

統計数理研究所共同研究レポート 466

統計教育実践研究

第 15 卷

2023 年 3 月

統計数理研究所

統計数理研究所共同研究レポート 466

統計教育実践研究

第 15 卷

2023 年 3 月

統計数理研究所

東京都立川市緑町 10-3

電話 (050) 5533-8500

この共同研究レポートは、2022 年度統計数理研究所共同研究（研究課題番号：2022-ISMCRP-5007）による研究成果報告書として作成されたものである。

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the author(s).

統計教育実践研究 Vol. 15

目次

I. 重点課題「ICT を活用した統計学教育とその評価に関する新たな展開」

1. Educational software を活用した統計教材の開発に関する基礎的研究
細田幸希（筑波大学大学院） 1
2. ICT ツールを活用した統計ワークショップの実践
ー 経済産業省「未来の教室：STEAM ライブラリー」ー
町田保（リトルスタジオインク株式会社） 9
3. オンライン生徒参加型プラットフォーム『センサス@スクール』の紹介と授業での導入事例
竹内光悦（実践女子大学）・末永勝征（鹿児島純心女子短期大学）・渡辺美智子（立正大学） 24
4. ICT を活用したロバストな工程設計における教育的効果（家庭科授業での事例？）
田中正敏（松本大学） 26
5. 統計リテラシー教育の世界的潮流と評価
深澤弘美（東京医療保健大学）・及川久遠（大和大学） 29

II. 【講演】高等学校『情報 I』の内容紹介

- いやにさせない統計・グラフを使う QC サークル教育
榎本竜二（中央大学） 35

III. 統計データ分析コンペティションを用いた統計教育

1. SSDSE と統計データ分析コンペティションの概要
田中雅行（独立行政法人統計センター） 37
2. 体力が基礎学力に与える影響について
水谷順一（フェリス女学院高等学校）・太佐美結（フェリス女学院高等学校） 39
3. 都市部と地方の教育格差の要因と課題
鈴木淳子（愛知県立一宮高等学校）・森下達也（愛知県立一宮高等学校） 41
4. ボランティア活動の決定要因
橋本想吾（慶應義塾高等学校）・今泉開（慶應義塾高等学校） 43
5. 市町村というミクロ的視点から投票率の実態を探る
渡辺安之（茨城県立並木中等教育学校）・林蔚欣（茨城県立並木中等教育学校） 45

IV. 統計教育の研究および授業実践事例 I

1. 統計的に探究することができる児童が育つ授業の在り方
日比野浩規 (名古屋市立神宮寺小学校) 47
2. データの特徴を適切に表現する力を育む学習指導
～第1学年「わかりやすく せいりしよう」の実践を通して～
白石円 (岩手大学教育学部附属小学校) 57
3. 統計的問題解決力の育成を目指した指導の在り方～標本調査の実践を通して～
稲垣道子 (岩手大学教育学部附属中学校) 59
4. 教師による授業への統計的探究プロセスの実装の試み
佐藤寿仁 (岩手大学) 68
5. 学校設定科目「データサイエンス」と課題研究との接続
林兵馬 (神戸大学附属中等教育学校) 72
6. 高校選択授業におけるデータサイエンス教育の実践事例
小森晴菜・馬場国博 (慶應義塾湘南藤沢中等部・高等部) 73

V. 統計教育の研究および授業実践事例 II

1. 探究への連携を見据えた数学の授業としてのデータサイエンス教育の試み
中田雅之 (神戸大学附属中等教育学校) 76
2. 「情報 I」統計ポスターを用いたデータサイエンス教育実践
～データの活用×情報デザインの体系化～
林宏樹 (雲雀丘学園中学校・高等学校) 78
3. 子どものワクワクを育てる「数式抜きのデータサイエンス」～身近なデータから世界を推測～
都丸希和・佐藤健太 (名古屋大学教育学部附属中・高等学校) 80
4. 「仮説検定」への接続を目指した「仮説検定の考え方」の教材開発とその実践
光永文彦 (西大和学園中学校・高等学校、大和大学)・大浦弘樹 (東京理科大学)
・吉川遼 (名古屋文理大学)・伏木田 稚子 (東京都立大学) 84
5. 高等学校における仮説検定にかかわる指導について—統計学習と探究活動の迫間をどう埋めるか—
稲葉芳成 (立命館宇治中学校・高等学校) 86
6. ICT を活用した推測統計の指導とその評価
橋本三嗣 (広島大学附属中・高等学校) 88
7. 情報 I におけるデータサイエンス教育の前倒し実践と今後の改善に向けて
荒川智浩 (国際基督教大学高等学校) 90

VI. 統計教育の研究および授業実践事例 III

1. 医療メッシュ統計を活用したデザイン力の育成
和泉志津恵 (滋賀大学)・佐藤彰洋 (横浜市立大学) 94
2. 公開型アダプティブオンラインテストのアナリティクス：
大学の基礎数学での COVID-19 以前と以降
廣瀬英雄 (久留米大学、中央大学) 100
3. 特異値分解から見た項目反応理論の新評価：大学数学の CBT によるテスト
廣瀬英雄 (久留米大学、中央大学) 109
4. プログラミングによる統計計算の実装を演習とする初等統計学教育の試み
西山慧 (京都大学教育学研究科) 116
5. 授業で使うデータに関する試みと提案
島谷健一郎 (統計数理研究所) 120
6. 高校生の探究活動の方法としての統計
橋本三嗣 (広島大学附属中・高等学校) 123

VII. 【講演】高等学校『情報科』の現状と大学教育への展開

1. 国公立大学の2025年度入学者選抜における教科「情報」の対応調査に関するレポート
下山恭平・大戸祥也（株式会社ラーンズ マーケティング・営業部） 125
2. 2025年度に入学する学生の統計スキルについて
大橋真也（順天堂大学） 140

VIII. 【講演】2022年度日本統計学会統計教育賞受賞者受賞講演

- 中等教育におけるデータサイエンス教育の実践と高大連携
林兵馬（神戸大学附属中等教育学校、第18回日本統計学会統計教育賞受賞者） 144

IX. 統計教育の研究および授業実践事例 IV

1. 【講演】基礎統計の教授法
狩野裕（大阪大学） 145
2. 文系学生のためのデータサイエンス教育：経済学部における実践と課題
田中昌宏（福岡大学） 147
3. どのように共通教育と専門教育の「統計学」の橋渡しをするか II？
－2013年度と2022年度の学生に対する質問紙調査に基づく考察－
関哲人・速水孝夫（北海学園大学）・米田力生（金沢大学） 148
4. 仮説検定・統計解析によるアンケート分析手法の実践報告
片江康裕（東京都立大島海洋国際高等学校） 154

X. 参考資料

- 「中高生・スポーツデータ解析コンペティション」について 158
日本統計学会統計教育分科会のお知らせおよび入会案内 159

Educational software を活用した統計教材の開発に関する基礎的研究

筑波大学大学院人間総合科学研究科・細田幸希

〒305-8572 茨城県つくば市天王台 1-1-1

E-mail : hosoda.koki.wy@alumni.tsukuba.ac.jp

1. 本研究の目的と方法

AI 実装社会の到来によって、学校教育におけるデータから意思決定・行動につながる統計的思考や判断、表現力を国民が広く身に付けるための教育の実効性の向上が指摘されている（日本学術会議数理科学委員会数学教育分科会，2020）。その中でも、テクノロジーを活用した学習環境の充実は、議論すべき課題の一つであると考えられる。

日本の統計教育研究及び教科書におけるテクノロジー活用の実態を調査した福田・大谷（2022）によると、日本では統計教育独自のテクノロジーである Educational software の開発が進んでいないことが課題の一つとして明らかにされている。また、Educational software では、ランダムな実験のモデルを構築しながらシミュレーションを実行したり（Biehler et al., 2013）、機械学習の手法を用いてデータ探索を行ったり（細田, 2022）することが可能であるにも関わらず、それらの機能を活用した教材開発や実践に関する研究もほとんど行われていない。

したがって、それらの課題を解決するために、Finzer（2017）によって開発された CODAP と呼ばれる Educational software に着目して教材開発を行う¹⁾。このような教材を開発することは、日本におけるテクノロジーを活用した統計教育研究及び実践に関する一つの指針を与えることにつながる。

以上より、本稿では、CODAP に焦点を当てて、Educational software を活用した教材を開発し、日本の統計教育への示唆を導出することを目的とする。上記の目的を達成するために、本稿では、先行研究を基に教材を開発し、解決過程の検討や生徒が作成したワークシート及びレポートの分析を通して開発した

教材の教育的価値を明らかにする。

2. 開発する教材の視点

(1) 日本の統計教育研究及び教科書におけるテクノロジー活用の課題

福田・大谷（2022）は、統計教育におけるソフトウェアツールの視点（Biehler et al., 2013, pp.650-653）を用いて、国内の先行研究や現行の学習指導要領に対応した小中学校の教科書分析を行い、テクノロジー利用を視座とする日本の統計教育の教授学習的課題を明らかにした。その結果、日本の統計教育研究では、R などの統計分析のためのコンピュータプログラムである Statistical software packages や、stathist などの Applets としてのテクノロジー活用がほとんどであり、シミュレーションとデータ分析の機能が内蔵された Educational software がほとんど活用されていないことが指摘されている。また、教科書内で扱われているテクノロジーに関して、Excel などの Spreadsheet が中学校で扱われているのみで、ほとんど扱われていない。さらに、小中学校の教科書の QR コード内で扱われているテクノロジーにおいても、Educational software は存在しないことが明らかにされている。

加えて、日本の統計教育におけるテクノロジーの活用例では、シミュレーションの側面が強調されていないこと（福田・大谷, 2022）や、機械学習の手法の指導において Python などのプログラミング言語による実装の側面に限定されていること（細田, 2022）が指摘されている。前者のシミュレーションに関するテクノロジーが活用されていないことは、不確実な事象から合理的な判断や意思決定を行うために必

要な統計と確率の接続が十分になされていないことを意味している。また、後者に関しては、機械学習を含めたデータサイエンスの手法の理解を目的とした場合、手法自体の仕組みや構造に関してブラックボックス化してしまうことが危惧される。したがって、これらの課題に対して Educational software の活用の観点から解決の方法を議論する必要がある。以上より、シミュレーションの実行と機械学習の手法の理解を目的とした Educational software の役割に基づき教材を開発する。

(2) Educational software を活用した教材の視点

Educational software とは、児童生徒の統計学に関する抽象的な概念やプロセスの理解を支援するために、様々な統計に関するソフトウェアプログラムを兼ね備えた柔軟性のある動的ツールを指す (Biehler et al., 2013)。Educational software の一般的な特徴として、ドラッグ・アンド・ドロップでグラフに変数を入れることで、簡単に一変数の分布と変数間の関係を探索することが可能な点が明らかになっている。以下では、さらに、前節で述べたシミュレーションの活用と機械学習の手法の一つである決定木の手法に関する理解に焦点を当てて、Educational software の特徴を明らかにし、本研究の教材の視点を設定する。

① Sampler によるシミュレーション機能を通じた標本分布の考えを理解する方法

近年の国際的な統計教育研究では、仮説検定や区間推定などの形式的な方法を用いずに、標本と母集団の関係及び標本抽出に関わる統計的概念 (例: 標本変動) とプロセスを関連付けながら、学校教育の早期から推測統計に必要な能力をボトムアップ型で育成することを目的とした「非形式的な統計的推測 (Informal Statistical Inference)」に関する研究が注目されている (Makar & Rubin, 2018)。また、Manor and Ben-Zvi (2015) は、ISI の背後にある論理を理解するためには、標本分布の考えを含む標本抽出を繰り返して行うプロセスを想像すること挙げている。ここでの標本分布の考えとは、統計量の値が何らかの確率

の範囲で分布しているという意味である。つまり、Manor らの研究では、このような標本分布の指導を通して、学習者が不確実性を制御したり定量化したりするために、標本統計量を確率的に記述する確率的見方 (probabilistic view) を育むことが重視されている。その背後には、標本分布は統計学を学習する上で難しい概念の一つであり、学習者の多くが計算等の手続的な知識を身につけることで終始してしまうことで、標本分布を深く理解することができていないという課題がある (Manor & Ben-Zvi, 2015)。

Manor らの研究では、この課題を解決するために、上述の確率的見方や標本変動といった標本分布の考えを、Educational software を活用しながら非形式的に学ぶことの必要性を指摘している。

非形式的に標本分布の考えを学ぶための Educational software の活用方法の一つとして、Sampler によるシミュレーション機能がある (図 1)。CODAP では、Sampler に予め母集団のデータを CSV ファイルに変換して取り込むことができ、図 1 のように学習者が Sampler を変更できないように教師がロックすることもできる。実際に、学習者は Sampler の標本の大きさ (Select) と標本の数 (Collect) を任意に変更しながら、繰り返し標本抽出を行う。そして、抽出された標本データは連動して作成されるデータ表に記録され、それらのデータの変数をドラッグ・アンド・ドロップでグラフに可視化しながら、標本分布の特徴や性質を探索することができる。

② Arbor を用いた決定木 (分類木) の構造を探索する方法

Beihler et al. (2022) によれば、産業界や社会科学などの多くの分野におけるビッグデータの普及によって、従来の統計教育からビッグデータの探索や意味の理解を目的としたデータサイエンス教育への発展の必要性が指摘されている。特に、多変数データを探索を通してデータの理解を深めるための新たなスキルとして、機械学習の手法の一つである決定木 (Decision tree) の指導が注目されている (Beihler &

Fleischer, 2021). 一方, 日本の統計教育では, これまで決定木の手法が数理学意決定力の育成に有益な方法として挙げられていたが, 授業内の効果的な指導方法については解明されていなかった(青山・清野, 2018). そこで, 細田(2022)は, 海外の先行研究の分析を通して, モデル構築とモデル評価の観点から決定木の指導方法を明らかにした.

決定木には分類木(Classification trees)と回帰木(Regression trees)の2つの手法が存在し, 特定のデータに基づいて分類または予測(回帰)を行うために, データを再帰的に分割するアルゴリズム(例: CART)を用いて, 階層的な樹木状のモデルを構築する(Breiman et al., 1984). 本稿では, 特に分類木に着目して教材開発を行う²⁾. まず, モデル構築の場面では, CODAP上で分析するデータ(CSVファイル)を開いた後, 「Arbor」と呼ばれる決定木の学習ツールに関するURL(<https://codap.xyz/plugins/arbor/>)を入力し取り込む. Arborには, RやPythonなどの専門的なデータ分析のツールとは異なり, 自動的に最適な木の分割やサイズを行うアルゴリズムはなく, 学習者がドラッグ・アンド・ドロップ操作によって分類木モデルを構築するように設計されている. ゆえに, Arborによる学習では, ただ一つの最適な分類木モデルを導き出すのではなく, 学習者がデータから探索的にモデルを構築しながら分類木の構造を理解することが目的とされている(細田, 2022).

次に, モデル評価の場面では, 決定木で用いられる評価に関する測定値の複雑さから, 混同行列を用いた誤分類率(misclassification rate)や感度(sensitivity)が採用されている(細田, 2022). また, 先行研究では, データを訓練データとテストデータに分けてモデルの一般性を評価する交差検証の導入に関する議論もなされている(Beihler & Fleischer, 2021). 実際のCODAPのArborでは, 上述の誤分類率や感度をモデル構築とあわせて計算し表示するアルゴリズムがあり, 学習者はモデル構築と評価を容易に行うことができる. したがって, CODAPのArborを活用するこ

とで, 学習者自身の価値観に基づいて探索的にモデルの構築や修正, 評価を行うことが可能となり, RやPythonなどのプログラム言語を用いた実装前の素地指導の方法として期待できる(細田, 2022).

以上より, 本研究における教材開発の視点として, 「学習者がCODAP上で統計グラフの作成や統計量の計算を行うだけではなく, Samplerによるシミュレーションを通して標本分布の考えを理解したり, Arborを用いて決定木の構造を探索したりできるようにすること」を設定する.

3. 教材の開発

本章では, 前章で設定した教材開発の視点を基に, 「推定に必要な標本の大きさを考えよう(教材1)」と「インターネットの利用調査(教材2)」を開発した. 教材1は, 細田(2021)で開発された好きなスポーツ観戦の教材を標本分布の探索や確率的見方の育成の観点から精緻化し新たに開発した. 教材2は, Beihler and Fleischer(2021)を参考に, インターネット利用に関する調査データを基づいて, 分類木モデルを構築することを目的として開発した.

以下では, 教材1は実際の授業の概要と生徒のレポート課題の分析結果を示す. また教材2は, 問題場面と教材の解決過程を示す.

(1) 推定に必要な標本の大きさを考えよう(教材1)

① 実際の授業とレポート課題の概要

授業は, 東京都の国立大学附属中学校第3学年の5クラス(201名)を対象に, 2023年1月の標本調査の単元に関する授業の一部として, CODAPのシミュレーションを通した標本分布の探索(2h)及び必要な標本の大きさの決定(2h)の計4時間扱いで行った. 生徒は, 母集団や標本の概念や無作為抽出, 標本の大きさといった標本調査の方法に関する内容については既習である.

まず, 授業前に観戦したいスポーツに関するアンケート調査を行った. 調査は, 性別と次の選択肢(ア: 野球, イ: テニス, ウ: サッカー, エ: バスケットボ

ール, オ: バレーボール, カ: ラグビー, キ: 水泳, ク: 陸上, ケ: その他, コ: 好きではない) の項目から一つ選んで答えるというものである. そして, アンケートに回答した 151 人のデータを母集団とし, 母集団において最も人気があると予想するスポーツの比率を一つ決め (例えば, 野球を選んだ人の割合), それを母集団の性質とみなし, 標本調査を通して推定するために必要な標本の大きさを決定するという問題の文脈を設定した. 問題は次の通りである.

【問題】

標本の特徴や性質から母集団 (151 人のデータ) の性質を推定しよう. また, 母集団の性質を推定するために必要な最小の標本の大きさとその理由について考えてみよう.

まず, 生徒は母集団の性質 (母比率) と推定に必要な標本の大きさに関する仮説を設定した後, 1 人 1 台タブレット型 PC を用いて, 図 1 の CODAP のファイルを起動し, シミュレーションを実行した. 次に, シミュレーションから得られた特定のスポーツに関する標本比率のデータをグラフ等で可視化したり, 統計量を求めたりしながら標本比率の分布の特徴を考察した. 特定のスポーツに関する標本比率の分布の特徴を探索する場面では, 様々な大きさにシミュレーションを行いながら, 生徒は標本の大きさと標本変動の関係を把握しながら考察していた (図 2).

分かった事 (野球) 50 回やれば結果

- 標本の大きさを増やると平均値と中央値の差が縮まる 中央値の位置は
- 大きく外れる値が減っていき → 比率が集まっていく →
- 中央値は 20 のまま (標本の大きさを変えても等しい)
 ↳ 70 で 20.96 にしたものの, あまり変わらない.
- データ所に大量に集まるのではなく, 15~25 の周辺にたまっていく
 (LL' LL'')
- 10 までみで計算していたから, 30 までいくと, 小人数が入ってくる. (あたりね)
- 標本の大きさを 50 を境に似たようなバラツキになった
 ↳ 桁 % が無くなった.

図 2 ワークシートにおける標本比率の分布に関する生徒の考察

生徒自身が標本比率の分布に関する考察と推定を行った後, 授業者がクラス全体で考察した内容を共有しながら, 改めて標本の大きさによる母集団の性質と標本の性質の関係についてまとめを行った. ここまでの授業で, 実際のアンケートデータを用いた標本抽出のシミュレーションを繰り返し行い, そこから得られた標本比率の分布の特徴を基に, 標本と母集団の関係について学習した.

