

プログラミング活用による 統計教育への問題解決型教材の開発に向けて

神部 順子¹⁾ 玉田 和恵²⁾

要 旨

「絶対の正解のない問いに対して、自ら立ち向かう力をつける」問題解決型のデータサイエンス教育の指導法及び教材開発に取り組んでいる。単に統計学を教えるのではなく、ICT ツールを活用した情報の思考・判断・表現能力を高めることを目的とし、特に、専門分野・学習者の学力到達度・地域性に対応すること、かつ、プログラミング活用によるデータサイエンス教育のための教材の開発を目指す。

キーワード：ICT 問題解決力 データサイエンス 縦糸・横糸モデル

1. はじめに

あらゆる分野の大学で、データサイエンス教育を実施することが求められている。これはSociety5.0時代に対応し産業競争力の源泉となるハイレベルなIT人材を育成・確保するためと言われている。そこで、さまざまな大学で実践可能なカリキュラム及び教材を開発する必要がある。現在、大学で行われている統計教育をベースにデータサイエンス教育が検討されているが、従来の統計教育では目的とするデータサイエンス教育を実現することは困難と考えられる。つまり全国の大学で、誰にでも同じようなカリキュラムや教材を提示しても、学習者の問題意識や、学力到達度によって、同様な効果が得られるとは考えられない。そのため筆者らは専門分野・学習者の学力到達度・地域性に応じたデータサイエンス教育のためのカリキュラム及び教材開発が求められていると考える。また、2020年小学校段階からプログラミングが必修化されることに伴い、各学校段階でどのような資質・能力を育成するのか、目標・内容・評価方法を明確化するための検討をすることも重要な課題である。具体的には「絶対の正解のない問いに対して、自ら立ち向かう力をつける」問題解決型のデータサイエンス教育の指導法及び教材開発である。

筆者らは、私立大学情報教育協会の情報リテラシー教育ガイドライン作成の一環で、縦糸・横糸モデルに基づくICT問題解決力育成カリキュラムを開発してき

た。その具体的指導計画の1つとして、統計的データ分析を題材に、松田・竹村・玉田(2019)が問題解決型の統計教育モデルを提案した。ここで、現状の高校数学「データ分析」や大学の統計教育について、統計知識の実践的活用力を十分に育成できておらず、改善が必要だと指摘している。その原因の1つに、目標とするパフォーマンスを発揮すべき学習活動をさせていれば、自ずと能力が身につくとの行動主義的な考え方があるとしている。そこで、統計教育を問題解決力の育成という観点から指導することを提案している。統計学に基づく統計教育が、理論的に正しい方法を演繹的に導き実行するのに対して、ICT問題解決力として指導する統計教育は、多様な分析方法を発想し、不適切な事例からの類推で帰納的に実行すべき方法を限定しつつ、ICTを活用して多様な方法を試し、結果を相互比較して、異なる結果が出る原因を考察し、より説得力のある結果に絞り込むものである。

今回は筆者らが所属する文系の学科の専門教科において、当該分野で必要となる内部知識を明確にし、問題解決の縦糸横糸モデルを適用して、指導法及びカリキュラムを検討することとした。具体的には学生の現状を把握し、明示的に統計を題材とした問題解決を指導することの効果について検証する。特に、プログラミング活用による統計教育への学生の反応を中心に追っていくこととした。

2. 方法

2.1. 調査対象と時期

本研究の対象は、2019年度学科専門科目「統計学概論」を受講する大学2年、3年生67名(履修登録者数)

2020年1月25日受付 2020年2月3日受理

1) 江戸川大学情報文化学科／情報教育研究所

2) 江戸川大学情報文化学科／情報教育研究所

表1 授業の流れ

回と 日付	内容
1 9/23	オリエンテーション 数字に強くなる
2 9/30	アンケート調査の基本
3 10/7	調査を行うための注意 計算力を強くする
4 10/21	Rのインストール 基礎統計量とは
5 10/28	Rでの基本的な操作確認 図や表を読む ゲーミング教材へのアクセス
6 11/11	表を読みこなすには
7 11/16	2つの項目を要約する 散布図
8 11/25	度数分布表・ヒストグラム データの標準化
9 12/2	基礎的な統計知識に関する確認 Rコマンドのインストール
10 12/9	相関に関するゲーミング教材(第1回目)
11 12/16	エクセルによるデータ分析 (シートのコピー/単純集計/データの種類による分析の違い/クロス集計/アンケート調査の際のデータ収集について)
12 12/23	統計知識修得状況の確認 (データの種類/グラフと表のそれぞれ/代表値や散布度の確認)
13 1/16	相関に関するゲーミング教材(第2回目) Rによるデータ分析(クロス集計表の完成を中心に)
14 1/20	クロス集計の検定 自らアンケート調査用紙を作成する

