値・最小値について理解する。 資料を度数分布表に整理して、その分布の様子を調べる。	最大値、最小値 範囲（レンジ） 階級、階級の幅 度数分布 度数分布表	table() max() min() range()
1	ヒストグラム	度数分布表を基にして、ヒストグラムや度数折れ線をかき、資料の傾向を読み取る。	ヒストグラム（柱状グラフ） 度数折れ線	hist()
1	相対度数	相対度数について理解し、相対度数を使って2つの資料の傾向を比べる。	相対度数	
1	代表値	平均値、中央値、最頻値について理解し、状況に応じて適切な代表値を用いる必要があることを理解する。	代表値（平均値、中央値、最頻値）、階級値	mean() median() names(which.max(table())) sort()
1	近似値と有効数字	近似値や誤差、有効数字の意味、 $a \times 10^n$ の表現方法について理解する。	近似値 誤差 有効数字	
2	コンピュータを使っての統計資料の活用	テーマを決め資料を収集・整理し、コンピュータを用いて傾向を読み取り、レポートを基にして説明し伝え合う。		hist() table() mean() median() names(which.max(table())) sort() range()
1	まとめ			

b) 中学2年…6章 確率 (全9時間)

時数	項目	指導内容	用語	Rで使用する関数
1	章の導入	実際にサイコロを振って、出る目の数が均等に予想し、多数回の実験を行い、サイコロの出る目の数を確かめる。		
2	確率	サイコロの出る目を予想し、コンピュータを用いて多数回の実験を行い、そのことを確かめる。	確率	<code>sample(c("omote", "ura"),100, replace = TRUE)</code> <code>saikoro<-ceiling(runif(6,0,6))</code>
2	確率の求め方	起こり得るすべての場合が同様に確からしいときの確率の求め方を理解する。 確率の範囲や余事象の起こる確率について理解する。	同様に確からしい	<code>sample(c("omote", "ura"),100, replace = TRUE)</code> <code>saikoro<-ceiling(runif(6,0,6))</code> <code>table()hist()</code>
3	いろいろな確率	樹形図や二次元表を用いて場合の数を求め、いろいろな確率を求める。 確率を用いてくじ引きの引く順番と当たる確率の関係を調べ、そのことを説明する。	樹形図	
1	まとめ	問題演習を行う。		

注) ここで

```
sample(c("omote", "ura"),100, replace = TRUE)
```

```
saikoro<-ceiling(runif(6,0,6))
```

について、1つめの処理は、omote および ura の2値を持つ集合(配列)に対して関数 sample を適用100回の無作為抽出をおこなうことによって、コイントスのメタファを実現している。2つめの処理は、runif が実数範囲での一様乱数を発生する関数であり、runif の引数は順に発生回数、乱数の下限値、乱数の上限値を示す。0～6の範囲で発生する乱数値に対して ceiling で天井関数を求めることによって、サイコロのメタファを実現している。

c) 中学3年…8章 標本調査 (全7時間)

時数	項目	指導内容	用語	Rで使用する関数
1	章の導入	いろいろな標本調査を例に、実生活の中で標本調査が使われていることを理解する。		
2	全数調査と標本調査	標本調査の必要性と意味を理解する。	全数調査, 標本調査, 母集団, 標本, 抽出, 推定	sample() append() mean()
1	標本調査による推定	標本調査における無作為抽出の必要性を理解する。 標本の大きさと標本平均の信頼性の関係を理解する。	無作為抽出, 乱数 標本平均	
2	標本調査の活用	コンピュータを用いて標本調査を行い、母集団の傾向をとらえ説明する。 標本の平均の平均を出すことによって、母集団の平均に近づくことを理解する。		
1	まとめ・章の問題	問題演習を行う。		

d) 高校1年…数学 I 第4章 データの分析 (全13時間)

時数	項目	指導内容	用語	Rで使用する関数
2	いろいろなデータの分析	度数分布表について理解すること。 ヒストグラムについて理解すること。	度数分布表 階級、階級幅、階級値 度数、相対度数、累積度数、累積相対度数 ヒストグラム 単峰性、多峰性 分布の歪み	round(rnorm(n=120,mean=163,sd=8)) table() hist()
5	データの傾向のとらえ方	代表値の意味について理解すること。 代表値を用いてデータの傾向を把握し説明すること。 四分位数、四分位範囲、四分位偏差などの意味について理解すること。 四分位数、四分位範囲、四分位偏差などを用いてデータの傾向を把握し説明すること。 箱ひげ図の意味について理解すること。 データの散らばりの大きさを数値で表し、傾向を読み取ること。	代表値 (平均値、中央値、最頻値) 5数要約、四分位数 (第1、第2、第3) 範囲 四分位範囲 四分位偏差 箱ひげ図 偏差、偏差平方分散、標準偏差 変量	round(rnorm(n=120,mean=163,sd=8)) mean() median() names(which.max(table())) var() fivenum() IQR() boxplot() range()
1	問題演習			
3	データの相関	散布図と相関係数 相関係数とその意味 表計算ソフトによる相関係数の計算 相関係数と因果関係の違い 2次元の度数分布表による変数間の関係の調べ方	散布図 正負の相関関係 共分散 相関係数 偽相関 2次元の度数分布表	cor() plot() var()
1	問題			
1	章末問題			