レポート課題は, 標本比率の分布の考察結果を踏まえて, 母比率を推定するために必要な最小の標本の大きさの特定とその理由を考えるものである. これは, ある意味, 標本調査による推定で生じる誤差を

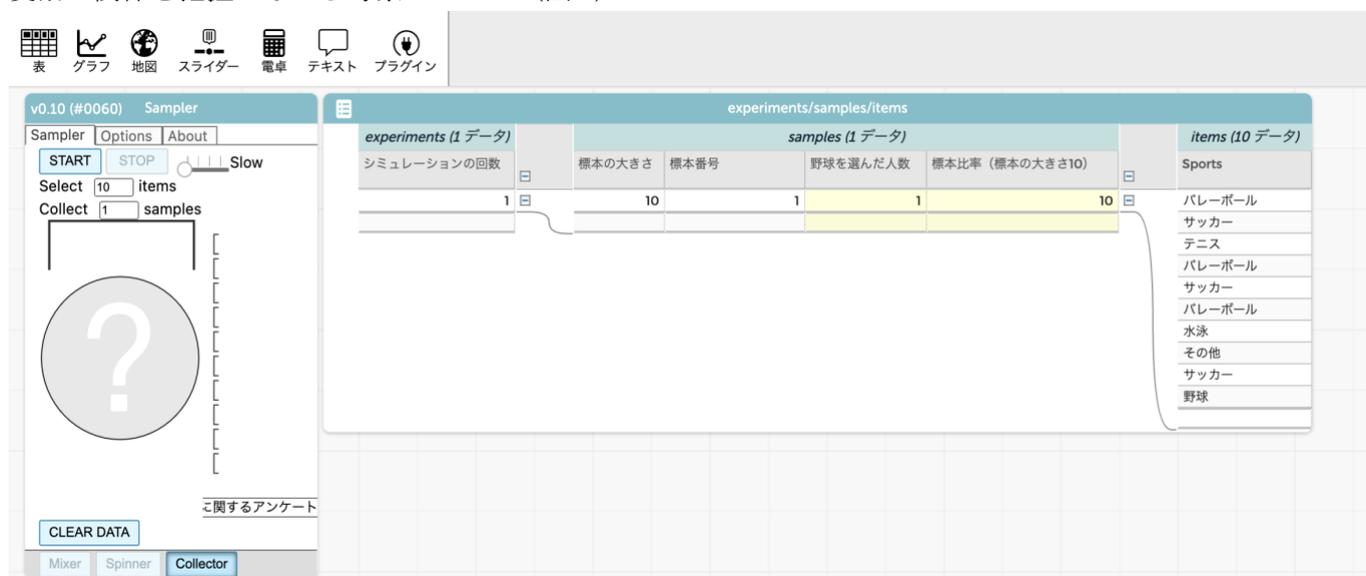


図 1 ソフトウェア CODAP の Sampler によるシミュレーション

確率的に捉え、定量化する問題と言えると考える。このレポート課題の提示にあたって、初めに授業者が生徒に標本分布の探索から得られた考察結果を振り返らせ、標本分布における母集団に関する不変的な特徴(シグナル)と無作為標本に関する偶然性を伴う確率的な変動(ノイズ)の認識を促した。そして、後者のノイズを考慮しながら必要な最小の標本の大きさを特定していくことを確認した。なお、レポートの形式は、授業支援クラウドの一つであるロイロノートテキストを使用した。

② レポート課題の分析

レポート分析では、母比率の推定に必要な標本の大きさの特定に関して、生徒はどのように誤差を認識し、確率的に捉えて定量化して解決していくのかをみていく。このような確率的見方の指針を捉えることで、教材の有効性及び非形式的な統計的推測に関する実態の一端を捉えられると考える。したがって、本稿では、全てのレポートの中から特に確率的見方の指針が明確であるものを分析していく。

以下で示すレポートでは、標本比率の平均値を根拠とする限界を示し、新たに標本比率の分布における相対度数に着目して標本の大きさを特定していく流れがまとめられている。具体的に、まず、CODAPのシミュレーションを用いて、標本の大きさ 10, 30, 50, 70, 100, 130 の標本を 15 個ずつ抽出し、その標本平均をそれぞれ求めている(表 1)。

表 1 標本比率の平均値

標本の大きさ ^o	標本比率の平均 ^o
10 ^o	20.7 ^o
30 ^o	20.7 ^o
50 ^o	20.8 ^o
70 ^o	19.0 ^o
100 ^o	19.7 ^o
130 ^o	19.6 ^o

そして、標本の大きさ 50 と 70 の間で標本比率の平均値が変わっていることまとめている。また、実際

の母比率(野球を選んだ人数の割合は 19.87)と求めた標本比率の平均値を比較している。ここで、生徒は以下のように、標本が母集団から一部取り出されていることによる誤差を考慮する必要があることを記述している。

素直にみるならば標本の大きさが 100 の時が一番正確だと見て取れる。しかし、標本調査というものは標本をピックアップしているので「誤差」はつきものである。つまり、自分がどれくらいの誤差まで許せるかによって、適している標本の大きさは変わるとのことだ。

さらに、標本比率の平均値には、外れ値が含まれている可能性があり、平均値で根拠づけてしまうことは適切ではないことを挙げている。したがって、この生徒は、標本比率の分布を描いてその相対度数を求め、外れ値とみなす範囲が含まれる割合を調べることを通して、外れ値が少ない標本の大きさを特定することを提案している。具体的にこの生徒は、標本比率が ± 5 の範囲を許容範囲とし、それ以外を外れ値とこの生徒は定義している。そして、定義した許容範囲と外れ値の定義を基に、標本の大きさ 10, 30, 50, 70, 100, 130 の場合の標本比率の分布を可視化している(図 3)。なお、図 3 の黒色で表されている相対度数は許容範囲であり、赤色で表されている相対度数は外れ値を表している。

最終的に、図 3 の分布から標本の大きさと外れ値の相対度数の値の関係を捉え、相対度数が 6.7%である標本の大きさ 70 を最低限必要な大きさであると結論付けている。さらに、標本の大きさが 100 と 130 の時の外れ値の相対度数が 0%でありながらも、選択しなかった理由について以下のように記述している。

標本の大きさ 100 と 130 の時は、確かにこれらはどちらも外れ値が 0%と出たのだが、151 人中 100, 130 人を取り出すのは 2/3 以上を取り出していることになるので、「コスト、手間がかかる」。つまり不適正だと考えた。

したがって、この生徒は外れ値の割合(相対度数)

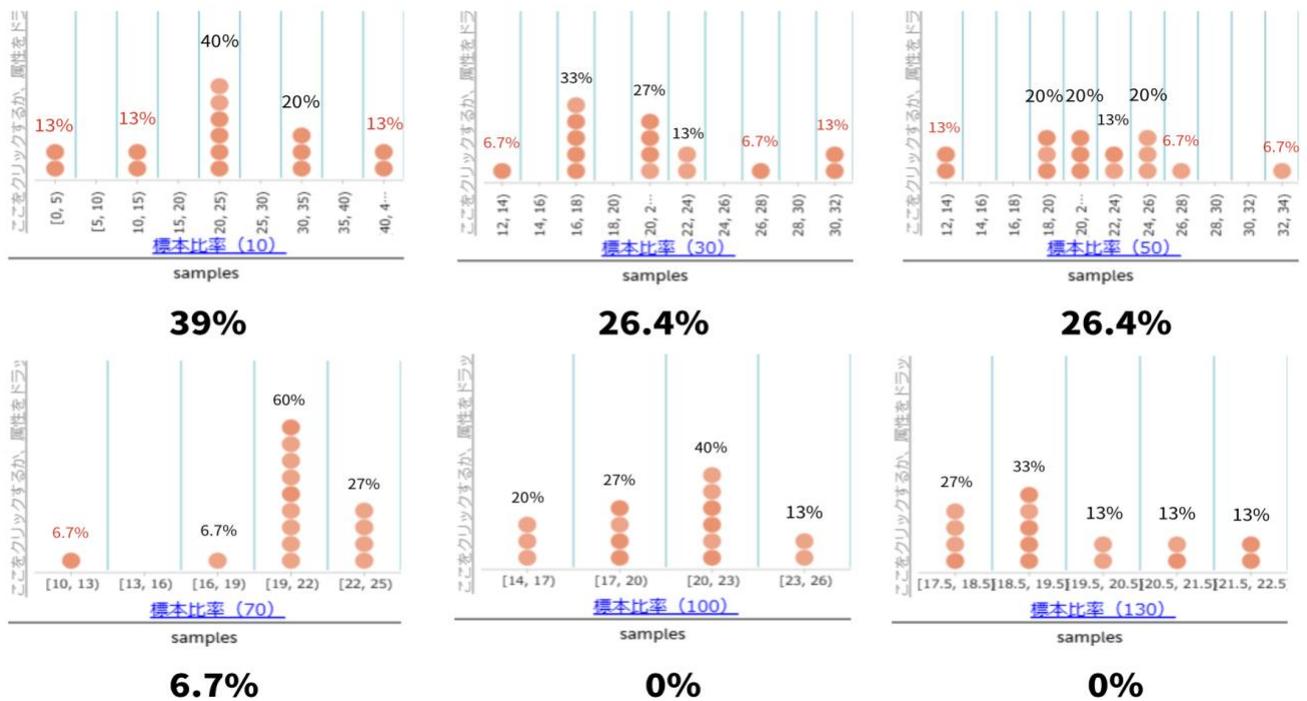


図3 許容範囲と外れ値を表した標本比率の分布

だけでなく、標本調査の意味や位置付けについても考慮しながら、標本の大きさを特定していたと考える。また、得られた標本分布を階級幅で整理した分布へと変換し、その区間に対応する相対度数(統計的確率)の値を求め、それを標本の信頼度として根拠付けることができていることがいえる。

(2) インターネットの利用調査(教材2)

① 問題場面の設定

本教材は、Beihler and Fleischer (2021) や細田 (2022) を基に、日本の文脈での設定を加えたものである。なお、本教材を扱う上で、対象となる生徒は、高等学校における数学I「データの分析」の内容及び数学A「条件付き確率」に関する内容までを学習した生徒を想定している。また、使用するデータは、筆者が収集したデータを用いる。

【問題】

近年、日本ではスマートフォンの普及によって、若者がテレビを見るよりもインターネットを利用する時間が長いということが様々な調査から明らかになっている。そこで、インターネットの利用に影響を及ぼす要因について、決定木(分類木)を作成して調べてみよう。

② 本教材の解決過程

本教材は、2章で述べた通り、CODAPの決定木に関する学習ツールであるArborを用いる。今回のデータの目的変数はインターネットの利用頻度であり、「頻繁に利用する」または「あまり利用しない」である。それに対し、15個の説明変数(性別、Twitter、LINE、インスタグラム、TikTok、YouTubeでゲーム実況に関する動画、YouTubeでファッション系の動画、YouTubeでお笑い芸人の動画、動画配信サービスでドラマや映画、スポーツ配信アプリ、音楽配信サービス、オンラインゲーム、自分専用のスマートフォンを持っているか、自分専用のタブレットを持っているか、自分専用のパソコンを持っているか)があり、生徒がドラッグアンド・アンド・ドロップ操作で影響を及ぼすと考えられる説明変数を選択しながら分類木モデルを構築していく。例えば、性別という変数を分割するルールとして選択した場合の分類木モデルは、図4の通りである。このモデルをみると、全体で117人中96人がインターネットを頻繁に利用していることが分かり、さらに「性別」で分割すると、男性は56人中50人がインターネットを頻繁に利用し、女性は61人中46人がインターネットを頻繁に利用

していることが分かる。

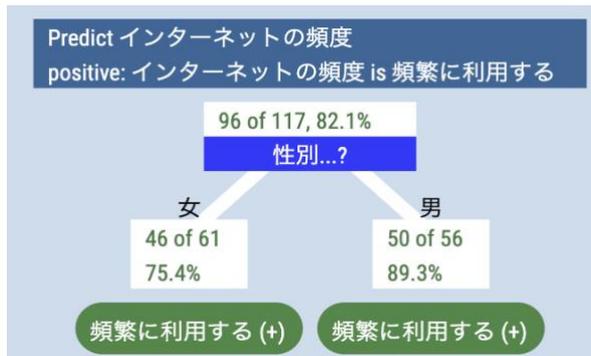


図4 「性別」で分割した場合の決定木モデル

さらに、説明変数を選択して追加していくことによって、決定木モデルを拡張することができる。構築したモデルの評価は、真陽性 (TP)、真陰性 (TN)、偽陽性 (FP)、偽陰性 (FN) から成る混同行列や誤分類率、感度を用いて行う (図5)。

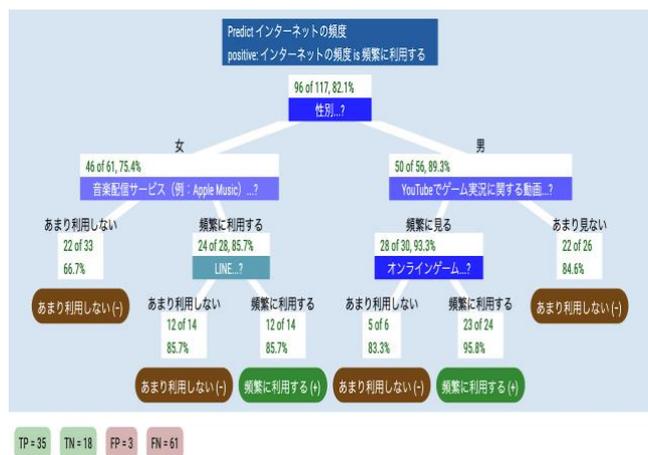


図6 混同行列の表示した決定木モデル

この混同行列では、Arbor上で説明変数を追加するごとに、予測値が上記の4つの指標に基づいて分類されていく。したがって、生徒は混同行列のそれぞれの指標の変化をみながら最適な決定木モデルを構築していく。加えて、誤分類率と感度についてもArbor上で構築した決定木モデルに対して自動で算出し表示することができる。ちなみに、図5の決定木モデルでは、誤分類率は0.547、感度は0.365である。

よって、先ほどの混同行列の指標とあわせて誤分類率や感度の指標を用いることで、決定木におけるモデルの評価の役割や意義を学習することができる。以上より、本教材は、生徒が探索的に決定木のアルゴリズムの理解を促すことに繋がると考える。

4. 考察

まず、教材1の実践と生徒のレポート分析から、標本抽出に関するシミュレーションから得られた標本分布の特徴を把握したり、代表値や相対度数による確率的見方を用いて標本の代表性を評価したりしていたといえる実態がみられた。このことは、CODAP上でシミュレーションの実行と得られた標本データの探索が容易にできることが要因の一つであると考えられる。中学校学習指導要領解説(文部科学省, 2018)においても、コンピュータを用いて母集団から標本を取り出し、標本の傾向を調べる指導が重視されているため、シミュレーションツールを搭載したEducational softwareを活用することには教育的価値がある。また、教材2では、データを用いた探索的アプローチから、決定木という新たな統計的手法の構築と評価の仕組みの理解を意図した新たな学習機会を提供できる面で一定の教育的価値があると考えられる。日本の統計教育では、質的データや量的・質的が混在した多変数データを用いた学習が少ないことを考えると、決定木のようなデータの非線形な関係を基に分類したり予測したりする新たなスキルを子どもたちが習得できるようにすることは、データサイエンスの視点から統計教育の強化が期待できる。

さらに、Educational softwareを活用することで、既存の統計的手法を提供するだけではなく、子どもたちが自身が統計的手法の構築や分析、比較を行うことが可能になる(Biehler et al., 2013)。今回開発した教材を通して、統計教育におけるテクノロジーの役割についても議論していく必要があると考える。

5. おわりに

本研究の目的は、Educational softwareを活用した教材を開発し、日本の統計教育への示唆を導出することであった。そのために、シミュレーションの活用と決定木の手法に関する理解に焦点を当てた教材を開発した上で教育的価値を示した。今後の課題として、今回開発した教材の精緻化と実証的研究を引き続き

進めることが挙げられる。

注

- 1) CODAP は、インターネットを通して無料で誰でも使用できるソフトウェアであり、学校教育でも比較的導入しやすいと考えたため、本研究では採用している。ただし、決定木学習ツールである Arbor は、ソフトウェアの使用言語が英語のみ使用可能になっている（2023 年 2 月 28 日現在）。
- 2) CODAP の Arbor では分類木だけではなく、回帰木モデルを構築するためのプログラムも搭載されている。今後は、回帰木に関する教材開発についても進めていく予定である。

引用参考文献

- Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. In M. A. (Ken) Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Third international handbook of mathematics education* (pp. 643–689), New York, US: Springer. https://doi.org/10.1007/978-1-4614-4684-2_21
- Beihler, R., & Fleischer, Y. (2021). Introducing students to machine learning with decision trees using CODAP and Jupyter Notebooks. *Teaching Statistics: An International Journal for Statistics and Data Science Teaching*, 43, 133-142. <https://doi.org/10.1111/test.12279>
- Beihler, R., Veaux, D. R., Engel, J., Kazak, S., & Frischemeier, D. (2022). Editorial: Research on Data Science Education. *Statistics Education Research Journal*, 21(2), Article 1. <https://doi.org/10.52041/serj.v21i2.606>
- Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.I. (1984). *Classification and regression trees*. Belmont, California: Wadsworth.

- Finzer, W. (2017). Common online data analysis platform, <https://codap.concord.org/> (最終確認日：2023.2.28).
- 福田博人・大谷洋貴 (2022). テクノロジー利用の視座からみた日本の統計教育研究と実践. 全国数学教育学会第 57 回研究発表会 発表原稿.
- 細田幸希 (2021). 中学校数学科におけるインフォーマルな統計的推測を促進するための教材の開発：統合的モデル化アプローチに焦点を当てて. 日本数学教育学会 第 55 回秋期研究大会発表集録, 49–56.
- 細田幸希 (2022). 高等学校段階におけるデータサイエンス教育に関する一考察：決定木に関する学習指導に焦点を当てて, 日本科学教育学会第 46 回年会論文集, 23–26.
- Makar, K., & Rubin, A. (2018). Learning about statistical inference. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in Statistics education* (pp.261-294). Cham: Springer. https://doi.org/10.1007/978-3-319-66195-7_8
- Manor Braham, H., & Ben-Zvi, D. (2015). Students' articulations of uncertainty in informally exploring sampling distributions. In A. Zieffler & E. Fry (Eds.), *Reasoning about uncertainty: Learning and teaching informal inferential reasoning* (pp.57–94). Minneapolis, Minnesota: Catalyst Press.
- 文部科学省 (2018). 中学校学習指導要領 (平成 29 年度告示) 解説 数学編, 日本文教出版.
- 日本学術会議数理科学委員会数学教育分科会 (2020). 新学習指導要領下での算数・数学教育の円滑の実施に向けた緊急提言：統計教育の実効性に向けて. <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-24-t293-2.pdf> (最終確認日：2023.2.28).
- 清野辰彦・青山和裕 (2018). 学校教育における設計科学的視座に基づく数理科学教育に関する研究—数理科学的意思決定をする際に有益となる手法の抽出とその検討—, 日本数学教育学会 第 6 回春期研究大会論文集, 35–40.

ICT ツールを活用した統計ワークショップの実践 — 経済産業省「未来の教室：STEAMライブラリー」 —



リトルスタジオインク株式会社代表取締役 町田保

世界はデータで出来ている～STEAM探究のための統計・データサイエンスの道具箱～

概要



PPDACサイクルに沿った、
データにもとづく課題解決を
身近な気づきから始まる
6つのストーリーで学習



概要

扱われる教科 / 単元

主たる対象

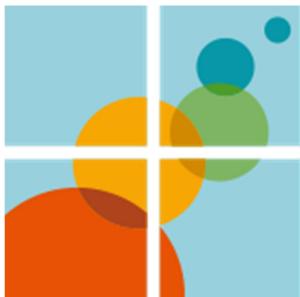
小学校/中学校/高校

- 算数・数学
- 理数
- 理科（物理）
- 総合的な探究の時間
- 情報Ⅰ・情報Ⅱ
- 国語
- 体育
- 家庭科（生活）
- 社会
- 現代社会
- 政治・経済
- 公民 / 地理・歴史

監修

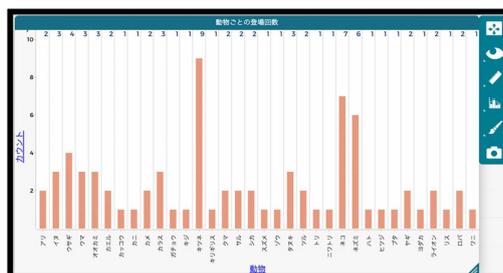
- 高部 勲（立正大学 データサイエンス学部教授・博士（統計科学））
- 竹内光悦（実践女子大学大学院人間社会研究科教授・博士（理学）・統計学会統計教育委員会委員長）
- 春 広計（統計数理研究所所長/筑波大学名誉教授・工学博士）
- 森井良直（宮崎大学教育学部教授・理学博士・日本統計学会統計教育分科会主査）
- 夏辺美智子（立正大学データサイエンス学部教授・理学博士・統計関連学会連合統計教育推進委員会委員長）

データ分析ツール



CODAP

データ分析には、小学生から高校生向けウェブベースのデータサイエンスツール Common Online Data Analysis Platform (CODAP) を活用しています。



レクチャー①

STEAM探究に役立つ PPDACサイクル



関連する科目・単元

算数・数学/理数/情報
総合的な探究の時間

主たる対象

小学校/中学校/高校

①STEAM探究に役立つPPDACサイクル

- 身近な課題(問いを解決するための科学的探究プロセス「PPDACサイクル」)について学習します。

レクチャー 2

小学生のランドセルは 重すぎる!?

ドットプロットと 箱ひげ図で データ分せき



関連する科目・単元

算数・数学/理数/情報
生活/総合的な探究の時間

主たる対象

小学生・中学生

②ランドセルって重くない?

- 小学生にとってランドセルはどのくらい重たいものなのか、ランドセルの重さや小学生の体格のデータをドットプロットや箱ひげ図を用いて分析・比較します。

レクチャー③

どんな動物たちが
物語に出てくる？

ドットプロットと
積み上げ棒グラフで
分せきしてみよう!



関連する科目・単元

算数・数学/理数/情報
国語/総合的な探究の時間

主たる対象
小学生

③物語の中にはどんな動物が出てくるんだろう？

- 物語に登場する数多くの動物の「種類」や「性格」「話の結末」などのデータを集め、積み上げ棒グラフや箱ひげ図を用いて分析します。

レクチャー④

ホームランには
コツがある!?

箱ひげ図・
クロス集計で
ホームランの条件を分析!