である。特に今回は第9回から第14回までの6回分(2019年12月～2020年1月)の講義として実施したものを分析する。授業の流れを表1に示す。

2.2. 指導内容

現在の学生は、コンピュータを操作することが日々当たり前前の行動となっている。表計算ソフトについても、すでに大学入学前に学んできている(学んだつもりでいる)。しかし、答えが用意されている例題が予め提示され、単に操作として作業を繰り返している結果、多様な分析方法を発想することの大切さなどについての指導はなされていない。

そこで、今回、2019年度前期に大学院レベルで試行されたゲーミング型のe-learning教材(山口2019)について、文系の学部生向けに改良したものを使用して実践を行った。問題解決の手順を明示的に示し、目的に

応じてデータ分析を進める考え方を指導している。これは問題解決の流れの理解とクロス集計を活用した分析方法を修得させることに特化した内容となっている。

授業では、開始時に授業の目的や作業の良さについて示した上で、実習をし、終了時に自分の作業について振り返りをしたコメントおよび、当日データ分析した内容を回収した。それらの提出状況について、整理していった。なお、学生らは1年次より、問題解決の縦糸・横糸モデルについての講義は受講済みであるが、問題解決に統計手法を活用することについて初めて学んでいる。また、プログラミング活用としては、R言語を用いて授業を展開した。その際、表計算ソフトエクセルとの違いや、コマンドを自分で打ち込み作業をしながら学習するものと(表1の第4回と第5回)、RのパッケージソフトRコマンドを用いて分析作業を行うもの(13回と14回)として展開していった。ほとんどの学生にとっては、R言語およびRコマンドについては別の授業でも触れられており、全く初めての作業ではなかった。

3. 調査結果

ゲーミング型のe-learning教材については3回実施した。そのうち、第10回と第13回でストーリー展開は同じものである。第5回では、教材へのログインの失敗やアクセスそのものが出来なかった学生もかなりおり、コメントをみると、途中で止まってしまったことへの不満がいくつかあった。内容以前の問題となったため、改善の必要があることがわかった。そこで、第10回と第13回で用いた教材については、その経験を踏まえた改良を加えたもので実施した。ICTツールを活用することによって情報の思考・判断・表現能力を高めることが目的であったにも関わらず、むしろ、機器操作の説明に時間を取られてしまうことへの対応は今後の課題となった。また、第10回と第13回の内容についての学生のコメントについて表2および表3に示す。第10回については、ゲーミング教材に関する感想を挙げている。第13回は同じゲーミング教材であるため、そこへの理解の深まりがあったかについてのコメントを挙げている。

第11回では、①シートのコピーをすることにより、データ分析後の内容と混同しない、②エクセルの機能ピボットテーブルを用いることでクロス集計が出来るようになる、③質的データと量的データの分析方法の違いを理解することを中心に展開し、ほとんどの学生が理解を深めたとコメントしている。しかし、第12回に、基本的な知識を確認として、平均値を求める際の

表2 第10回ゲーミング教材に関する感想

番号	内容
1	画面の表示について、前のページの問いが何だったかを忘れてしまった結果、現在のページで何を答えたらいいのかがわからなくなった
2	エクセルデータをダウンロードができて、とりあえず集計してみる作業になることに戸惑った
3	ゲームとして自分が選んだ選択肢が間違いと指摘してくるが、何が間違いなのかわからず、説明がほしかった(2名)
4	クロス集計をするための2項を選ぶのは直観的に選んだ結果、たまたま正解であったため、理論的に説明できるようになりたかった
5	ゲーム教材の最後にログが表示されたので、振り返ることができたのがよかった

ポイントを自らの言葉で書かせてみたところ、52名中21名しか正確に示すことは出来なかった。これは知識が定着していないことを示している。

さらに、第13回のRを用いたクロス集計についての提出物(40名)では、複数の項目についての正しい集計及びコメントを回答したものは20名である。集計結果を貼り付けてもコメントがない学生は13名、不十分な分析となった学生は7名である。このことから、学生の多くは操作の習得のみに満足しており、本研究が目指す「自ら目標を立て、分析方法を決め、実施し、振り返る」作業には至っていないことが明らかになった。