注) ここで

```
round(rnorm(n=120,mean=163,sd=8))
```

について、rnorm は正規乱数を発生する関数であり、平均値163、標準偏差 8 の正規分布に従う乱数を120個生成して、関数 round で丸めを行うことによって、120名の身長についてサンプルデータを作成することができる。また、

```
names(which.max(table()))
```

は、R で実装していない最頻値を求める手続きである。関数 table は引数とした配列の度数分布表を作成し、関数 which.max では、最大の度数をとる階級値およびその度数のベクトルを抽出する。このベクトルに対して、「見出し」を提示する関数 names で階級値を取り出すことによって最頻値を得る。

高校2年…数学B 第4章 確率分布と統計的な推測 (全45時間)

時数	項目	指導内容	用語	Rで使用する関数
3	確率変数と確率分布	確率変数 簡単な場合について、確率分布を求め、表やグラフに表わす	確率変数 確率分布	
2	確率変数の平均	確率変数Xの平均の意味 Xを1次式で変換した時の平均の性質	平均 期待値	mean()
4	確率変数の分散と標準偏差	確率変数Xの分散と標準偏差の意味とその性質 確率変数Xを1次式で変換した時の分散と標準偏差 簡単な確率分布について、Xの標準偏差が求められること	分散 標準偏差	var() sd()
4	和の平均と分散	確率変数の和の平均の性質 確率変数の独立 独立な確率変数の積の平均 独立な確率変数の和の分散	確率変数の独立	mean() var()
3	二項分布	二項分布の意味 二項分布の平均、標準偏差とその計算	二項分布	rbinom() dbinom()
1	問題			
3	連続的な確率変数	連続的な値をとる確率変数 連続的な値をとる確率変数の確率分布	連続型確率変数 離散型確率変数 確率密度変数 分布曲線	
8	正規分布	正規分布 正規分布による二項分布の近似	正規分布 標準正規分布 正規分布表	curve(dnorm(x),from=-4,to=4,xlab="",ylab="") rnorm()
1	問題			
8	母集団と標本	標本調査の考え方 母集団と標本 母平均、母標準偏差と標本平均 標本平均の標準化	全数調査 標本調査 母集団、標本、抽出 標本の大きさ 乱数さい、乱数表 無作為抽出、任意抽出、 無作為標本、 復元抽出、非復元抽出 母平均、母標準偏差 標本平均	sample() mean() var() sd() sample() rnorm() runif()
6	推定	母平均の推定 母比率の推定	推定 区間推定 信頼度、信頼区間 母比率	
2	章末問題			

注) ここで

curve(dnorm(x),from=-4,to=4,xlab="",ylab="")

について、関数 curve は曲線の描画であり、関数 dnorm は標準正規分布の確率密度関数を提示する。この処理を実行することによって、引数 from および to で与えた ± 4 の範囲で標準正規分布の分布形状を視覚化できる。

どの既存の授業の中で展開すべきであるかなどの指針はまだ示されていない。著者らの研究は、現行の数学の授業体系の中にプログラミングを導入し ICT 教育を行うための試みである。ICT を導入することで、数学の単元である確率統計の理解が阻害されるようであれば ICT 導入の意義はない。ICT の導入にあたっては学生の学習理解の促進に寄与するためのものであるという前提に立っていることが条件である。

今後は、教師の R 言語習得のプロセスを検討し、数学と情報が苦手な生徒に対して、効果的な授業を提供できるように授業形式の改善を行うことが課題となる。

●参考文献

- (1) エミール・デュルケーム著, 佐々木交賢訳, 「教育と社会学」, 誠信書房, 1982.
- (2) 文部科学省, 「教育の情報化ビジョン～21 世紀にふさわしい学びと学校の創造を目指して～」(http://www.mext.go.jp/b_menu/houdou/23/04/_icsFiles/afieldfile/2011/04/28/1305484_01_1.pdf), 2011.
- (3) 内閣官房内閣広報室, 「世界最先端 IT 国家創造宣言の変更について」(<http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20140624/siryoul.pdf>), 2014.
- (4) 総務省編, 「平成25年版 情報通信白書 ICT 白書」, 2013.
- (5) 日本経済新聞, 「ビッグデータ分析に人材の壁、25万人不足見通し I T 各社、育成へ本腰」(<http://www.nikkei.com/article/DGXNZO57421630X10C13A7EA1000/>), 2013.

●英文タイトル

A new direction of mathematics teaching method for human resources development of IT professionals – Evaluations and opinions to experimental lessons by using the GNU R –

●英文要約

"The big data era" has come. In Japan, for global competitiveness and increasing business, the industries and the government have a serious shortage of data scientists and analysts. The educational circle are entrusted with a mission for fostering the manpower. Especially, the probability theory and the statistics become important. In late years, the government is reshaping the education policy for making good use of information science. The math education will be encouraged to big change by the policy.

The authors proposed a teaching method of the probability theory and statistics for the junior high school and high school level by using the GNU R statistics analysis software. The authors expect that the programming language R shows mathematical theories as an intermediary for better understanding the math.

In this year, the authors' proposing is proved by practical lessons. Students' evaluations to the lessons are analyzed by the principal components analysis. The authors' (role of teachers and overlooker) and visitors' opinions make ideas for developing the teaching method.

● Key Words

Digitization of education

Probability theory and statistics

Big data