関連する科目・単元

算数・数学/理数/情報
生活/総合的な探究の時間

主たる対象
中学生

最近ホームランが増えているってホント？

- 「データの分析」によって、最近メジャーリーグでホームラン数が増加していることについて、箱ひげ図やクロス集計を用いて解説します。

レクチャー⑤



関連する科目・単元

算数・数学/理数/情報
社会/地歴・公民
総合的な探究の時間

主たる対象

中学生

⑤新型コロナウイルスの感染者数の傾向は？

- ・新型コロナウイルス感染者数の増減の「正しい傾向を読み取る」ために、「時系列データの4つの変動成分」について解説します。

レクチャー⑥



関連する科目・単元

算数・数学/理数/情報
社会/地歴・公民
総合的な探究の時間

主たる対象

高校性

⑥暑い日はアイスクリームが売れる？

- ・全国の家計調査データと気象データをもとに散布図を作成し、相関関係を調べることで、どんな気象条件下でアイスクリームが売れているかを分析します。



関連する科目・単元

算数・数学/理数/情報
社会/地歴・公民/技術・家庭
総合的な探究の時間

カレーといえば牛肉？豚肉？

- カレーに使われる肉が日本の各位地域によって異なる特徴をしますことを、都道府県ごとの家計調査データを可視化し、分析します。

物語の中にはどんな動物が出てくるんだろう

授業実践

令和5年2月14日（火）に

東京都目黒区中目黒小学校

6年生クラスに

出前授業を行った実践例になります。

★本資料では、Steam Library掲載のデータになりますが、当日（3 / 1 2）は実際のデータで発表します。

計画表

5W1Hで考えよう

調べる対象(What)

グリム童話、イソップ童話、日本昔話、宮沢賢治作品
物語の種類ごとに20話

調べる方法

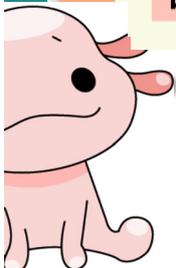
作品を読む

調べる項目

物語の種類、タイトル、動物の役回り、性格、
動物の種類、動物のよび名、動物のけつまつ

調べる日

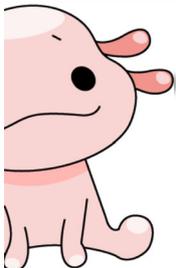
放課後、休みの日



ちょうさ票

物語の中の動物

- タイトル
- 物語の種類
- 動物の種類
- 動物のよび名
- 動物のけつまつ
- 役
- せいかく



データ表

物語の中の動物

事例 (76データ)

索引	タイトル	物語の種類	動物の種類	動物のよび名	動物のけつまつ	役	せいかく
1	馬をうらやんだろば	イソップ	ウマ	馬	どちらとも言えない	主役	すなおな
2	セロひきのゴージュ	宮沢賢治	カッコウ	かっこう	どちらとも言えない	わき役	おひとよしな
3	うら島たろう	日本昔話	カメ	亀	どちらとも言えない	わき役	やさしい
4	しかおどりのはじまり	宮沢賢治	シカ	鹿	どちらとも言えない	わき役	おろかな
5	した切りすずめ	日本昔話	スズメ	雀	どちらとも言えない	わき役	せい実な
6	文福茶がま	日本昔話	タヌキ	文福茶釜	どちらとも言えない	主役	のんきな
7	セロひきのゴージュ	宮沢賢治	タヌキ	狸の子	どちらとも言えない	わき役	せいじつな
8	セロひきのゴージュ	宮沢賢治	ネコ	三毛猫	どちらとも言えない	わき役	いじわるな
9	よだかの星	宮沢賢治	ヨダカ	よだか	どちらとも言えない	主役	すなおな
10	ししの分け前	イソップ	ライオン	獅子	どちらとも言えない	主役	いじわるな
11	馬をうらやんだろば	イソップ	ロバ	ろば	どちらとも言えない	主役	なまけものな
12	いなばの白うさぎ	日本昔話	ウニ	和邇	どちらとも言えない	わき役	いじわるな
13	おく様きつねのけっこん	グリム	キツネ	古狐	どちらとも言えない	主役	おろかな
14	くまと旅人	イソップ	クマ	熊	どちらとも言えない	わき役	おひとよしな
15	おく様きつねのけっこん	グリム	ネコ	猫	どちらとも言えない	わき役	すなおな
16	どんぐりと山ねこ	宮沢賢治	ネコ	山猫	どちらとも言えない	主役	いじっぱりな
17	ねこのじむ所	宮沢賢治	ネコ	猫3匹	どちらとも言えない	わき役	おろかな

データ表

物語の中の動物

事例 (76データ)

索引	タイトル	物語の種類	動物の種類	動物のよび名	動物のけつまつ	役	せいかく
1	馬をうらやんだろば	イソップ	ウマ	馬	どちらとも言えない	主役	すなおな
2	セロひきのゴージュ	宮沢賢治	カッコウ	かっこう	どちらとも言えない	わき役	おひとよしな
3	うら島たろう	日本昔話	カメ	亀	どちらとも言えない	わき役	やさしい
4	しかおどりのはじまり	宮沢賢治	シカ	鹿	どちらとも言えない	わき役	おろかな
5	した切りすずめ	日本昔話	スズメ	雀	どちらとも言えない	わき役	せい実な
6	文福茶がま	日本昔話	タヌキ	文福茶釜	どちらとも言えない	主役	のんきな
7	セロひきのゴージュ	宮沢賢治	タヌキ	狸の子	どちらとも言えない	わき役	せいじつな
8	セロひきのゴージュ	宮沢賢治	ネコ	三毛猫	どちらとも言えない	わき役	いじわるな
9	よだかの星	宮沢賢治	ヨダカ	よだか	どちらとも言えない	主役	すなおな
10	ししの分け前	イソップ	ライオン	獅子	どちらとも言えない	主役	いじわるな
11	馬をうらやんだろば	イソップ	ロバ	ろば	どちらとも言えない	主役	なまけものな
12	いなばの白うさぎ	日本昔話	ウニ	和邇	どちらとも言えない	わき役	いじわるな
13	おく様きつねのけっこん	グリム	キツネ	古狐	どちらとも言えない	主役	おろかな
14	くまと旅人	イソップ	クマ	熊	どちらとも言えない	わき役	おひとよしな
15	おく様きつねのけっこん	グリム	ネコ	猫	どちらとも言えない	わき役	すなおな
16	どんぐりと山ねこ	宮沢賢治	ネコ	山猫	どちらとも言えない	主役	いじっぱりな
17	ねこのじむ所	宮沢賢治	ネコ	猫3匹	どちらとも言えない	わき役	おろかな

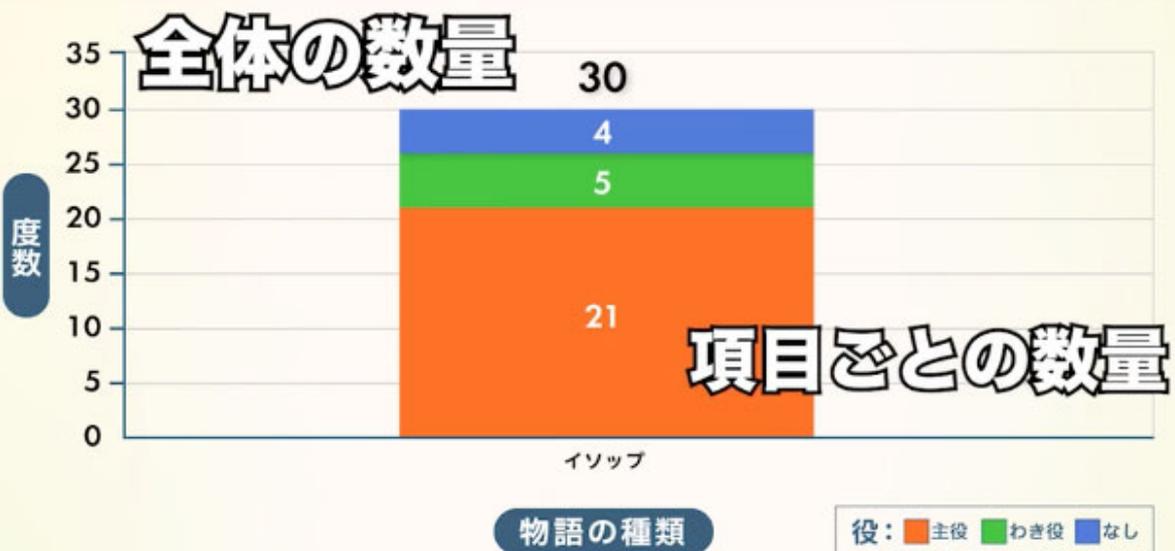
動物の役どころ（集計表のドットプロット）



T2_0326

Little Studios, Inc. 積み上げ棒グラフ

動物の役どころ



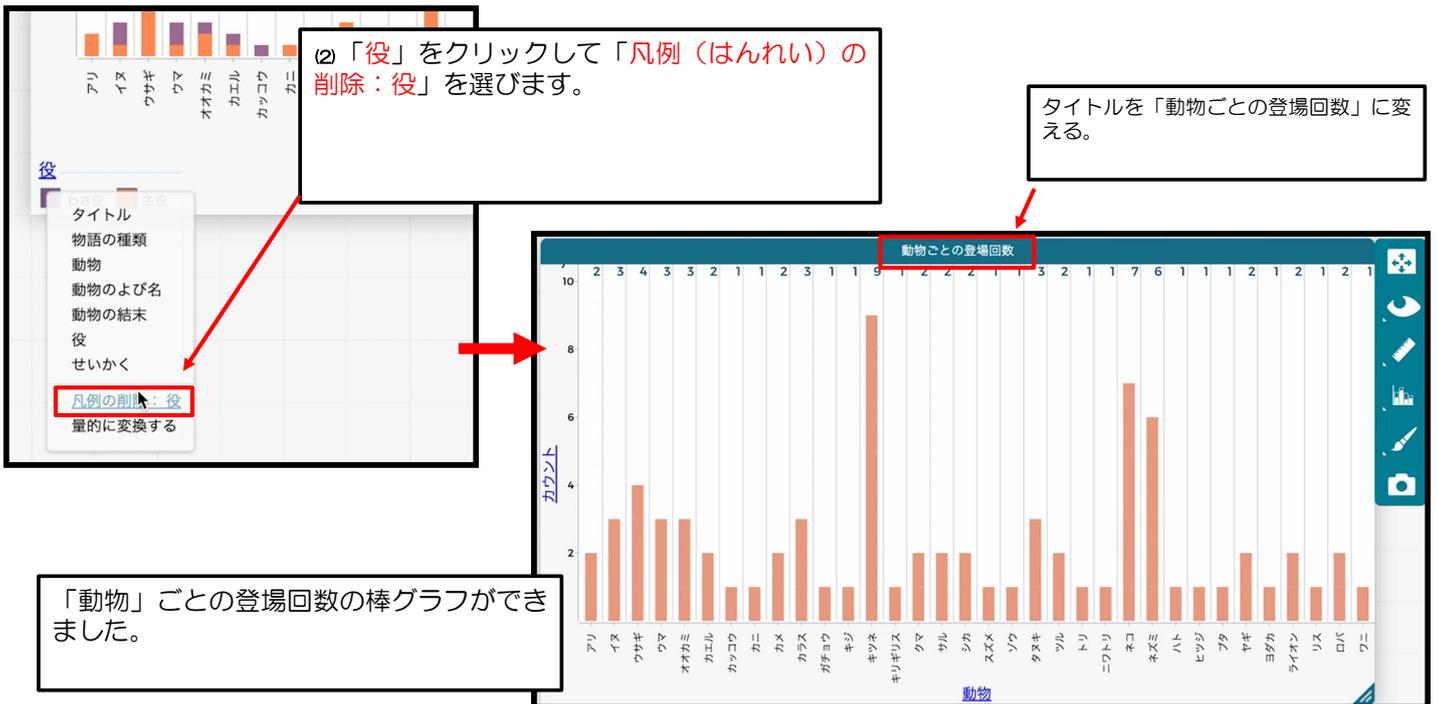
※今回わたしたちが調べたデータの中での結果です。

どんな動物が物語に多く登場する？

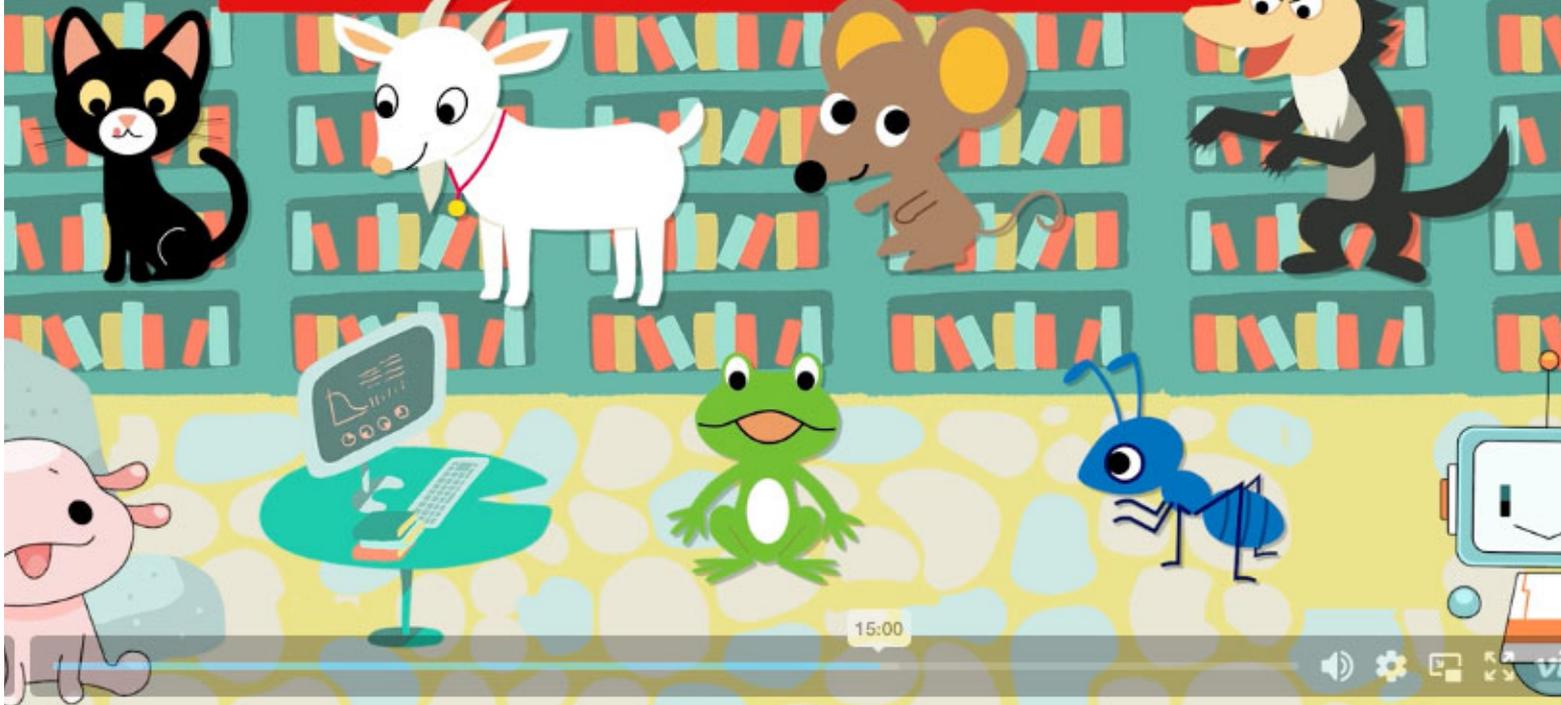


物語にはどんな動物が出てくるんだろう？

CODAPの操作（そうさ）手順



物語の種類によって登場する動物に違いはある？



物語の種類ごとに登場する動物（ドットプロットと集計表）



積み上げ棒グラフと棒グラフのまとめ

- ・積み上げ棒グラフは、棒グラフの棒の中に複数の項目を積み上げて表すグラフで、「全体の数」と「項目の数」を同時に表すことができる。
- ・積み上げ棒グラフは並べることで、全体どうしも、項目どうしも比べることができる。
- ・項目が多い時は、集計表やドットプロットの方がわかりやすい。
- ・積み上げ棒グラフや棒グラフは数量の大きさの順番に並べ替えるとわかりやすくなる。

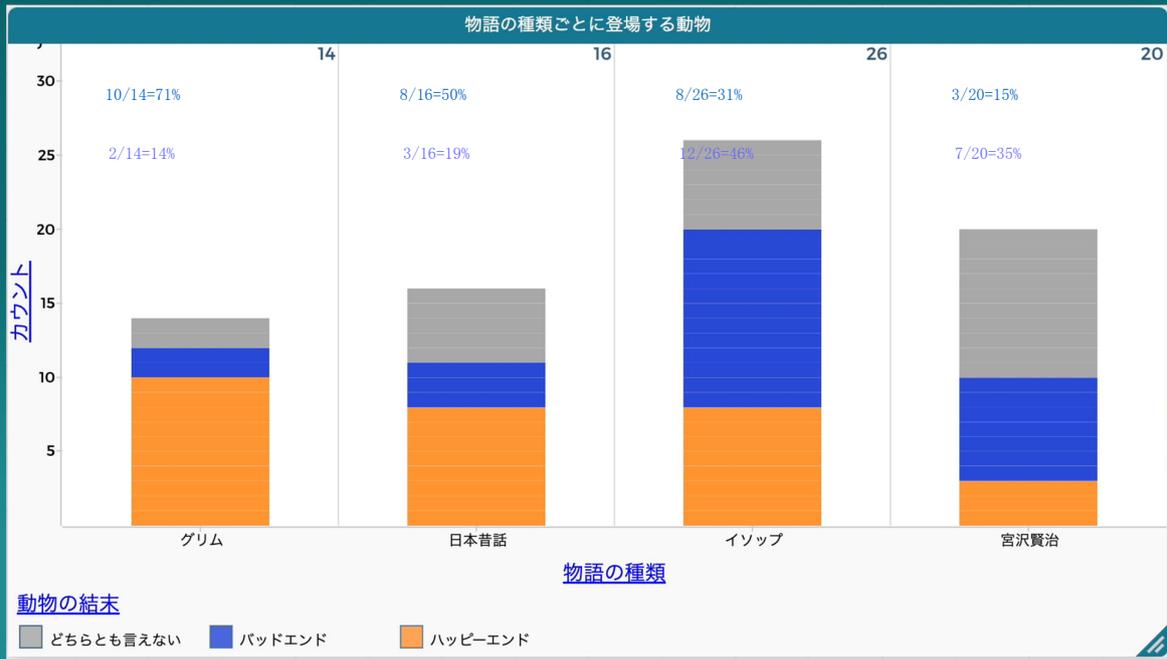


物語に登場する動物の結末

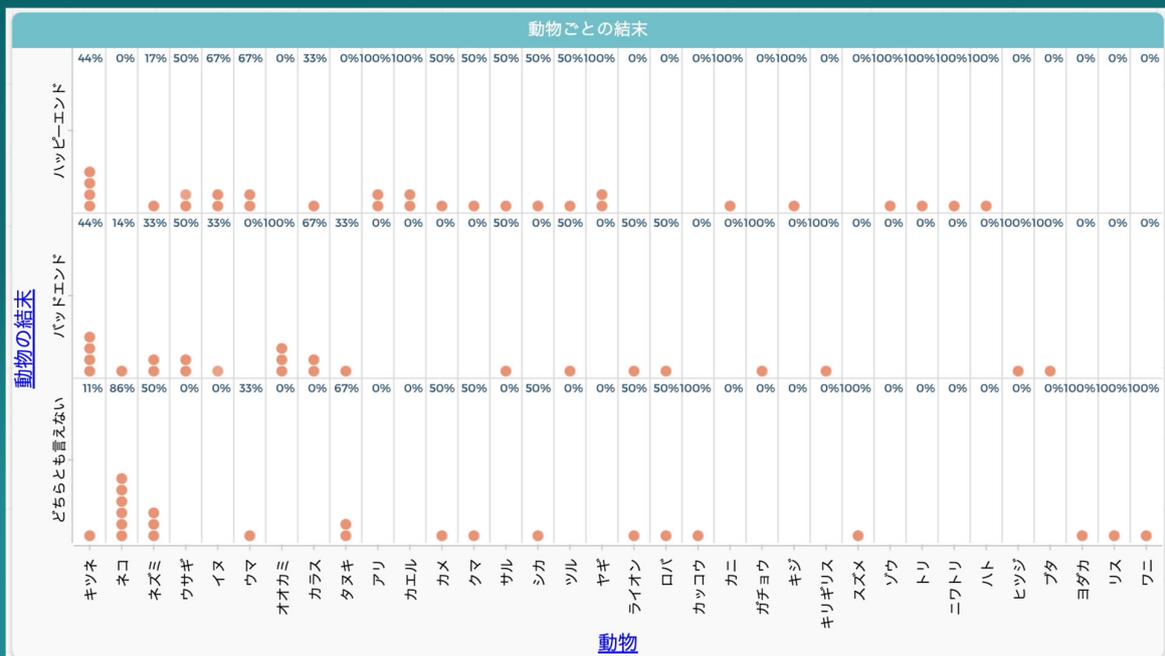
ハッピーエンド? バッドエンド?