この科目は11名の留学生在が履修登録している。留学生は日本語(特に漢字の読み書き)に困難さを伴っているため、第10回では出席した7名全員が「ゲーミング教材について、何をやっているのかわからなかった」と述べている。第13回ではそのうちの2名は理解できたとコメントしている。

また、最終回である第14回ではカイ二乗検定についての解説を行い、クロス集計表で整理されているデータを入力表に落とし込む作業およびRコマンドによる分析について追加の解説した。学生の提出を確認したところ、49名中8名を除けば内容を理解した上での作業が完了していた。またこの回で初めて扱ったカイ二乗検定について10名がコメントしていた。その内容を表4に示す。

学生のコメントをみていくと、全般的には回を重ねるごとに理解は深まり、データサイエンスについて学習の動機付けは高まっていった様子が見られる。このことは、問題解決の縦糸・横糸を用いて問題解決の流れを明確に意識し、ゲーミング教材やこれまでのエクセル以外のソフトとしてRを用いたことにより、これまでとは異なった学びとして関心を寄せ、そこでの操作によつての達成感を得ることによる効果があったも

表3 第13回の授業コメント

番号	内容
1	年明け久しぶりの講義でこれまで何をやっていたのか思い出すのに時間がかかった(2名)
2	Rコマンドでエクセルデータをインポートする際の注意点がやっと飲み込めた
3	エクセルのピボットテーブルを学んだ際より今回のRコマンドによるクロス集計することによって、質的データと量的データの違いを意識して分析した(3名)
4	アンケート作成にはデータがどのように分析したのかを考えてから取り組むべきだと思ふようになった
5	なぜクロス集計が大切なのかを実感できた(2名)

表4 カイ二乗検定を学習しての感想

番号	内容
1	Rコマンドの操作方法習得が不十分のまま終わった(3名)
2	カイ二乗検定に関する計算が難しい(2名)
3	復習して、他にもやってみたい
4	やっと慣れてきた
5	エクセルでもカイ二乗検定ができるようになりたい
6	数字を絶対値にする意味がやっとわかった
7	他の検定方法も知りたい

のと考えられる。

しかし、本来意図した到達点までには、まだ遠い結果であったとも言える。学生からの自由な発想を導き出す授業展開への工夫や、考察を重ねる重要性の説明についても、さらに検討を重ねる必要がある。

4. まとめと今後の課題

今回は、ゲーミング教材およびR言語を用いたデータサイエンス教材に取り組んだが、機器の操作方法の習得へ意識が奪われている学生がかなり見受けられた。

2022年度より高等学校ではプログラミングとデータサイエンスを扱う「情報I」が共通必修科目となる。これらのことを踏まえ、今後、高大連携を見据えたプログラミング活用によるデータサイエンス教育について検討する必要がある。そのためには、高等学校までの情報教育の実態を把握し、そのことに対応した大学の学びになるべきである。今回は2年生および3年生と異なる学年が受講する科目での試みであったが、実際は学年間の差ではなく、これまでの学びの差があった。このことへ配慮するべきであったと考える。さらには、大学での学びのゴールをどこに設定するかについて、

より明確にする必要がある。本研究は、単に従来型の統計学を教えるのではなく、特に、専門分野・学習者の学力到達度・地域性に対応したデータサイエンス教育のためのカリキュラム及び教材を開発することを目指す。そのために、まず文系の大学生を対象に、問題解決の縦糸・横糸を用いて統計を問題解決に活用する力を涵養する指導法を開発する必要があることも明らかになった。

謝辞

本研究の実施にあたり、日本学術振興会科学研究費補助金(基盤研究(C)No. 17K1145001代表: 神部順子)の助成を受けたものである。また、科学技術融合振興財団(FOST)助成金(課題名: ICT問題解決力育成の

ための情報リテラシー教育モデルとゲーミング教材の開発)の支援を受けた。ここに記して感謝する次第である。

参考文献

- 松田稔樹・竹村徳倫・玉田和恵(2019)「問題解決の縦糸・横糸モデルに基づく統計教育の改善(1)－教育目標の明確化とカリキュラム設計－」, 日本教育工学会研究会報告集, JSET19-1, 495-502
- 山口敏和・竹村徳倫・小原裕二・神部順子・玉田和恵・松田稔樹(2019)「問題解決の縦糸・横糸モデルによる統計教材の開発～二変数間関係」, 日本教育工学会2019年秋季全国大会講演論文集, 235-236