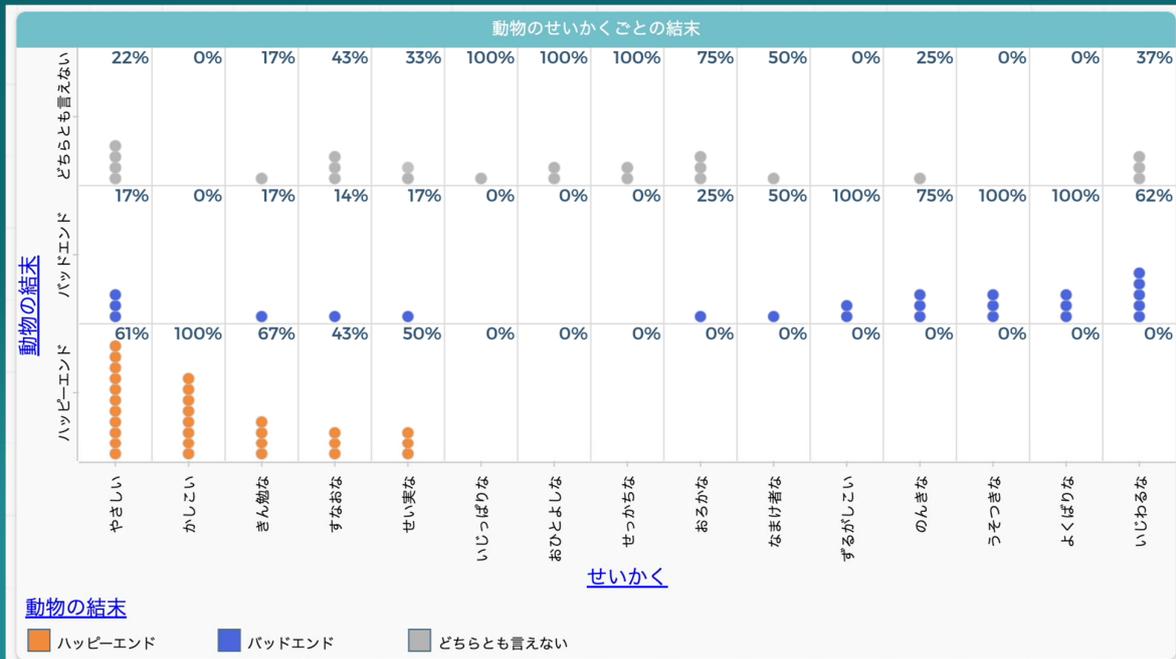
物語の種類ごとの動物の結末（積み上げ棒グラフ）



動物ごとの結末（集計表のドットプロット）



動物のせいかくごとの結末 (ドットプロットと割合)



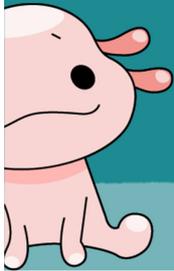
今回の調査 (ちょうさ) のまとめ ①

- 動物が主役のタイトルが、一番多いのはイソップ童話で、次が宮沢賢治作品。
- グリム童話は、動物が主役のタイトルは少ない。
- 物語に登場する回数が多い動物は、1番目がキツネ、次がネコ、3番目はネズミ。
- 宮沢賢治作品にはネコが多く出てくる。
- イソップ童話は、出てくる動物の種類が多い。
- グリム童話と日本昔話に出てくる動物は、ハッピーエンドの割合が大きい。
- イソップ童話に出てくる動物は、バッドエンドの方が、ハッピーエンドより割合が大きい。

今回の調査（ちょうさ）のまとめ ②

- ヤギ、カエル、アリの結末は、ハッピーエンドだけだった。
- キツネは、ハッピーエンドとバッドエンドの割合が同じ。
- やさしいせいかくの動物は、ハッピーエンドの割合が大きいが、バッドエンドになる割合も小さくはない。
- かしこい動物は、ハッピーエンドだけだった。
- キツネは、いじわるなせいかくとかしこいせいかくのキツネが、2タイトルずつ登場する。

4種類の物語の全ての作品を読んだのではなく、自分たちの周りにある作品を選んで調査したので、データには、かたよがりがあり、一般化はできません。



オンライン生徒参加型プラットフォーム『センサス@スクール』の紹介と授業での導入事例

竹内光悦・実践女子大学人間社会学部
末永勝征・鹿児島純心女子短期大学、渡辺美智子・立正大学
連絡先 Email : takeuchi-akinobu@jissen.ac.jp

1. はじめに

日本統計学会統計教育委員会では、ISI・IASE に関係する国際統計リテラシープロジェクト (ISLP、International Statistical Literacy Project) を中心として、国際的に展開されている「Census at School (以下、CAS)」の日本版「センサス@スクール」の開発・運用を行っている。CAS は世界 10 ケ国 (2013 年当時) で運営される Web ベースのデータ収集・相互利活用のプラットフォームであり、オンラインで児童・生徒が参加できる形式をとっている (CensusAtSchool NEW ZEALAND、2023)。本システムのベースはオーストラリア統計局とのライセンス協定により、提供されたシステムを日本語にローカライズし、公開してきた。2022 年まで、統計数理研究所のサーバで運用し、後援には総務省、文部科学省を頂いており、これまでも初等中等教育の教員をはじめ、多くの方々に使われてきた (青山、2008)。

しかしながら実装から 10 年以上たち、システムの老朽化、特にサーバ上での運営に関して、セキュリティ等の問題が指摘され、システム改修が急務となった。そこで本報告では、2022 年度に行った本システム改修の詳細について報告し、授業での導入事例についても触れる。

なお本システム改修において、日本統計学会 ISI 東京大会記念基金には多大な支援を頂いた。また統計数理研究所にはキャラクターの利用許可など、多くの協力・支援を頂いた。ここに記して感謝の意を示したい。また本報告では、本改修に関して代表的な活動者を発表連名者として列記したが、日本統計学会統計教育委員会としての活動であることにも注意されたい。日本統計学会および同統計教育委員会の関係者の方々、ISI 東京記念基金運営委員会の方々など、多くの方々の支援で継続運営できたことに改めて感謝したい。

2. 『センサス@スクール』について

2.1 CAS の設立までの流れ

CAS は、各国で国勢調査を始めとする政府統計調査に関する回答率の低下が問題になっていた時期に、学校教育の中で共通質問票による模擬的なセンサスを Web 上で実施し生徒に調査票回答からデータセットの取得、データクリーニング、統計の作成等を学習させる経験を通して統計調査に関する理解促進を図るために始められた教育プロジェクトで、IASE における ISLP によって、2003

年から 2006 年にかけて RSS 統計教育委員会委員長 Nevile 氏を中心に積極的に、各国に参加が求められたものである。参加国は、当時の ISLP サイトから各国のサイトにリンクが相互に貼られ、また、ISI 大会、IASE サテライト会議時に開催される ISLP 委員会で報告が行われていた。

2.2 日本における『センサス@スクール』について

日本における『センサス@スクール』については、三浦由己元統計学会会長が ISLP の前身である 1994 年設置の ISI 特別委員会 World Numeracy Project の発議者の一人であったことから、ISLP、IASE と連携して活動を行っていた経緯から、2008 年の統計教育委員会の会議で日本も参加することを決定し、統計教育委員会にセンサス@スクール研究部会を設置し、システム開発・運用を行うことになった。

当時の統計数理研究所が配布していた『センサス@スクール』のパンフレットには、その特徴として、以下の 7 項目が示されていた。これらは今回の改修では踏襲している。

- 教室でデータをとる統計授業はプライバシーの問題も発生する懸念もあるが、その心配がない。
- 参加国の各学年の教員が同じ調査票データを題材にした授業事例・ワークショップ等を各国サイトでオープンに共有する仕組みがある。
- 自分自身のデータも加えられ、かつ自分たちと同世代の対象から得られたデータを扱っているということで興味を持たせることができる。
- データそのものは生徒にとって身近なものであるため、生徒にとっても分析しやすい。
- 地域ごと、あるいは国ごとのデータの比較をすることができる。
- 質問項目に回答する中で、調査票の作成方法を始め、調査方法などについても学ぶことができる。国勢調査など社会で行われている調査の仕組みについて理解できる。

3. 『センサス@スクール』の改修について

3.1 改修の基本方針について

上述したように、セキュリティ等の問題から、単なるシステムおよびサーバの移動では解決でき

ず、また今後のサーバ管理、時代とともに変化するプログラムの老朽化などの課題があることから、改修の基本方針として、本システムの1)サーバ維持の継続性、2)既知のセキュリティ対策と柔軟性、3)管理・運用の利便性確保、などの点から、これまでのPHPとJavaScriptベースのシステムを作り直すよりも、クラウドサービスの一つであるGoogleのサービスを利用したものを検討した。これにより、調査はGoogleフォームによりデータを収集することは容易であり、かつGoogleが提供している機能を使うことで、セキュリティ対策も行うこととした。ただしこの場合、一部の質問項目はJavaScriptの実装が必要となりレンタルサーバでCASサイトと共に運用した。また簡単なグラフ描画機能がこれまでのシステムでも実装されているが、それらもGoogleが提供する「Google Charts」を使うことで実装することとした。

3.2 『センサス@スクール』改訂版

2023年1月から上記の方針に従い、改訂版の『センサス@スクール』を日本統計学会統計教育委員会の運営のもと、以下のサイトで公開している(日本統計学会統計教育委員会、2023)。

<https://census.k-junshin.ac.jp/>

改訂版のサイトでは小学生・中学生の生徒・児童にも使われやすいように、キャラクター(統計数理研究所のご協力に感謝)を用いて、サイト全体で統一感を意識した(図1)。主なコンテンツにはCASの紹介、本サイトを利用する際に重要な概念となる「PPDACサイクルの紹介」などの紹介文に加え、データ収集ページであるWeb調査票(図2)へのリンクと説明および回答方法のマニュアル、実際に収集したデータを取り出してグラフを作るページ(図3)へのリンクと説明および回答方法のマニュアルなどがある。

4. 授業での導入事例とまとめ

本改訂版の公開が2023年1月からであることもあり、まだ主だった導入事例はできていない。そこで著者の所属している大学で、大学生にアンケートの回答を授業内でを行い、その感想等を聞いた。この結果については、当日会場に紹介するが、おおむねこのような活動に好意的であり、身近なデータでデータ分析を学ぶことの利点を感じているコメントもあった。オンラインであることもあり、データ収集については、自宅等でも実施でき、学校では公開されているデータ分析について、PCやタブレット端末を手元に議論することも想定される。今後、ますますの本システムの利用を期待したい。

参考文献

[1] 青山和裕(2008)国際的な生徒参加型データを

活用した統計授業の提案、日本科学教育学会年會論文集、32、333-334。

[2] CensusAtSchool NEW ZEALAND (2023) CensusAtSchool New Zealand - TaurangaKiTeKura Aotearoa、<https://new.censusatschool.org.nz/>(最終確認日:2023年2月28日)。

[3] 日本統計学会統計教育委員会(2023)国際的な生徒参加型データを活用する学習サイト「CensusAtSchool JAPAN」センサス@スクール、<https://census.k-junshin.ac.jp/>(最終確認日:2023年2月28日)。



図 1. センサス@スクールのトップページ



図 2. センサス@スクールの調査票

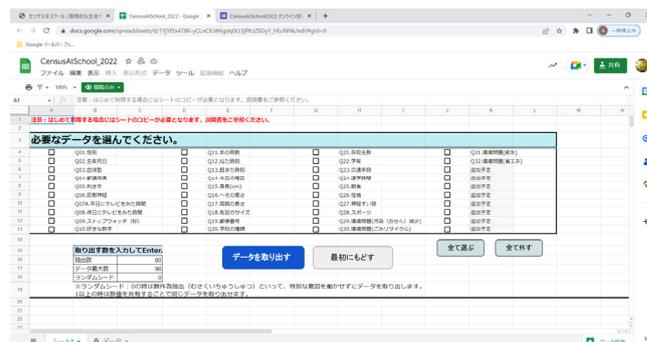


図 3. Google スプレッドシート版のデータ表示

ICT を活用したロバストな工程設計における教育的効果（家庭科授業での事例?）

田中 正敏 (松本大学)

〒390-1295 長野県松本市新村 2095-1

TEL: 0263-48-7243 Fax:0263-48-7290

E-mail: masatoshi.nishiwaki@t.matsu.ac.jp

1 はじめに

近年、コロナ禍によるオンライン授業（ネットワークシステム構築（ZOOM あるいは TEAMS など）のソフトウェアの利用）が余儀なくされつつある。このとき、講義内容や進行についても、従来では考えられないような環境で対処し、工夫をせざるを得ないのが実情である。例えば、統計学教育において、非対面的な教育で、効果的な統計学やデータサイエンスの授業をどう実現するか、また、その評価もどのようにするかが大きな課題であると考えられている。

ところで、統計用語として、日常的に使われる言葉としては「平均」である。例えば、グループで飲みに行き、勘定は割り勘で支払う、また、スーパーマーケットで、果物や野菜などをバスケットに入れる場合、できるだけ粒を揃える（大小・多少などの差が少なく、揃っていること）ようにする、と言う、ここでの「割り勘」や「粒を揃える」という言葉も統計で言う「平均」である。一般的に、「平均」という意味はよく知られているが、「拡がり」や「ばらつき」という言葉、統計用語では「分散・標準偏差」はあまり日常的で利用されていないのが現状である。しかしながら、「拡がり」や「ばらつき」という言葉もよく考案してみると出現している。前述したように、「粒を揃える」という言葉は見方を変えれば「ばらつきを小さくする」という解釈になる。このように、統計に関わっている学校の先生や大学の研究者たちが小・中学生にこの「拡がり」や「ばらつき」のイメージや意味を理解してもらおうことを考えている。

さて、ある工業製品の部品の外形寸法や待ち行列におけるサービス活動時間などには「ばらつき」があり、そのパラメータである「ばらつき」が、本来、あらゆる品質問題の源として考えられてきてい

る。このとき、「ばらつき」の源を同定したり、それらを削除したりすることは、品質改善の目的において最優先課題である。しかしながら、いつも、ばらつきを削除できるとは限らない。例えば、人（サービスを楽しむもの（患者、顧客））や人の流れ（組立ラインの労働者など）を扱うときに、我々の管理を超えるばらつきを受け入れなければならない場合が存在する。さらにその上、「ばらつき」の源が管理できても、それを削除することが返って費用を被る結果となるかもしれない。

したがって、そのようなことの対策の1つにロバストな工程設計がある。これは工程に「ばらつき」がさらされても、欠陥品をカバーできる工程を設計することを意味する。以上のような問題を扱うのが品質管理である。

本研究は、「ばらつき」という統計用語を用いて、どのようにすればシステムの機能が耐性のある状態（ロバストな工程設計）にあるのかという問題を対処することである。具体的には、小・中学生における家庭科の時間での調理実習において、パン作りを考えている。このとき、生徒にはパンに対するし好の程度がその製品属性（味、見た目、匂いなど）を持っている。しかし、ここでは、1つだけの製品属性として、パンの噛み応えに絞って評価を行うことを考えている。パンに焼き方は、温度と時間で決める。一般的に、焼く温度が低ければ、焦げ付きへの変化がゆっくりであるが、焼く温度が高ければ焦げ付きへの変化が早くなる。ここで、我々は、うまく焼く（噛み応えがあり、味もそれなりに確保できている）ためのパラメータには焼き時間と温度のパラメータを設定することである。また、温度と時間を設定するための機材としてフライパンを用い、さらに、調理実習結果からアンケート調査を行い、ばらつきにおけるロバストな工程設計の意味を理解させる。

最後に、オンライン授業による遠隔操作に対しても、この調理実習を録画機材で録画を行い、ファイルに保存しておけば、学生はZOOMあるいはTEAMSなど利用して、必要なときに繰り返しそのファイルを再生することができ、教育的効果にも貢献できるはずである。

以上のことにより、ICTを活用した統計学教育の一例である「ばらつき」の考え方を「ロバストな工程設計」という品質設計の概念などで理解できるものと考えている。

2 ロバストな工程設計 [1]

ある工程におけるパラメーターによるばらつきが、その起源を同定したり、その問題を削除したりすることは、品質改善の目的において最初の優先順位になる。しかしながら、いつも、ばらつきを削除できるとは限らない。さらにその上、ばらつきの源が我々の管理下であっても、それを削除することが返って費用を被る結果となるかもしれない。そのため、ロバスト的な工程を設計しなければならない。

「ロバスト」な工程とは工程に「ばらつき」がさらされても、欠陥品をカバーできる工程を設計することである。

例えば [1]、あるレストランでコーヒーショップやカフェテリアがある。ここでは、クッキーを準備している。買い求める顧客は多くの重要な製品属性（ファクター）を持っている。例えば、味、見た目、価格などである。ここでは一番重要なものはクッキーの噛み応えにしている。このとき、このレストランでは、焼き時間が短いとあまりにも噛み応えはないけれども、味の方はよい。しかし、焼き過ぎる硬くてばりばりで、噛み応えはあるけれども、味の方は返って悪くなる。うまく焼く（噛み応えがあり、味も確保できている）ためのパラメーターには2つの焼き方がある。それは、焼き時間と温度で決める。

また、我々は、クッキーの噛み応えを以下のように定義する。

$$\text{クッキーの噛み応え} = F_1(\text{焼き時間 (分)})$$

$$+ F_2(\text{設定温度 (}^\circ\text{C)})$$

いま、温度と時間に対する噛み応えの対応関係は以下の図1に示される。

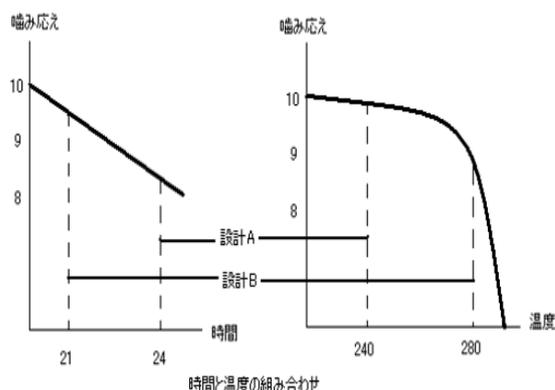


図1. 温度と時間に対する噛み応えの対応関係

このとき、種類の焼き方（設計A：焼き時間が24分で温度設定が240°C，設計B：焼き時間が21分で温度設定280°C）を示す。その結果を図2に示す。図2の説明として、（時間，温度）の組み合わせ

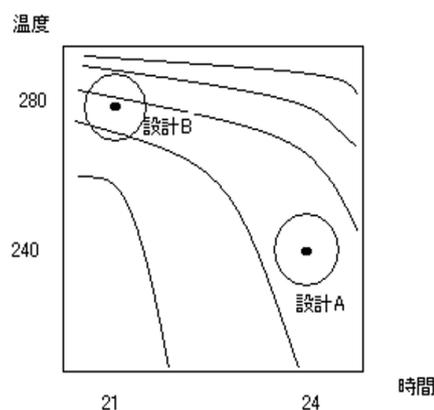


図2. 3次元の平面図

に対する点、ばらつきとして選択される（時間，温度）の組み合わせと一致する円を示す。このとき、設計Aがロバスト的であるのは、円がその領域内に対

べて入っているからである。

よって、図 2 より、設計 A の方が多少のばらつきがあっても均一にとれたクッキーが作れることになる。

3 実験・考察

図 3 に示されているように、フライパンを使って、パンを焼く例を示す。



図 3. フライパン

次に、パンの材料を図 4 に示す。これらの材料を混ぜてパン生地を作る。



図 4. パンの材料

今回も、2 種類の焼き方（設計 A：焼き時間が長く、温度設定が低い場合、設計 B：焼き時間が短く、温度設定が低い場合）を示す。



図 5. 結果（もちもち）

図 5, 6, および、実験に手伝っていただいた数名の被験者のアンケート調査結果から、次の結果を得



図 6. 結果（パリパリ）

た。もちもち感の方はあまり焦げなかった。これは、分散が大きい場合でも対処できることを意味し、一方、パリパリ感の方は焦げた。つまり、これは、分散が小さいので層が変わることを意味する。

以上のことから、「ばらつき」や「拡がり」のイメージや意味が深まることが考えられる。

4 まとめ

ICT を活用した統計学教育の一例である「ばらつき」の考え方を「ロバストな工程設計」という品質設計の概念で示した。今後は、データの収集に力を入れていき、分析を行いたい。

[謝辞]

本研究は統計数理研究所（2022-ISMCRP-4201）から一部支援を受けている。

参考文献

- [1] Cachon Terwiesh, “Matching supply with demand”, McGraw-Hill Higher Education, pp.192-193, 2009.

統計リテラシー教育の世界的潮流と評価

発表者：深澤弘美（東京医療保健大学）h-fukasawa@thcu.ac.jp
共同研究者：及川久遠（大和大学）oikawa.hisao@yamato-univ.jp

1. はじめに

日本では、「個別最適な学びと協働的な学びの一体的な充実」という新しいキーワードのもと教育の見直しを実施され、GIGA スクール構想などの新たな教育が急速に進展している。各学校の各教室では、多種多様な学びを具現化し、個々の児童生徒のニーズや特性に応えることが求められる。統計教育においても、算数・数学科のみならず情報、理科、社会などの科目を横断した多様な学びの内容および学び方の選択肢も必要になる。本稿ではニュージーランドの教科書を参考に、統計リテラシー教育の内容とその評価の考え方について考える。

2. ニュージーランドの初等教育の教科書

「Maths Aotearoa（ニュージーランドの数学）」（図1）は、2019年より順次新シリーズが販売されたニュージーランドの初等教育用の教科書である。2022年秋に全ての学年の販売がスタートした。これらの教科書は、これまで多くの小学校及び中学校で使用されてきた算数・数学（ニュージーランドの科目名「数学と統計」）の教科書 Pearson Mathematics シリーズの最新版である。表1に示すように、学年ごとに8冊で構成されているが、習熟度別教育を実施しているため、到達度のレベルによって学年とは異なる教科書を使うこともある。

ニュージーランドでは教科書の使用が義務付けられていないため、我が国のような教科書検定等の制度も存在しない。教科書を使っていない学校もあるが、Pearson Mathematics シリーズはシェアの高い教科書である。ニュージーランドの国定カリキュラムの内容をすべて含んでいるため、他の教材を必要とせず、単独での使用が可能であると同時に、既存の教材を補足や補強のために合わせて使用することもできる。それぞれ、教師用、デジタル版、ワークブックが出版されている。著者である Charlotte Wilkinson 氏は、数学

教育を専門とする教育コンサルタントで、ニュージーランドと英国で1年生から8年生までの初等教育の経験を持つ教師でもあり、教師教育にもかかわってきた。なお、2022年夏時点で販売されている同シリーズの初等教育向け教科書および中等教育向け教科書については、西村他（2022）「算数・数学の教科書の世界的潮流に関する調査研究」で詳細を紹介している。

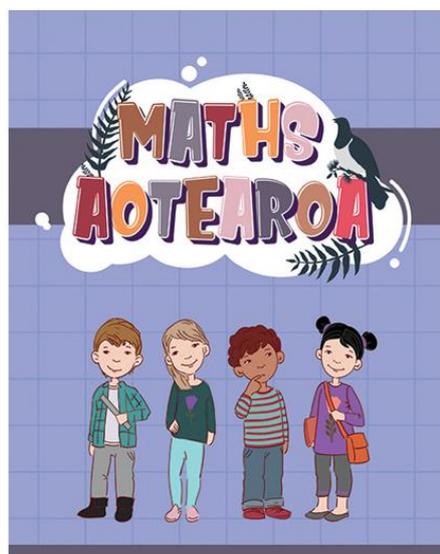


図1 初等教育向け教科書

表1 初等教育向け教科書「Maths Aotearoa」

番号	レベル	NZの学年	日本の学年
1A	レベル1	Y1-2	幼稚園
1B			小学校1年生
2A	レベル2	Y3	小学校2年生
2B		Y4	小学校3年生
3A	レベル3	Y5	小学校4年生
3B		Y6	小学校5年生
4A	レベル4	Y7	小学校6年生
4B		Y8	中学1年生

ニュージーランドの国定カリキュラムについては、教育省が開設するポータルサイト TKI

(Te Kete Ipurangi – the online knowledge basket – is New Zealand’s bilingual education portal, <https://www.tki.org.nz/>)に詳細情報が公開されている。

2. 1 小学校低学年の教科書

日本の小学校低学年（小学校2年生まで）に対応する教科書であるシリーズ番号2Aは、ニュージーランド国定カリキュラムレベル2に準拠している。統計の内容としては、29章「データの収集と表示」、第30章「データの分類」が13ページにわたって掲載されている。確率は6ページである。「データの分類」では、図2に示すベン図、クロス表、樹形図を使って図形を種類ごとに分類する方法を学習する。

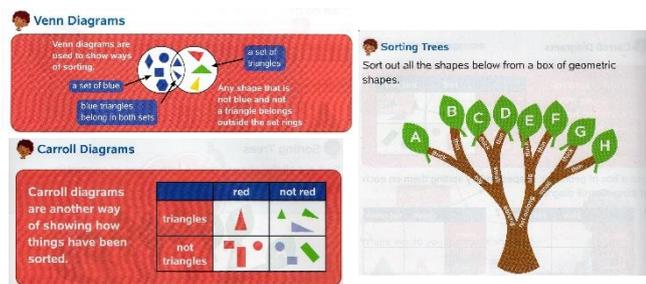


図2 データの分類（小学校2年生）

2. 2 小学校中学年の教科書

(1) 小学校3年生 (2B)

第29章「データの並べ替え」の学習目標は以下の3つは、章の冒頭で紹介されている(図3参照)。

- 表にデータを整理して並べ替える
- ITツールを用いて表を作成する
- データを意思決定や問題解決に活用する

演習としては、図4に示すように紙またはコンピュータを使って右のデータを整理する。

コンピュータを使った演習については、図5から図8に示す通り、操作方法を含む表の作成、データの並べ替え、およびデータの読み取りに関する内容が、丁寧に紹介されている。学習文書作成ソフト(Microsoft Word)で文書中に表を挿入する方法、表計算ソフト(Microsoft Excel)で表を作成し並び替える方法(罫線を引くなどの操作方法を含む)が手順を追って画像付きで紹介されている。

**CHAPTER
29**

Sorting Data

Working through this chapter you will:

- learn to use a table to sort information
- learn to use technological tools to design tables
- learn to use your data to make decisions and answer questions.

Designing Tables

A table is a way of sorting information or data to be collected. It is made up of columns and rows.

Use a Word document on a computer. Go to Table and click Insert Table.

Think about the information you need to collect to decide how many columns and rows your table will need. (On a computer you can add or delete columns and rows if you change your mind!)

Each column in a table must have a heading.

Design a table to collect the following data (either on paper, using a ruler, or on a computer):

everybody's name, age, favourite colour and if they are left-handed or right-handed.

図3 第29章データの並べ替え（小学校3年生）

Compare your table with someone else. Are they exactly the same or could they be slightly different?

Look at the following table:

Name	Girl or Boy	Age	Birth Month	Eye Colour
Emily	Girl	16	December	Blue
Kristina	Girl	12	November	Brown
Jonathon	Boy	14	November	Blue
William	Boy	11	April	Blue
Rachel	Girl	15	October	Blue
Emma	Girl	13	May	Brown
Arwen	Girl	11	March	Brown
Bethan	Girl	6	March	Brown
Megan	Girl	11	February	Blue
Julie	Girl	9	February	Blue
Sarah	Girl	8	November	Blue
Nathan	Boy	8	November	Blue
Joshua	Boy	10	January	Brown
Ben	Boy	7	December	Brown

Read each sentence and use the table to say whether the statement is true or false.

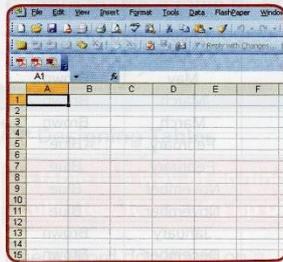
- 1 William has blue eyes.
- 2 If you are born in March you have brown eyes.
- 3 There are more boys born in November than girls.
- 4 The oldest person is a girl with blue eyes.
- 5 The youngest person is a boy with brown eyes.
- 6 William was born in April.
- 7 The youngest boy was born in December.
- 8 All eight-year-olds were born in November.

Other ways of sorting help you to see information even more clearly.

図4 並べ替えの演習（小学校3年生）

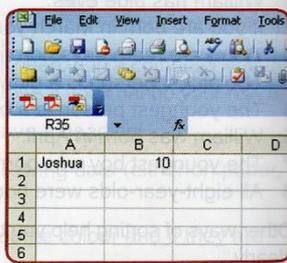
Designing Tables on a Spreadsheet

Open a spreadsheet on a computer. You will see the spreadsheet is made up of columns and rows like a table. Each space in the table is called a cell.



You can type words or numbers in each cell.

Type your name in one cell and your age in another cell. What do you notice?



192

図5 表計算ソフトの使い方① (小学校3年生)

Type the information of the first 7 rows of the table on page 191 onto a spreadsheet.



Place the cursor between the letters on the top row, then click and drag to widen the columns if necessary.



193

図6 表計算ソフトの使い方② (小学校3年生)

When all your data is in the table, use the cursor to left click and drag across the table from top left to bottom right to highlight the whole table.

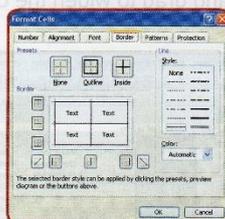
Click on the following sequence to draw a grid for your table.

Click Format



Click Cells

Click Borders



Click Outline and Inside

Click OK.

Your table should now have the grid lines between the rows and columns drawn in.

Using a table on a spreadsheet allows you to sort your data in the table.

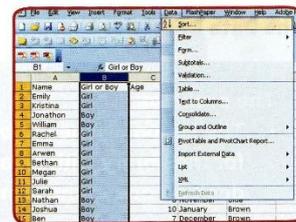
194

図7 表計算ソフトの使い方IV③ (小学校3年生)

Using the icons you can sort the data in a column.

Select the 'Girl or Boy' column by clicking on the letter at the top of the column.

Select Sort and either Ascending or Descending.



You will be asked whether you want to expand the selection.



What happens to the data if you choose yes?
What happens to the data if you choose to use just the selected data?

Which do you think you should use, and why?

What happens to the word data?
What happens to the numerical data?

195

図8 表計算ソフトの使い方④ (小学校3年生)

(2) 小学校 4 年生(3A)

小学校 4 年生の統計の学習は、PPDAC のサイクルに沿って、統計的調査を実行し、集めたデータをドットプロット、幹葉図を用いて視覚化する方法を学ぶ。対象とする多変量のデータを種類ごとに並べ替える方法として、図 9 のカードを使う。性別、誕生月、身長、好きな色の 4 項目を記載して、並べ替えを行い、図 9 に示す間に答える演習が紹介されている。



図 9 データカード (小学校 4 年生)

並べ替えの練習 (小学校 4 年生)

- ① 6 月に生まれた男子生徒は何人ですか
- ② 最も身長の高い生徒は男子生徒ですか女子生徒ですか
- ③ 黄色が好きな女子生徒は何人ですか
- ④ 身長が 1 番小さい女子生徒が生まれた月は何月ですか
- ⑤ 1 番身長の高い男子の好きな色は何でしょう
- ⑥ 赤が好きな 6 月生まれの人は何人ですか
- ⑦ 5 月に生まれた男子生徒のうちブルーが好きではない人は何人ですか

図 10 並べ替えの演習 (小学校 4 年生)

2. 3 小学校高学年の教科書

(1) 小学校 5 年生(3B)

第 29 章「データの収集と表示と読み取り」では、収集したデータから様々なグラフ表現などを学習する。グラフから適切な読み取りを行い、データをもとに結論をまとめる。探索課題では、小説等の文章で以下の 2 点を調査する。

- ・アルファベット「e」は最も使用される母音である
- ・アルファベット「s」は最も使用される子音である

第 30 章「コンピュータを使った統計」は、スプレッドシートに図 11 のデータを入力して並べ替え、データの整理・集計の方法を学ぶ。図 12 に示す間に答えるために、表計算ソフトの操作方法も詳細に紹介されている。

	A	B	C	D	E
1	Name	Jan	April	June	Sept
2	Johan	\$30.00	\$45.60	\$34.50	\$23.00
3	Jonathon	\$45.00	\$32.00	\$27.00	\$23.00
4	William	\$34.60	\$24.60	\$32.70	\$18.00
5	Nathan	\$46.20	\$12.50	\$32.20	\$21.80
6	Kris	\$32.80	\$24.70	\$14.80	\$21.60
7	Mary	\$26.40	\$32.10	\$21.40	\$24.60
8	Roseanna	\$12.80	\$14.60	\$32.90	\$24.00
9	Emily	\$15.70	\$31.70	\$21.90	\$19.50

図 11 データの集計 (小学校 5 年生)

コンピュータを使った演習 (小学校 5 年生)

- ① 一人当たりの年間支出はいくらですか
- ② 1 番お金を使うのは誰か
- ③ 最もお金を使わないのは誰か
- ④ 各四半期の合計支出額はいくらか
- ⑤ 全部でいくらのお金が使われたか
- ⑥ 最もお金を使う人と使わない人の違い
- ⑦ 1 月と 9 月の差はいくらか
- ⑧ クリスとウィリアムズの差はいくらか

図 12 コンピュータを使った演習 (小学校 5 年生)

(2) 小学校 6 年生 (4A)

この教科書では、統計に関連する内容として「時系列グラフ」が独立した章として構成されている。この章では、時系列グラフを正しく読み取りグラフの内容に応じてデータが急激に変化する原因を検討する。読み取りのポイントとして、外れ値、急激な変化、長期的傾向(トレンド)、季節的なパターンや周期について学習する。

3. 統計リテラシーの評価

統計リテラシーの達成度評価については、ペーパー試験や口頭試問による内部で実施する方法以外に、統計検定など外部評価を取り入れるなど、各学校で工夫がなされている。しかしながら、ニュージーランドの教科書等を参考にすると、ICT ツールの操作スキルやデータの基礎的な操作(たとえば、並べ替えなど)についても、一定の質を保証できる教育を実施し適切に評価することが必要と考える。そしてそれらの評価には、全員が同じ問題を解き、その正答率で評価するだけではなく、達成度のレベルにあった評価を実施することも必要であろう。

具体的な評価の方法については、項目反応理論を取り入れるなど現在検討を進め、問題データベースの構築および評価ルーブリックの作成中である。ルーブリックについては次節で 2017 年に開発した PPDAC に沿って実施した統計的問題解決のプロジェクトを評価するルーブリックと統計リテラシーの知識及び技能を評価する方法を紹介する。

3. 1 統計的問題解決力の評価

深澤・和泉・櫻井(2017)では、統計的問題解決評価ルーブリック(SPART: Statistics Project-based Assessment Rubric Table)を開発し、PPDAC サイクルに沿って探求的学びを評価する方法を提案した。図 13 に示す通り、データの操作に関してはデータの「収集」だけではなく「クリーニング」という評価項目を設けて表 2 のルーブリックを基に評価を行っているが具体的な知識や技術のレベルを問うものではない。また、統計グラフについても、表 3 に示す通り全体での評価である。しかしながらこのルーブリックを用いて評価を実施することにより、思考力を含めて統計的問題解決力を評価することが可能となる。

3. 2 統計リテラシーの評価(知識・技能)

統計リテラシーの達成度をペーパー試験等で実施する場合には、前節で示した統計的問題解決力の評価とは異なり、思考力と問うことには工夫が必要である。統計検定では、設問に関する記述を選択する問題や、現実事象の場面を理解しながら具体的に問題を解決する過程を等も問題などが出題されている。また、項目化反応理論などを用いて受験者のレベルにあった問題を出题する際には、さらなる工夫が必要となる。例えば表 4 のように、「箱ひげ図の理解と利用」について評価する際には、評価の段階別に問題を用意し、受験者のレベルに合わせた問題を準備することが考えられる。

4. まとめ

ニュージーランドの初等教育の最新の教科書を調査することにより、初等教育の初期の段階からコンピュータやカードを用いてデータの操作を行う授業が充実し、統計リテラシー教育においてこれらが重要であり、日本の教育では不十分な点があることがわかった。また、ニュージーランドの教科書には、生徒がコンピュータを用いる場合と用いない場合など多様な学習方法を支援する配慮がなされている。多様な視点や方法から指導を進めることの必要性や意義に関連した示唆的内容である。

わが国でも統計教育の重要性が広く認識され、学校教育においても充実が図られているが、諸外国と比べると不十分な点もある。更なる統計教育の充実を図るには、内容を逐次更新することと同時に、正しく教育の評価をする仕組みを構築し、生徒がステップアップして学習を積み重ねられる環境を整備すること、および、生徒の達成度を教育の改善にフィードバックすることが重要と考える。継続して、統計リテラシー領域の学習内容については、箱ひげ図以外の評価ルーブリックを作成し、レベルごとの問題データベースの構築をすすめたい。

参考文献

西村他(2022), 算数・数学の教科書の世界的潮流に関する調査研究 (<https://textbook-rc.or.jp/wp-content/uploads/2022/10/80160bf7b38e4d1fd6ab6a113f8bd3e.pdf>).

深澤・和泉・櫻井 (2017), 統計教育における評価指標の作成と試行—高校, 大学の教養レベル—, 第 13 回統計教育の方法論ワークショップ—学習指導要領の次期改訂に向けた統計教育の新展開—, 東京.

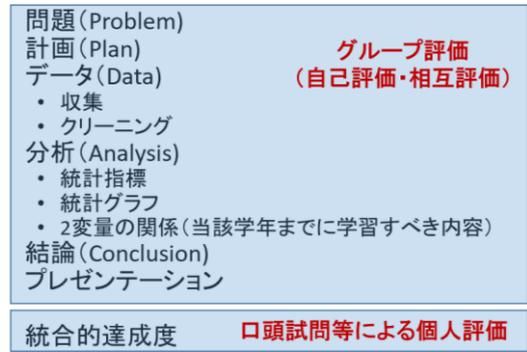


図 13 SPART 評価項目

謝辞

本研究は, 統計数理研究所・共同研究利用・重点型研究 (2022-ISMCRP-4203: 統計基礎リテラシー評価におけるコンピュータ適応型テストに関する研究の助成を受けた。統計数理研究所の船渡川氏より貴重なコメントを頂いた。ここに記して感謝の意を表したい。

表 2 SPART 評価ルーブリック

		1	2	3	4	5	
P(問題)		問題を理解できなかった	提示された問題を扱うことはできた	提示された問題を自らの問題として扱うことができた	自らテーマを設定し, そのための課題を考え, 具体的に問題を探ることができた	自らテーマを設定し, そのための課題を考え, 具体的に統計の問題として扱うことができた 仮説を立てることができた	グループ評価
	P(計画)	必要な計画を理解できなかった	プロジェクトの計画を立てることができなかった	計画の概略を作成することができた データの仕様に関して及んでいた データの収集, クリーニングに関する計画を立てた	適切で概略のある計画を立てた	適切で概略のある計画を立てた 分析の見直しを立てることができた	
D(データ)	収集	必要なデータを理解できなかった	問題に合ったデータを収集できなかった	データが収集できない, 不適切などの理由により問題の解決につながらない不十分な部分があった	問題解決につながるデータを集めることができた	問題解決につながるデータを集めることができた データの不備や欠損, 次への課題についても発見的に考えることができた	個人評価
	クリーニング	問題に合ったデータのクリーニング実行できなかった	データのクリーニングが不適切などの理由により問題の解決につながらない不十分な部分があった	データのクリーニングが不適切などの理由により問題の解決につながらない不十分な部分があった	問題解決につながる適切なクリーニングを実行した	問題解決につながる適切なクリーニングを実行した データの不備や欠損, 次への課題についても発見的に考えることができた	
A(分析)	統計指標	統計の内容を活用できなかった 中心の位置 (平均値, 中央値, 最頻値), ばらつき (範囲, 四分位数, 標準偏差, 分散) などの記述的要素もできなかった	平均, 中央値, 範囲, 四分位数, 標準偏差, 分散など公式通り計算するまでであったが, 要領的な理解にとどまり公式を盲目的に使っただけであった	平均, 中央値, 範囲, 四分位数, 標準偏差, 分散など, 中心, 範囲を示す指標を計算し, データに合わせて使い分けることができた	必要となる統計的概念と手法を考え, 広範囲にわたる統計の内容を適切に使った 図形式, 予測に触れるなど若干のカリキュラムを超えた内容も使った 統計的な説明は簡潔であった	広範囲な統計の内容が問題解決に活用され, それ以上のものも活用された 一部には適用されない手法にも試みた オリジナルのアイデアが見られた	
	統計グラフ	雑で正しくない	棒グラフ, ヒストグラムなど分布を示すグラフ, 構成比率を示す円グラフなどを用いてデータを視覚化することが試みられた グラフの作りかた, かつらに不適切, ヒストグラムと棒グラフの違いなどが区別できていなかった	適切に分布を比較するためのグラフ (棒グラフ, ヒストグラム) を作成できた 箱ひげ図も活用できた	分布を示すグラフ (ヒストグラム, パレート図) と指標を適切に活用し, 正しい解釈を示した	グラフと指標を正しく活用し, 条件の違いやグループでの比較, 対象の分類にも試みた	
	2変量の関係	2変量の関係について分析できなかった	2変量の関係についても分析したが, 不適切な箇所もあり, うまく適用されたとはいえない	2変量の関係について分析を試みたが背後にある問題と向き合っていない箇所があった	2変量の関係について検討し適切に分析した	2変量の関係について適切に分析し, 予測, 問題解決に活用した 図形式にも言及した	
C(まとめ)	結論に至らなかった	不十分な点があるが結論をまとめることができた	結論を適切にまとめることができた	問題の意味を広く認識し, 結論をもとにさらにそれを広げようとした 結果を明確に説明できた	一般化することにより問題を解決した 結果は要領し, 問題やデータの背景および自らの結論の意味, 影響についても説明した		
プレゼンテーション	計画性がなく論議が雑でできない つながりがない短い論議や発表であった 結果がほとんどまとめられていないもしくはその試みすらない	十分に整理できなかった 論議的な要素が欠けていた 比較的重要ではない事柄に焦点が当てられ, 重要な事柄に関する論議や発表がなかった 作業の順に発表するだけで, その順番を再度論理的に構築し直していない	努力を継続し, 適切なプレゼンテーションだった 理にかなったまとめができたが, 若干, 雑でないところや強要すべき箇所を削っている箇所があった	関係なまとめ, 満足できるプレゼンテーションだった 作業は論理的な順番で構成され, 結果の一般化に対して十分に正当性がある	プレゼンテーションは, わかりやすく, 簡潔, グラフが適切に活用された 結論とその一般化は, 理にかなった文章で表された まとめは, 作業の主要部分で得た結果を反映した内容であった 異なる意見についても, まとめに盛り込まれていた		
総合的達成度	字が理論や手法をほとんど理解できていない 字んたことをほとんど使うことができない 多くの助けをかりた	理論や手法を十分に理解できなかった 基本的な方法を真似して使うことができるが, これを超えるための助けが必要 先生やグループの仲間, お手本 (書籍, ネット) 上の情報などの助けをかりた ほとんどの部分は作業をできなかったが, いくつか問題解決のために挑戦できた	当該プロジェクトについて, そこで使われた手法について理解している プロジェクト全体の理解し込んだ上で実行できる 多少の助けをかりながら, 形式的なやり方を実行できた 多少は自身の考えをもとに, 形式的なやり方を実行できた	当該プロジェクトで使われた統計的手法を異なる状況への対応, 応用, 拡張はできない プロジェクトの全体を理解して実行でき, 詳細な議論やまとめを構成できた 問題解決のサイクルをうまく実行できた 多少の助けをかりながら, 自ら実行し発表させた 参考書など資料を有効に活用できた	当該プロジェクトで使われた統計的手法を異なる状況への対応, 応用, 拡張のアイデアを持ち, さらに実行を試みた 当該プロジェクトについて詳細に説明, 文章化できる 大々(特筆される)問題解決のサイクルが(ときには複数回)うまく実行できた 自ら作業できた 参考書など資料を有効に活用できた		

表 3 SPART 統計グラフに関する評価

評価	統計グラフ
1	雑で正しくない
2	棒グラフ, ヒストグラムなど分布を示すグラフ, 構成比率を示す円グラフなどを用いてデータを視覚化することが試みられた グラフのタイトル, ラベルに不備, ヒストグラムと棒グラフの違いなどが区別できていなかった
3	適切に分布を比較するためのグラフ (棒グラフ, ヒストグラム, 箱ひげ図) を作成できた
4	分布を示すグラフ (ヒストグラム, パレート図) と指標を適切に活用し, 正しい解釈を示した
5	グラフと指標を正しく活用し, 条件の違いやグループでの比較, 対象の分類にも試みた

表 4 箱ひげ図の理解と利用に関する評価

評価	箱ひげ図の理解と利用
1	箱とひげの意味を理解し, 5数との対応, 四分位ごとの割合を正しく読み取ることができる
2	時系列に並ぶ箱ひげ図や, 性別ごとの箱ひげ図を正しく比較できる
3	複数の箱ひげ図から全体を考察し正しく分布を読み取ることができる
4	度数分布表, ヒストグラムとの対応を理解し, 正しく分布を読み取れる
5	ひげや箱の間に位置する値についても正しく理解し, 箱ひげ図を正しく読み取り問題解決に活かすことができる

いやにさせない統計・グラフを使うQCサークル教育

榎本竜二 中央大学 非常勤講師

東京都八王子市東中野 742-1・(042)674-2210(代)

email: enomoto-ryu.027@g.chuo-u.ac.jp

1. 全員が学ぶことになった統計学

高等学校では2022年4月から新しい学習指導要領にもとづくカリキュラムが始まった。教科等を越えた全ての学習の基盤として生まれ活用される資質・能力としては、従来から「言語能力」と「問題発見・解決能力」は規定されていたが、今回から新たに「情報活用能力」が入った。

もちろん従来からも情報活用能力という言葉は使われてきたが、「基盤として」入ったことが大きな変更である。文部科学省は情報活用能力を「より具体的に捉えれば」と前置きをし、「学習活動において必要に応じてコンピュータ等の情報手段を適切に用いて情報を得たり、情報を整理・比較したり、得られた情報を分かりやすく発信・伝達したり、必要に応じて保存・共有したりといったことができる力」と説明している。ところが、これに続いて「さらに、このような学習活動を遂行する上で必要となる情報手段の基本的な操作の習得や、プログラミング的思考、情報モラル、情報セキュリティ、統計等に関する資質・能力等も含むものである」としている。小学校学習指導要領解説にまで同じ事が言及されていることから、「統計」の必要性が小学校段階から重視されるということが明らかになったのである。

一方で高等学校でも「数学B」で統計学が事実上必修となり、必ず統計的推定を学ぶことになった。さらに教科「情報」の科目が一本化し、データサイエンスを含む「情報I」が必修科目となった。

2. 知識・技能が問題解決に使えない

学習の仕方も従来の一斉講義から、児童・生徒が自ら学ぶような「主体的・対話的で深い学び」が重視されている。自分たちの考えたこと、集めた情報の正しさをどう証明していくのかという部分で統計学が役に立つ。ところが、統計学を知っていると分析しやすい・理解しやすいという「しやすい」以前の話で、なぜそ

れをするのかわからないという高校生が非常に多い。統計を計算として習い、グラフ作成は表計算ソフトの一つの機能として習うからだ。問題を出されれば何の式を使うかはわかるし、整理して表は作れる(知識)。統計の計算は出来るし、表からグラフは作れる(技能)。ただ、知識・技能はあっても問題解決行動に結び付かないのである。

また、理系の先生方が共感するような気持ちも伝わらない。子どもたちは計算(過程や結果)に楽しさも美しさも感じない。プログラミングのコードのステップが減ってもスゴイとは思わないし、コンパイラのパフォーマンスに感動したりもしない。何のためにしているのかわからないまま、情報の授業なのに理屈もわからずモンテカルロ法の計算をしている。アンケート結果の分析の段階で突如、統計学を前面に打ち出して教えてしまっている。結局は数学ぎらい、コンピュータぎらいを増産している結果になってしまっているのである。

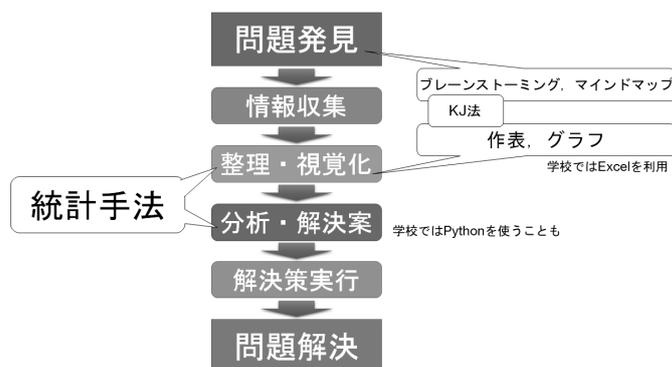


図1 高等学校における問題発見・解決学習

3. 情報は数学との関連は強いが数学ではない

教科「情報」は基本的に2単位なので、学問の前提までやっていると時間が足りない。Excelで合計を計算させるのに、ガウスや公式のことなど知らなくても関数名と引数を知っていれば答えが得られる(理屈はブラックボックスでもよい)。大切なのは、「なぜその範囲の合計を出さなければならないのか」本質を理解させることである。

数学における問題解決は、問題例（事例）を帰納的に考察しながらも、結局は過去（既習）の数学的知識や考え方をを用いて演繹的に解決する（答えは1つ）。しかし、統計学は集められたデータや観測された現象から「たぶんこうなのでは？」という推論を導く帰納的に解決していくものである（答えは1つとは限らず状況で選択しうる）。数学的方法論を情報に持ち込むと、何のためにしているのかがどんどん解らなくなってしまう。

4. 非製造分野で広がったQCサークル

QC（Quality Control）サークルは日本では「小集団改善活動」と呼ばれる。もともとは品質管理の考え方なので工場などの製造部門の品質向上を目指したものであったが、次第に事務所や研究開発部門（はては教育分野）まで広がっていった。

経営層からの売上増大や費用対効果などの押し付けはトップダウンによる改善の考え方である。これに対して、QCサークルは働く現場の同じ空間（事務所）、同じ勤務内容の集団からのボトムアップによる改善活動である。同じ問題意識を持ちやすい集団であるため、互いが共通に感じる問題を解決する活動にモチベーションが持てる。高所・大所からの改善指摘は単なるスローガンになりがちで、解決する問題が大きすぎてどこから手を付けるのか、どう解決に導くのかわからない。QCサークルはそれぞれの部署で感じた直接的な不満や問題を小さな単位で改善活動を提案する。しかも問題の分析にも改善案を持ち引き出すためにも、ツールと呼ばれる図解法やグラフ、統計の考え方が使われる。覚えるべき統計知識のために作られた問題では無く、解決すべき問題のために統計知識を活用するという、本来のあるべき姿としての学習活動が行える。

5. 知識・技能が問題解決に結び付く

どのグラフは何のために作られ、何に向いているのかをわかっていない高校生が多いという現実がある。QCサークルで定番として使われる「7つ道具」のグラフや図には明確な目的がある。すべての道具を使うわけでは無いが問題解決のツールとして目的別に利用し問題点を見つけ、改善提案を構築していく。QCサークルは、活きた統計手法の使い方の実践なのである。ツールを「使う」前提があることで考え方と計画の

組み立てが明確になる。まず、何を「問題」とすべきか、どう分析するか（何で分析するか）、どう対策を立てるか、実施後どう評価するかが、ツールを使うことで集団全員が納得できる。

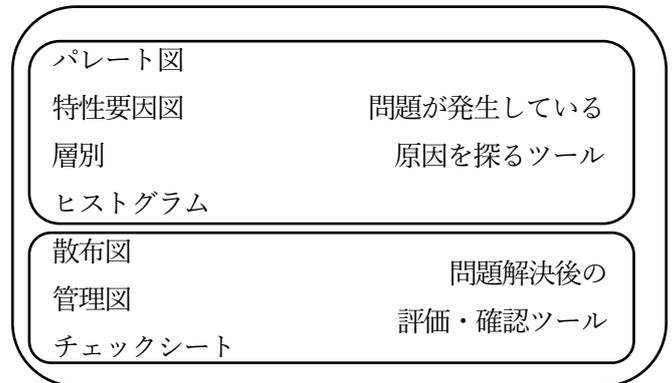


図2 QC 7つ道具

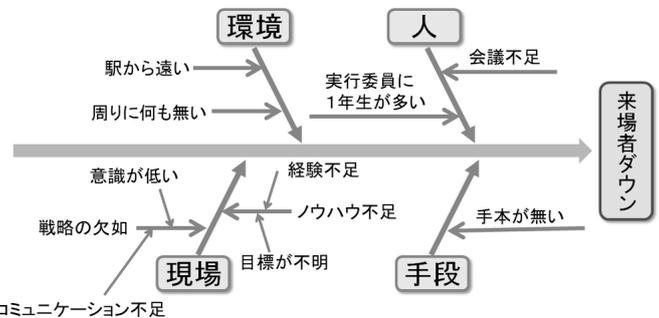


図3 QC 7つ道具の1つ特性要因図

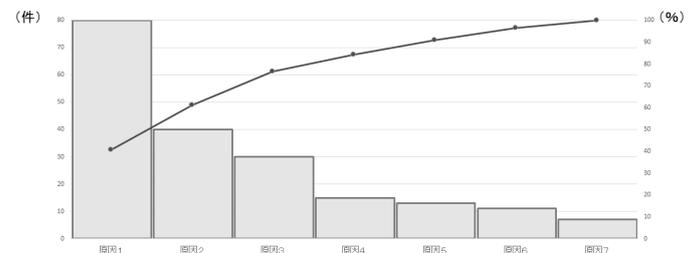


図4 QC 7つ道具の1つパレート図

小さな問題なので、比較的すぐに解決できる。そのため、身近な問題を自らが集団で問題解決をするにはQCサークルが有効であるという認識が生まれ、ツール(統計手法)を使うと問題が明らかになると理解し、ツールを使うと定量的な根拠になると納得できる。サークル内にも外部にも説得力を持って提案することができる。また、この活動を行うためにはICT機器が有効で、情報活用能力が問われるということから高等学校の情報の授業にQCサークルを用いる教育を行うことは有効と思われる。

SSDSE と統計データ分析コンペティションの概要

田中 雅行・独立行政法人統計センター
〒162-8668 東京都新宿区若松町 19-1
TEL：03-5273-1368
E-mail：statcompe@nstac.go.jp

本報告では、SSDSE 及び統計データ分析コンペティションの概要等について解説する。

1. SSDSE の概要

SSDSE（教育用標準データセット：Standardized Statistical Data Set for Education）は、データ分析のための汎用素材として、統計センターが作成し毎年更新の上、公開している。様々な分野（人口・世帯、経済基盤、教育、健康・医療、福祉・社会保障など）の公的統計を、都道府県別又は市区町村別にまとめた、表形式のデータセットであり（ファイル形式は、CSV 及び Excel）、これをダウンロードすることで直ちにデータ分析に利用することができる。政府統計の地域別データを手軽に利用できるような編集しており、項目ごとの解説もあり、データサイエンス演習、統計教育などに活用できる。

2018 年に提供を開始し、現在、以下の 5 種類が無償ダウンロードできる。今年は 2 月 27 日に SSDSE-D のデータ更新を行い、残りの SSDSE も 4 月から 7 月にかけて、最新データに更新、順次提供を行うこととしている。今後も、利用者の要望を参考に内容を充実させていく予定である。

SSDSE-市区町村（SSDSE-A）：

市区町村別、多分野データ（項目数 124）

SSDSE-県別推移（SSDSE-B）：

都道府県別、時系列、多分野データ(107)

SSDSE-家計消費（SSDSE-C）：

都道府県庁所在市別、家計消費データ(226)

SSDSE-社会生活（SSDSE-D）：

都道府県別、自由時間活動・生活時間データ(121)

SSDSE-基本素材（SSDSE-E）：

都道府県別、多分野データ(90)

2. 統計データ分析コンペティションの概要

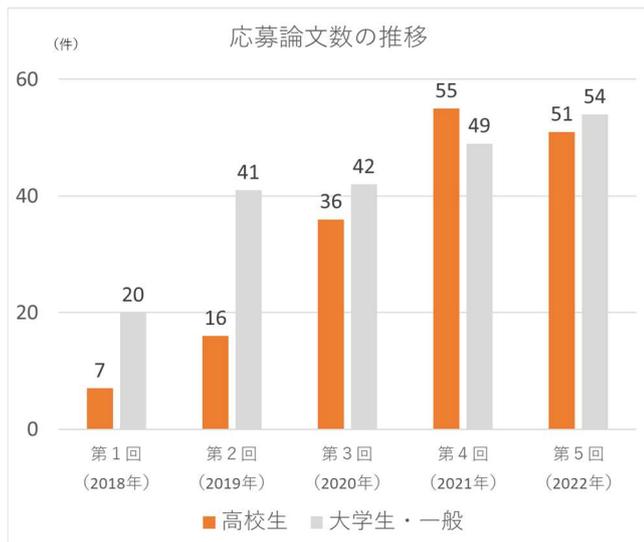
独立行政法人統計センター（以下、統計センター）では、総務省統計局等との共催により、統計リテラシー向上のための取組として統計データ分析コンペティションを 2018 年から毎年開催している。2021 年からは文部科学省の後援も得ている。

このコンペでは、高校生、大学生等を対象に論文を公募し（参加無料）、統計センターが提供する日本の地域別統計をまとめたデータセット SSDSE を用いた統計データ分析により、そのアイデアと解析力を評価するものである。

2022 年のポスター（2023 年も 4 月公開予定）

3. 統計データ分析コンペティションの応募状況

昨年 2022 年の応募論文数は、高校生の部 51 本、大学生・一般の部 54 本であり、この応募論文数は増加傾向にある。



4. 受賞論文の概要等

高校生の部の受賞者及び受賞論文の概要は、以下のとおり。

(1) 総務大臣賞

太佐 美結（フェリス女学院高等学校）

「体力が基礎学力に与える影響について」

個別の体力テストの点数から抽出した体力因子が学力に与えている影響を分析し、家庭の経済力や学習環境なども含めた学力との関連性を明らかにすることで、小中学生の運動意欲の低下防止や、運動や学習を行う動機付けについて示した。

(2) 優秀賞

今泉 開（慶應義塾高等学校）

「ボランティア活動の決定要因」

都道府県別のボランティア活動率を用い、重回帰分析や散布図を用いた分析を行うことで、ボランティア活動の決定要因が都道府県単位で存在し、その決定要因がそれぞれのボランティア活動に大きな影響を与えることを示した。

(3) 統計数理賞

林 蔚欣（茨城県立並木中等教育学校）

「市区町村というミクロ的視点から投票率の実態を探る」

市区町村別の投票率を、得票数合計と人口から計算し、それを様々な変数で回帰することによ

り、離婚件数や財政力指数、完全失業率、民生費割合等がマイナス、教員数等がプラスの影響であることを見出した。

(4) 統計活用奨励賞

森下 達也（愛知県立一宮高等学校）

「都市部と地方の教育格差の要因と課題 ～日本の教育現場において～」

都道府県別の進学者数、県民所得、教育費支出、教育関係事業所数等を用い、都道府県別の教育格差やその要因を分析するとともに、教育のデジタル化として、県別のデジタル教科書整備率や電子黒板整備率を比較し、都市部におけるデジタル整備が進んでいないことを示した。

(5) 審査員奨励賞 16本

【参考URL】

- ・統計データ分析コンペティション 2022
<<https://www.nstac.go.jp/statcompe/>>
- ・SSDSE（教育用標準データセット）
<<https://www.nstac.go.jp/use/literacy/ssdse/>>
- ・総務省統計局報道資料
<<https://www.stat.go.jp/info/guide/public/houdou/pdf/ho221018.pdf>>

体力が基礎学力に与える影響について

水谷順一・フェリス女学院高等学校
 太佐美結・フェリス女学院高等学校
 フェリス女学院高等学校

(住所/神奈川県横浜市中区山手町178 TEL/045-641-0242)

1. 研究目的

小学5年生と中学2年生を対象とした「全国体力・運動能力、運動習慣等調査」(以下、体力テスト)によれば、平成30年までは緩やかな体力増加傾向で推移したが、平成30年以降は急減少している。このまま体力低下が進むと健康やメンタル面を通じて学習面に影響が出てくることが指摘された。一方、小学6年生と中学3年生を対象とした「全国学力・学習状況調査」(以下、学力調査)の結果は、同一指導要領で授業が行われているにもかかわらず、都道府県ごとに大きな差が見られる。

そこで本研究では、個別の体力テストの点数から背後に潜んでいる体力因子を抽出し、各体力因子が学力に与えている影響を分析する。さらに、家庭の経済力や学習環境なども含めて学力との関連性を明らかにすることで、小中学生の運動意欲の低下防止や、運動や学習を行う動機付けについて提案することを目的とする。

2. 研究方法

本研究は図1のフローにより実施した。

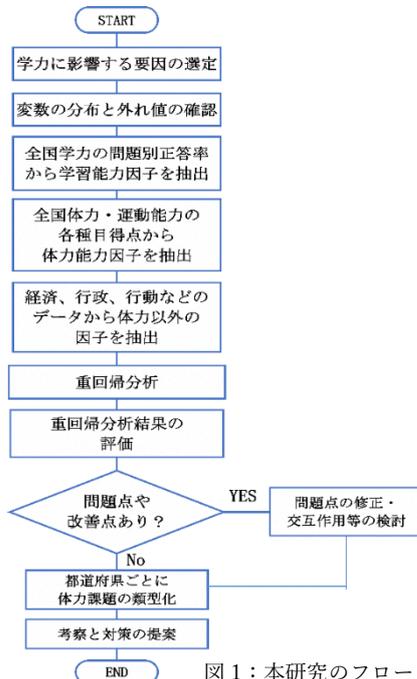


図1: 本研究のフロー

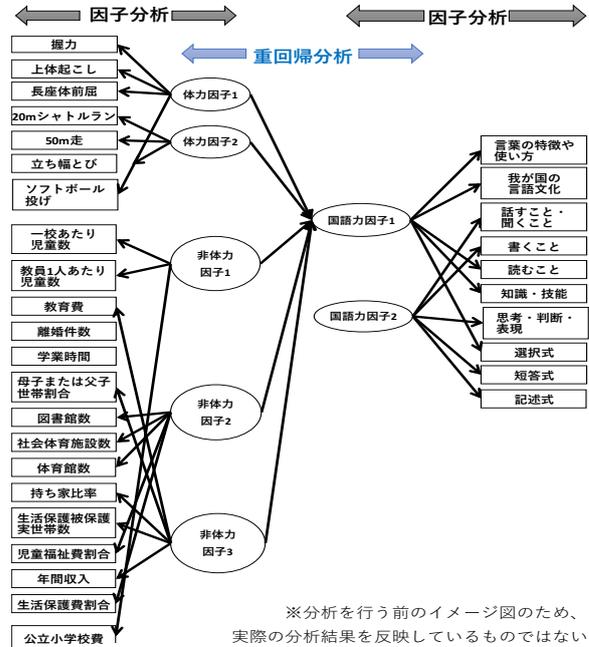


図2: 分析のモデルイメージ

3. 分析結果

(1) 学力、体力および体力以外因子の抽出

因子分析に当たり、因子抽出は主成分法、因子数は固有値1以上とし、斜交回転(プロマックス回転)を使用した。

・学力因子(重回帰分析の目的変数)の抽出

小学生の算数、中学生の国語・数学では、1因子のみであったが、小学生国語では、因子負荷量から『思考的国語力』と『知識的国語力』と解釈可能な2因子が抽出された。

・体力因子(重回帰分析の説明変数)の抽出

体力テストについては小・中学生共に4つの因子が抽出された(表1)。因子負荷量より、「持久力」、「柔軟性と筋力」、「スピード力」、「瞬発力」と解釈した。

・体力以外の因子(重回帰分析の説明変数)の抽出

小学生と中学生共に3因子が抽出され、因子負荷量も類似していた。それぞれ「行政支援」、「家庭環境」、「学習環境」の因子と解釈した(表2)。

表1: 体力テストの因子分析結果(小学生)

小学生	体力
-----	----

変数	因子1	因子2	因子3	因子4	共通性
	持久力	柔軟性と筋力	スピード力	瞬発力	
シャトルラン女子	1.034	-0.225	0.006	-0.023	0.880
反復横とび女子	0.983	0.007	-0.186	0.009	0.936
反復横とび男子	0.915	0.088	-0.107	-0.023	0.925
シャトルラン男子	0.906	-0.165	0.195	-0.145	0.869
立ち幅とび女子	0.702	0.045	0.229	-0.154	0.729
上体起こし女子	0.675	0.448	-0.038	0.349	0.870
ソフト投げ女子	0.586	0.036	-0.106	-0.493	0.815
長座体前屈女子	-0.174	0.952	0.372	0.026	0.939
長座体前屈男子	-0.207	0.932	0.400	-0.086	0.926
握力男子	-0.058	0.823	-0.321	-0.305	0.898
握力女子	0.134	0.771	-0.466	-0.123	0.930
上体起こし男子	0.494	0.548	0.144	0.243	0.836
50m走男子	-0.154	0.077	0.989	-0.089	0.932
50m走女子	0.321	0.078	0.760	0.146	0.830
立ち幅とび男子	0.417	0.066	0.457	-0.414	0.734
ソフト投げ男子	0.053	0.083	0.028	-0.889	0.861
因子寄与	7.476	5.756	2.871	2.709	

表2：体力以外要因の因子分析結果(小学生)

変数	小学生			共通性
	因子1	因子2	因子3	
	行政支援	家庭環境	学習環境	
児童1人当たり公立小学校費	0.982	-0.148	0.249	0.782
小学校教員1人当たり児童数	0.958	-0.148	0.043	0.881
小学校1校当たり児童数	0.886	-0.071	-0.096	0.875
人口1人当たり教育費	0.853	-0.228	-0.048	0.796
児童福祉費割合	0.837	0.304	0.291	0.686
体育館数	0.733	0.123	-0.187	0.732
社会体育施設数	0.723	0.170	-0.279	0.840
図書館数	0.673	0.216	-0.174	0.661
生活保護被保護実世帯数	-0.102	0.889	-0.382	0.857
離婚件数	0.365	0.854	0.208	0.896
年間収入	-0.381	0.854	0.075	0.868
母子または父子世帯割合	-0.104	0.839	0.334	0.881
生活保護費割合	-0.045	0.744	-0.193	0.562
持ち家比率	0.392	0.629	-0.174	0.669
学習時間(在学者、平日)	-0.071	0.133	-0.862	0.691
消費支出に占める教育費	-0.353	0.203	0.528	0.629
因子寄与	6.818	4.268	3.372	

(2) 基礎学力への影響(重回帰分析結果)

(1)で抽出された因子の因子得点を説明変数として、強制投入法による重回帰分析を行った。分析結果から中学生では瞬発力との交互作用について考察した。

表3：学力に対する重回帰分析の結果(小学生)

標準化係数	小学生			
	変数名	思考的国語力	知識的国語力	算数
	持久力	0.321	0.851 **	0.314
	柔軟性と筋力	0.074	-0.195	0.015
	スピード力	0.263	0.136	0.493 **
	瞬発力	-0.169	-0.034	-0.055
	行政支援	-0.263	0.239	0.121
	家庭環境	0.154	-0.213	-0.054
	学習環境	0.293 +	0.247 +	0.358 *
	R ²	0.470	0.594	0.422
	Adjust R ²	0.375	0.522	0.318
	p値	0.000 **	0.000 **	0.002 **

表4：交互作用を入れた場合の重回帰分析結果(中学生)

標準化係数	中学生		
	変数名	国語	算数
	瞬発力	0.181	0.187
	柔軟性と筋力	-0.007	-0.139
	持久力	0.009	-0.028
	スピード力	-0.042	-0.043
	行政支援	-0.159	-0.361 *
	家庭環境	0.379 *	0.472 **
	学習環境	-0.032	0.044
	瞬発力×家庭環境	0.431 *	0.374 *
	R ²	0.468	0.527
	Adjust R ²	0.356	0.427
	p値	0.001 **	0.000 **

4. 分析結果の解釈と考察

(1) 小学生についての考察

・知識的国語力：「持久力」が知識的国語力に影響を与えている。過去の研究より、持久力のある人は、学習や記憶に関わっている脳の「海馬」が大きく、記憶力が優れていることが確認されている。知識習得には記憶の積み重ねが必要であり、相応の記憶力が必須条件となる。習慣的に運動し、短い距離のみならず、長い距離のランニングなどの有酸素運動などを取り入れることが有用と考える。

・算数：「スピード力」が算数に影響する可能性が示された。幼児期から小学生時代の運動神経を形成するゴールデンエイジに鬼ごっこや公園を駆けまわり、スピード力がついた児童ほど、素早い計算処理能力や論理組み立て能力が形成される可能性が示唆される。

・学習環境因子：学習環境因子は5%または10%有意であり、学力との関係性は否定できない。表2より、学習時間について負の因子負荷量が抽出されたため、学習環境(逆符号)と考えた。標準化回帰係数がプラスになっていたため、学習時間が少ないほど、正答率が高いという矛盾した結果が読み取れる。しかし、小学生の体力が学力に影響を与える可能性があることから、体を動かすことや外遊びに力を入れている児童は、学習の時間を運動時間にあてられるとも考えられる。以上の結果から、小学生では、家庭環境や学習環境よりも体を動かす活動が学力を伸ばすことがわかり、学内だけでなく、学外でも放課後のグラウンド開放や定期的な地域の運動交流会などの機会を設けることを提案したい。

(2) 中学生についての考察

中学生では瞬発力と家庭環境の交互作用が、5%有意で、学力に影響を与える可能性が示された。中学生では、瞬発力をはじめとした体力要因は、家庭環境が整った上で学力に良い効果をもたらしていると考えられる。中学生になると、自身のみならず家庭環境にも目を向ける。さらに、学習量が増えるため、家庭環境により、家庭内のみではこなしきれないこともあると考えられる。中学の学習内容を小学校に、高校の学習内容を中学校に少しずつ移行するなど、その時々でこなすことのできる学習量を調節することも提案したい。

5. 結論

本研究を通じて、小中学生の運動意欲の向上や運動を行う動機付けが提案できたと考える。今後は継続的な分析および個人レベルでの分析方法を検討する必要がある。

都市部と地方の教育格差の要因と課題

森下 達也・愛知県立一宮高等学校（2年）

鈴木 淳子・愛知県立一宮高等学校（教員）

〒491-8533 愛知県一宮市北園通6丁目9番地

E-mail : suzuki9059@aichi-c.ed.jp [@は半角]

1. はじめに

昨今、学校教育段階からデータ活用能力を育成することについての期待が大きくなっている。高等学校では、今年度新たに始まった必修科目「情報Ⅰ」において、従来の「情報の科学」「社会と情報」よりもデータ分析分野の割合が多くなっていることもそのあらわれであろう。本校では、1学期にデータ分析の基礎について学んだ後、各自でテーマを設定してレポートにまとめる学習活動を行った。

「統計データ分析コンペティション」の紹介も授業内で行ったところ、夏休み中に追加の分析を行った上で論文形式にまとめて応募した生徒が統計活用奨励賞を受賞した。以下はその概要である。

2. 論文概要

本研究は日本の教育現場における、都市部と地方の教育格差の要因と課題について分析し、是正に向けた取り組みの手掛かりを探ることを目的とした。このテーマを設定した理由は、自身の経験において、学習塾や英会話教室など家庭での学習を補填する機関が都市部に多くみられ、また都市部には中高一貫校が多く、中学・高校時代に上の学年の内容を学習できるなど、より専門的に学ぶ機会を持っていることに格差を感じたからである。また、今回デジタル教育の実態や格差についても調査した。

3. 先行調査・研究の検討

まず、現状に教育格差が存在するという認識が高校生にあるか確認した。2021年に日本財団が行った18歳意識調査(1)によると、「教育格差を感じるか」という質問に「感じる」と答えた人は48.9%を占めた(図1)。また、「今後、教育格差は広がると思うか」という質問に「思う」と答えた人は51.2%

を占めた(図2)。以上のことから、全国の高校生の約2分の1が教育格差を感じており、すでに社会問題と呼んでも良い状況になっていることが分かる。

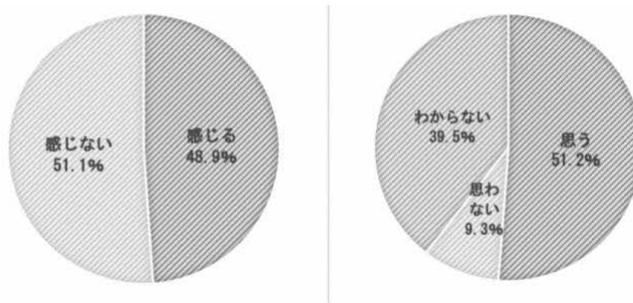


図1 教育格差を感じるか (左)

図2 今後教育格差は広がると思うか (右)

4. 研究の手順

本研究のデータは、高校生および高等学校を対象としたものを扱った。また、教育格差の生じる要因を「個人の経済面」と「学校教育における設備や学習環境の充実度」という2つの側面から分析した。さらに、それをもとに格差は正へ向けた対応と喫緊の課題を検討した。仮説の整理にあたっては、「ロジックツリー」を活用した(図3)。

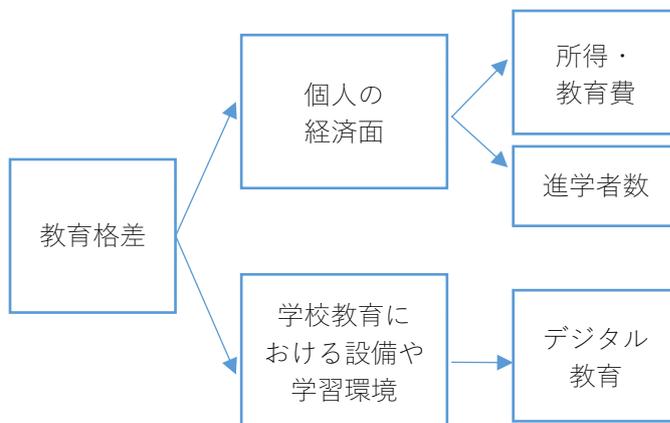


図3 仮説の整理

5. データ分析の結果

5.1 個人の経済面の視点から

まず、個人の経済面において所得と教育費の関係を調べた。

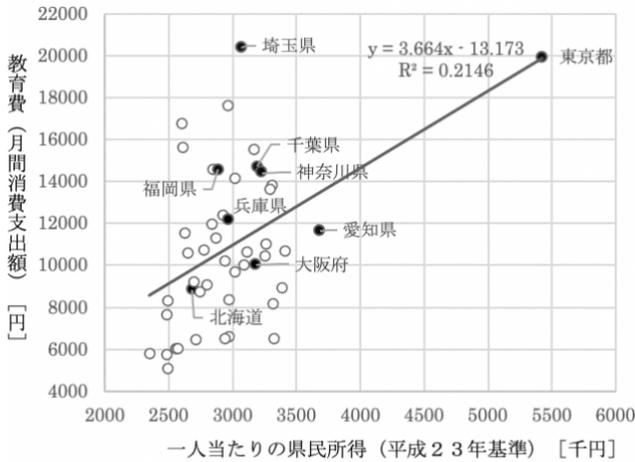


図4 一人当たりの県民所得と教育費の関係

一人当たりの県民所得と教育費には相関係数0.4632のやや強い正の相関がみられ、所得が多いほど教育費も多くなるといえる(図4)。しかし、都市部か地方かで県民所得の差があるとは読み取れず、仮説を裏付けることはできなかった。

次に、都市部と地方では、親の教育への意識に差があるのではないかという仮説を立て、進学者数、進学率、教育費、大学数、教育・学習支援事業所数の関係を調べることで仮説の検証を行うこととした。図5は相関を調べたもののうちの1つである。その結果、人口の多い地域、すなわち都市部に教育・学習支援事業所が多く、必然的に教育費も高くなっていること、また、教育費が高い地域で進学率も高くなっていることが読み取れたことから、都市部と地方の教育費の差が、教育格差を生むと判断した。

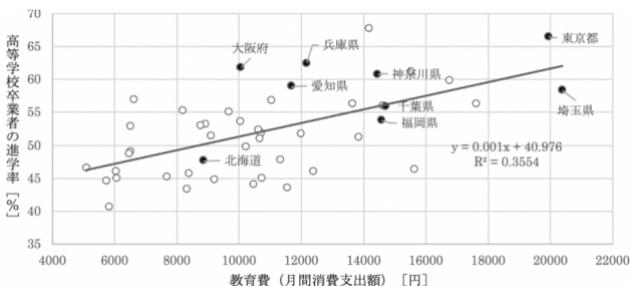


図5 教育費と高等学校卒業者の進学率の関係

5.2 学校教育における環境面の視点から

さらに、高等学校におけるデジタル教育および

その学習環境の整備の実態を表す指標として、県別の「デジタル教科書の整備率」「普通教室の電子黒板整備率」について調べた。ここでは、「都市部の方が、デジタル教育およびその学習環境の整備が進んでいる」と仮説を立てた。しかしながら、デジタル教育およびその学習環境の整備率は地方の方が高いことが分かった。そこで、「高等学校数が多い都道府県では整備が行き届かない」「そもそもすべての都道府県で整備率が低い」という2つの仮説を立て、高等学校数も加味してデータを整理した結果、2つの仮説が正しいことがわかった。

	整備率 [%]	高等学校数	整備済み高等学校数
東京都	33.9	428	145
愛知県	17.5	222	39
大阪府	21.0	256	54
石川県	93.3	56	52
佐賀県	90.7	46	41
岐阜県	81.1	81	66

表1 【デジタル教科書】都市部3都府県(東京都、愛知県、大阪府)と整備率上位3県(石川県、佐賀県、岐阜県)の高等学校数の比較

6. 考察と課題

以上の分析結果から、都市部の方が地方よりも教育において有利な面が多いと結論づけた。都市部には大学や教育・学習支援事業所が多いことも、高等学校卒業者の進学率に影響を与えていると分かった。教育格差の是正には、地方の世帯が教育にかかる費用を増やせるような施策、また、地方の高校生への教育支援・学習支援が必要である。

デジタル教育およびその学習環境の整備については、都市部、地方関係なく格差が生じていた。これは、各都道府県で導入への取り組みが大きく異なるからだと考えられる。デジタル教科書の整備率、普通教室の電子黒板整備率の全国平均はともに50%を下回っている。地域による教育格差を広げぬよう、またデジタル教育において日本が世界に遅れをとらぬよう、早急な整備と対応を強く願う。

参考文献

(1)日本財団：18歳意識調査「第33回-教育格差-」

ボランティア活動の決定要因

慶應義塾高等学校三年 今泉 開

1. データ分析の目的

本研究のデータ分析の目的は、ボランティア活動の決定要因を実証分析の観点から検証する事である。近年、日本におけるボランティア活動者率は他の先進国と比較すると平均より低いのが現状であり、ボランティア活動の活性化が求められる。また都道府県によってボランティア活動者率が大きく異なることから、都道府県単位でボランティア活動者率の決定要因が存在すると推測できる。そのため、都道府県単位のデータを用いてボランティア活動の決定要因を特定し、それに基づいた施策を策定すればボランティア活動者率を高めることができると考えた。筆者がボランティア活動に従事した際、人手不足が深刻化していたことが個人的な研究の動機である。

2. データ分析の手法

本研究は主に三つの分析により構成される。まず一つ目に、ボランティア活動者率と種別ボランティア活動者率の間でピアソンの積率相関係数を算出する事で、ボランティア活動者率の決定要因が各活動に与える影響を調べた。二つ目に、ボランティア活動者率の決定要因として可能性が示唆される変数を先行研究の結果に基づき抽出した後に経済的要因、社会資本要因、社会環境的要因に分類し、各要因から一部の変数を取り上げてボランティア活動者率との関係を、散布図を用いて分析を行なった。三つ目の分析は、経済的要因、社会資本要因、社会環境的要因に含まれる全ての変数を説明変数、ボランティア活動者率を応答変数とした重回帰分析であり、実際にボランティア活動の決定要因として影響を与えうる変数であるか検討した。また、多重共線性を調べるために各説明変数間の VIF 指数を算出した。

3. データ分析の結論

一つ目の分析から、ボランティア活動の決定要因が都道府県単位で存在し、その決定要因がそれぞれのボランティア活動に大きな影響を与えることが示された。重回帰分析、散布図を用いた分析の結果から、人口密度と3次活動の時間がボランティア活動者率の決定要因として統計的に有意な結果を示し、社会資本要因や能動的活動に関する変数を含む、それ以外の変数は統計的に有意な結果を示さなかった。以上二つの結果から、公共インフラの整備など、近年政府が実施している取り組みはボランティア活動を活性化させる効果が期待されないこと、またボランティア活動の決定要因について都道府県単位のデータを年齢や性別に分類しさらなる分析・検証が必要であることが示された。

4. データ分析の意義 (利点)

ボランティア活動者率と能動的活動の関連性について、社会学的領域では能動的活動がボランティア活動の決定要因となることが示唆されていたが、これは定性的な分析であり、実証分析の観点から能動的な活動が決定要因であるか否かを検討したのは本研究以外管見の限り存在しない。

先行研究ではボランティア活動の決定要因を全て個票データに基づいて分析していたが、本研究では都道府県単位のデータを使用して分析を行なった。結果からも、都道府県単位でボランティア活動の決定要因について分析することの重要性を見出した。

また、本研究ではボランティア活動に従事する個人の属性の違いは考慮せず、都道府県ごとに測定可能な説明変数を用いて分析を行うことで、全ての人が普遍的に影響を受けるボランティア活動の決定要因に着目して分析を行なった点で他の先行研究と差別化する事ができる。

市区町村というミクロ的視点から投票率の実態を探る

渡辺安之・林蔚欣（茨城県立並木中等教育学校）

〒305-0044 茨城県つくば市並木4丁目5-1 029-851-1346

1. 研究の背景と目的

私は学校の探究テーマ「民主主義はオワコンか」において、Economist Intelligence Unit の出す民主主義指数や経済指標などの相関を見るマクロ的解析を行なっているが、民主主義が一概に定義できない点を踏まえると、ミクロ的視点での分析が必要だと感じた。また民主主義を「参加と責任のシステム」と言えることを知り、選挙に興味を持った。本研究は、これまでの先行研究の現状と課題を踏まえ、第49回衆院選(小選挙区)の投票率を対象に、SSDSE-A-2022のデータセットを活用して市区町村というミクロ的視点から地域属性がどのように投票率に影響するのかを解析し、投票率に影響する地域属性を明らかにすることを目的とする。

2. 解析に使ったデータと手法

研究の目的を達成するため、本研究は以下に示した手順とデータ解析手法を用いた。

- ① SSDSE-A-2022にある124の市区町村の地域属性データを基に、先行研究や地域属性データの特徴を加味して地域属性解析用データセット(94項目×1682市区町村、福島県の59市町村を除き)を作成した。
- ② 全1682市区町村の投票率について、基本統計量やヒストグラムを用いた概要確認を行なった。
- ③ 個々の地域属性と投票率の散布図(94個)を用いた地域属性変数スクリーニング評価により、地域属性変数を94項目から32項目に絞り、投票率と32項目の地域属性変数の相関係数を表にまとめた。
- ④ 32項目の地域属性変数同士の相関行列の相関係数(絶対値0.65以上)を用いて地域属性変数を19項目に減らした。
- ⑤ さらに19項目の地域属性変数同士の相関行列とVIFを用いて、それぞれの変数が重回帰分析に使用するに相応しい(有意水準10未満)ことを確認した。うえ、重回帰分析を行なった。

- ⑥ 宮崎県と茨城県を対象に得られた重回帰式を用いた予測と比較検討で投票率に影響する地域属性変数を考察した。

3. 主な解析結果

1682市区町村の投票率の平均は58.77で、標準偏差は8.15だった。全市区町村において投票率が最も高い地域は、熊本県球磨村(94.74)と宮崎県西米良村(93.63)、最も低い地域は東京都稲城市(35.91)と茨城県神栖市(41.67)であることがわかった。市区町村間の投票率の差は58.83ポイント(熊本県球磨村94.74-東京都稲城市35.91)で、これは都道府県レベルの地域間投票率の差である15.25ポイント(山形66.34-茨城51.09)と比べると、投票率には大きな地域差があることを確認できた。

投票率に影響を与える地域属性変数を特定するため、94個の相関散布図によるスクリーニング評価、相関行列による多重共線性変数排除、相関行列とVIFによる多重共線性変数確認など一連の解析を実施した。その結果、投票率に影響を与える12個の地域属性変数(表1)を特定し、式1の重回帰式が得られた。重回帰式(式1)から、1682市区町村の投票率(目的変数)は10個の地域属性(説明変数)に影響されていることを示した。
$$Y = -16.552X_1 - 3.421X_2 - 2.897X_3 - 2.486X_4 + 1.444X_5 + 0.369X_6 - 0.361X_7 - 0.317X_8 + 0.166X_9 + 0.065X_{10} + 83.333 \quad (式1)$$

* X_1 : 離婚件数の割合; X_2 : 財政力指数; X_3 : 完全失業者数の総人口に対する割合%; X_4 : 医師数の総人口に対する割合%; X_5 : 教員数(小中義務合計)の総人口に対する割合%; X_6 : 15歳未満人口の割合%; X_7 : 民生費(市町村財政)の占める割合%; X_8 : 15~64歳人口の割合%; X_9 : 非労働力人口の総人口に対する割合%; X_{10} : 人口密度(可住地面積)

さらに、得られた重回帰式(式1)を評価するため、宮崎県と茨城県を対象に投票率を予測し実際のもの

と比較した。その結果（図1）から、茨城県のような県全体として投票率が低い県にはこの重回帰式が使えないことを示唆した。また、予測精度のよい宮崎県（図1）における市町村地域属性変数と投票率との関係の検証から、得られた主な結果として表2にまとめた。

4. 考察と今後の課題

本研究は、国や都道府県というマクロ的視点の従来研究と違い、市区町村というミクロ的視点から投票率の実態を探った新規性のほか、先行研究に見られなかった多く

の地域属性（離婚率、教員数、財政力指数、民生費、15歳人口の割合など）に対して独自の結果が得られた。厳密な因果関係の解析や交絡因子の考察、数年分のデータを使ったより長期的視点での解析などは今後の課題として残されているが、多くの地域属性と投票率との関係を明らかにしたことは有益である。これらの結果は今後の投票率向上の方策につながることを期待したい。

表1 特定された地域属性変数とその重回帰分析結果

回帰統計									
重相関 R	0.775								
重決定 R2	0.600								
補正 R2	0.597								
標準誤差	5.199								
観測数	1682								
分散分析表		自由度	変動	分散	観測された分散比	有意 F			
回帰		12	67621.351	5635.113	208.511	0.000			
残差		1669	45105.530	27.025					
合計		1681	112726.881						
	切片	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
X ₁ : 離婚件数の割合%		83.333	4.336	19.219	0.000	74.828	91.837	74.828	91.837
X ₂ : 財政力指数		-16.552	2.998	-5.520	0.000	-22.433	-10.671	-22.433	-10.671
X ₃ : 完全失業者数の総人口に対する割合%		-3.421	0.733	-4.668	0.000	-4.859	-1.983	-4.859	-1.983
X ₄ : 完全失業者数の総人口に対する割合%		-2.897	0.281	-10.302	0.000	-3.449	-2.345	-3.449	-2.345
X ₄ : 医師数の総人口に対する割合%		-2.486	0.733	-3.394	0.001	-3.923	-1.049	-3.923	-1.049
X ₅ : 教員数(小中義務合計)の総人口に対する割合%		1.444	0.249	5.796	0.000	0.956	1.933	0.956	1.933
X ₆ : 15歳未満人口の割合%		0.369	0.083	4.446	0.000	0.206	0.532	0.206	0.532
X ₇ : 民生費(市町村財政)の占める割合%		-0.361	0.022	-16.364	0.000	-0.405	-0.318	-0.405	-0.318
X ₈ : 15~64歳人口の割合%		-0.317	0.054	-5.921	0.000	-0.422	-0.212	-0.422	-0.212
X ₉ : 非労働力人口の総人口に対する割合%		0.166	0.051	3.245	0.001	0.066	0.267	0.066	0.267
X ₁₀ : 人口密度(可住地面積)		0.065	0.007	9.221	0.000	0.051	0.079	0.051	0.079
X ₁₁ : 婚姻件数の割合%		1.639	1.523	1.076	0.282	-1.348	4.627	-1.348	4.627
X ₁₂ : 第1次産業就業者数の総人口に対する割合%		0.046	0.035	1.307	0.192	-0.023	0.115	-0.023	0.115

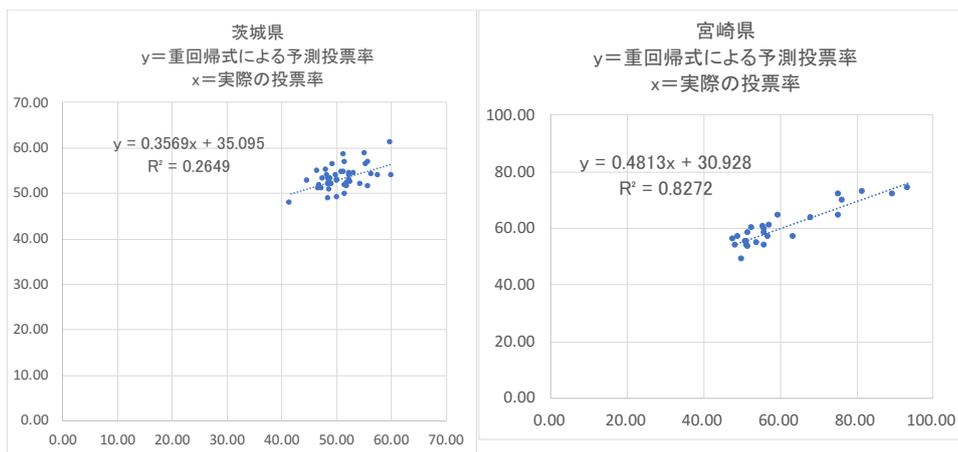


図1 重回帰式による予測精度や予測傾向の比較(茨城県と宮崎県)

表2 本研究の解析から得られた主な結果のまとめ

分類	本研究で得られた示唆
本研究の解析から得られた結果(これまでにない新しい知見)	<ul style="list-style-type: none"> ・教員数が多いほど投票率が高い傾向。 ・離婚率が高いほど投票率が低い傾向。特に興味深いのは、宮崎県内において投票率ベスト1の西米良村とベスト2の諸塚村は共に離婚率が0であった。 ・民生費が高いほど投票率が低い傾向。 ・財政力指数が高いほど投票率が低い傾向。 ・15歳の人口が高いほど投票率が高い傾向。
先行研究などに概ね一致する発見	<ul style="list-style-type: none"> ・完全失業者が多いほど投票率が低い傾向にある ・人口密度が高いほど投票率が低い傾向にある
先行研究などと一致しない発見	<ul style="list-style-type: none"> ・医師数の差は投票率にあまり影響がなさそう

統計的に探究することができる児童が育つ授業の在り方

名古屋市立神宮寺小学校
日比野 浩規

I 主題設定の理由

1 社会的背景の側面から

世界は、感染症をはじめとする疾病や異常気象、地震などの災害、経済やビジネスの複雑化などによって、VUCA(Volatility、Uncertainty、Complexity、Ambiguity：予測困難で不確実、複雑で曖昧)となる時代に直面している。OECD(2020)では、Education 2030プロジェクトにおいて、生徒が生き抜く、よりVUCAな時代を歩んでいくための重要な力の一つとして、「生徒エージェンシー(Student agency)(以下、エージェンシー)」を挙げている。コンセプトノート(2019)では、生徒の対象を小学校の児童、幼児など、すべての学習者と捉え、その定義や特性について以下のように述べられている。

エージェンシーとは、生徒が社会に参画し、人々、事象、および状況をよりよい方向へ進めるといった変革を起こすため、目標を設定し、振り返りながら責任ある行動をとる能力として定義づけられる。つまり、働き掛けられるというよりも自らが働き掛けることであり、型にはめ込まれるというよりも自ら型を作ることであり、また他人の判断や選択に左右されるというよりも責任をもった判断や選択を行うことを指す。エージェンシーは人格特性ではなく、伸ばすことも学ぶこともできるものである。

一方、日本財団(2019)が行った「18歳意識調査」(図1)の結果から、白井(2020)は、諸外国と日本財団(2019)が行った「18歳意識調査」とを比較して日本の教育におけるエージェンシーの弱さが示唆されることを課題として取り上げている。

	自分を大人だと思う	自分は責任がある社会の一員だと思う	将来の夢を持っている	自分で国や社会を変えられると思う	自分の国に解決したい社会課題がある	社会課題について、家族や友人など周りの人と議論ができる
日本	29.1%	44.8%	60.1%	18.3%	46.4%	27.2%
インド	84.1%	92.0%	95.8%	83.4%	89.1%	83.8%
インドネシア	79.4%	88.0%	97.0%	68.2%	74.6%	79.1%
韓国	49.1%	74.6%	82.2%	39.6%	71.6%	55.0%
ベトナム	65.3%	84.8%	92.4%	47.6%	75.5%	75.3%
中国	89.9%	96.5%	96.0%	65.6%	73.4%	87.7%
イギリス	82.2%	89.8%	91.1%	50.7%	78.0%	74.5%
アメリカ	78.1%	88.6%	93.7%	65.7%	79.4%	68.4%
ドイツ	82.6%	83.4%	92.4%	45.9%	66.2%	73.1%

[図1 18歳意識調査]

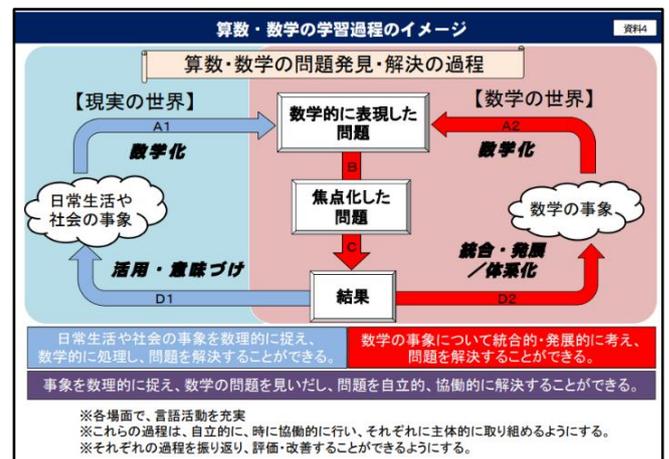
また、内閣府は、VUCAとなる時代における日本のビジョンとして、Society5.0の社会を掲げている。この社会においては、膨大なデータが世の中にあふれ、

経済だけでなく、生き方も変化することが求められている。実際にあらゆる業種や領域においても、ビッグデータを活用して、他者へ情報発信したり、意思決定したりすることが重要視されており、そのためのプロセスは必須となってきた。

これらの社会的背景から、児童が将来、エージェンシーを発揮しながらデータを活用できることが、よりよい社会と幸福な人生の創り手となっていくための重要な要素の一つであることが分かる。そして、学校教育においては、その授業の在り方を明らかにすることが喫緊の課題であると考えた。

2 小学校における統計教育の側面から

小学校学習指導要領(平成29年告示)解説算数編では、社会の急速な変化に対応するために必要な資質・能力を育成することの重要性が述べられている。また、こうした資質・能力を算数科で育成していくためには、学習過程の果たす役割が極めて重要であるとされており、算数・数学の学習過程のイメージについて、以下の図で示されている。(図2)



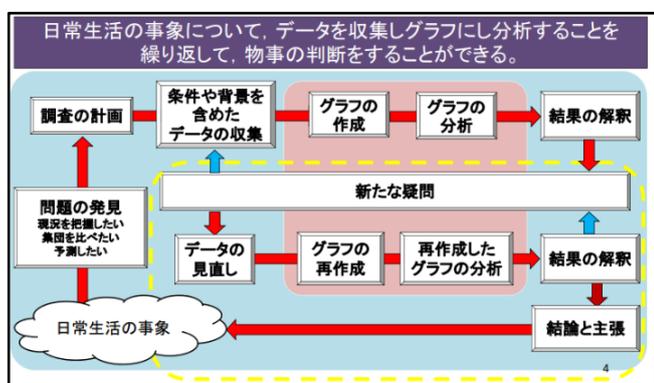
[図2 算数・数学の問題発見・解決の過程]

また、質・量ともに不足している統計教育の現状を改善することが目指され、指導事項「データの活用」領域が新設された。そして、統計的な内容の改善・充実として、「様々な場面において、必要なデータを収集して分析し、その傾向を踏まえて課題を解決したり意思決定をしたりすること」が強調されている。さらに、統計的な問題解決活動における「統計的探究プロセス(PPDACサイクル)」(図3)が示されている。

問題	・問題の把握	・問題設定
計画	・データの想定	・収集計画
データ	・データ収集	・表への整理
分析	・グラフの作成	・特徴や傾向の把握
結論	・結論付け	・振り返り

[図3 統計的探究プロセス]

算数・数学ワーキンググループ(2016)では、統計的な問題解決活動について、より詳しく紹介されている(図4)。このことから、統計的な問題解決の方法は、図2のうち、現実の世界を通る過程(プロセス)となるものであることが分かる。



[図4 統計的な解決活動の詳細]

特に、高学年においては、一連の統計的探究プロセスを意識し、自分たちで問題を設定し、調査計画を立てることや、分析を通じて判断した結論についても別の観点から妥当性を検討できるようにすることも扱うと述べられている。また、統計は社会における必須ツールとなっており、統計的な問題解決のよさを感じ、進んで学習や生活に生かそうする態度を養うことも述べられている。

しかし、現行の小学校算数科の教科書では、問題は設定され、整えられたデータや分析方法があらかじめ提示されているものがほとんどである。そのため、児童が目的意識をもち、調査方法や分析方法を考えたり、結論について別の観点から再検討したりするなど、本来、小学校段階における統計教育で目指すべき姿を引き出すことが難しくなってしまうと考えられる。

実際に、勤務校の教員から「データの活用」領域について、「従来の指導から脱却できない」「知識・技能の習得が中心となってしまう」「児童に問題設定をさせることが難しい」といった困惑の声が聞かれた。

このことから、「データの活用」領域における統計的探究プロセスを活用した授業の在り方について明らかにすることが求められていると考えた。

3 研究主題について

以上のことから、「統計的に探究することができる児童が育つ授業の在り方」を主題として設定した。

本研究でいう統計的に探究することができる児童とは、日常生活や社会に変革を起こすための目標を設定し、統計的探究プロセスを活用した問題解決を通して、責任をもった判断や選択を行うことができる児童と捉えることにする。つまり、「エージェンシー」と「統計的な問題解決」の二つの側面を満たす児童のことである。

II 課題の焦点化

本研究につながる課題について、「統計的探究プロセス」「カリキュラムマネジメント」の観点から述べる。

1 統計的探究プロセスの観点から

小学校第6学年「資料の調べ方」において、児童に統計的探究プロセスを活用させた筆者の実践研究から、統計的探究プロセスの観点において以下の課題を捉えた。

- ・ 「計画」の段階で、よりよい結論を見出すために、多面的な視点で扱うデータや分析方法まで、見通しをもって考えることが容易ではないこと。
- ・ 「分析」の段階で、適切な判断や選択につながる結論を見出すために、試行錯誤しながらデータを分析することが難しいこと。

上記の課題を踏まえ、本研究では、「計画」「分析」の段階で手立てを講じることにした。

また、「分析」の段階では、GIGAスクール構想で、一人一台タブレットが配布されたことを生かして、統計分析ソフトを活用させ、適切なグラフを選択させたり、階級幅を工夫させたりすることが容易にできるようになることで、試行錯誤しながら多面的に考察することができるようにする手立てを明らかにしたいと考えた。

2 カリキュラムマネジメントの観点から

小学校第6学年「資料の調べ方」において単元計画(10時間完了)を考え、実践を行った際、日常生活や社会において解決したい問題を話し合って目標を設定することや、分析結果や結論を発信するための資料作りに時間をとられた。そのため、知識・技能に関する指導時間が圧迫され、教師と児童の双方に負担が生じてしまった。また、エージェンシーの側面において、地域社会への参画を通じて、人々や物事、環境がよりよいものとなるよう影響を与えるという責任感をもって学習に取り組むことが重要となるが、これを実現することは算数科の時間だけでは難しいと感じた。

関連して、小学校学習指導要領解説(平成 29 年告示)解説総則編には、カリキュラムマネジメントの充実について述べられており、その側面の一つとして、「児童や学校、地域の実態を適切に把握し、教育の目的や目標の実現に必要な教育の内容等を教科等横断的な視点で組み立てていくこと」と述べられている。とりわけ、日本学術会議の数理科学委員会数学教育分科会の提言「新学習指導要領下での算数・数学教育の円滑な実施に向けた緊急提言：統計教育の実効性の向上に焦点を当てて」(2020)では、「小学校学習指導要領(平成 29 年告示)解説算数編に示された算数・数学の学習過程のイメージ(図 2)のような数学的活動には、より深く追究する時間が必要であり、また、他教科と連携する必要が生じると想定されることから総合的な学習の時間と連携することが望ましい」と述べられている。

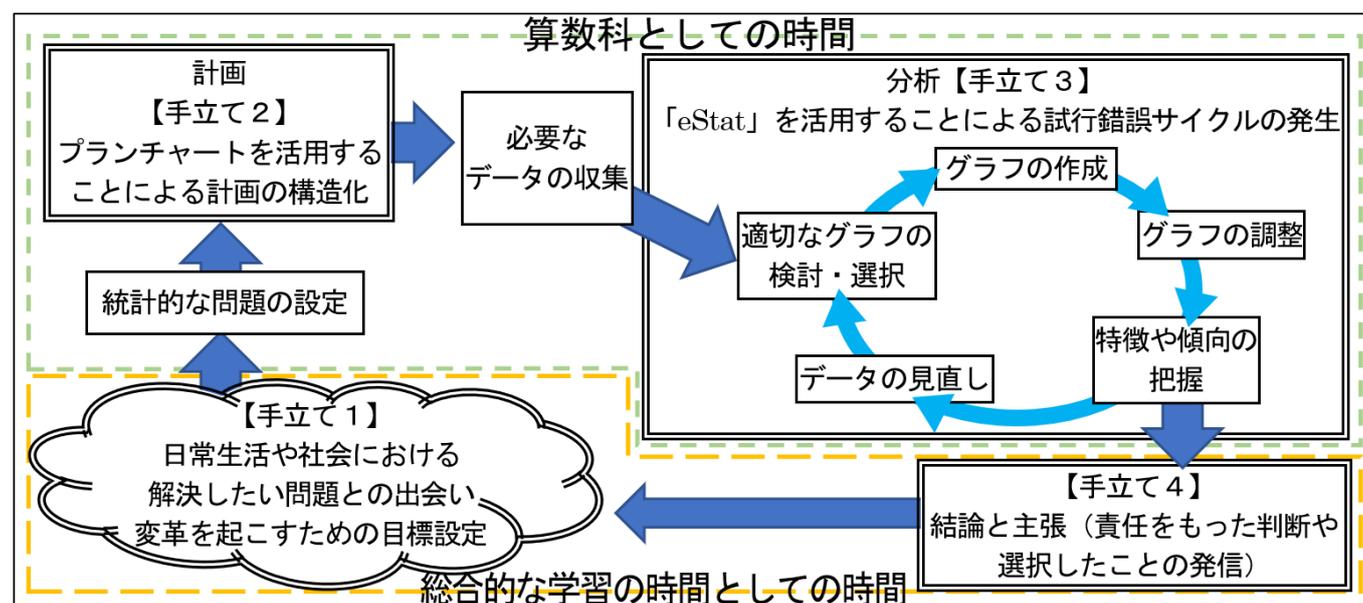
上記の課題を解決するために、総合的な学習の時間との関連を図った統計的探究プロセスを活用した単元構成を明らかにすることが必要であると考えた。

Ⅲ 研究の構想 (対象・小学校 6 年生)

1 研究仮説

総合的な学習の時間との関連を図った統計的探究プロセスを活用した単元構成において、変革を起こすための目標設定や自らの判断や選択を発信する態度を高めること(エージェンシーの側面)や、「計画」の段階で、計画を構造化したり、「分析」の段階で試行錯誤したりすること(統計的な問題解決の側面)ができるような指導の工夫を行えば、統計的に探究することができる児童が育つだろう。

2 研究構想図 (図 5)



[図 5 研究構想図：統計的に探究することができる児童が育つ授業の在り方]

3 研究の手立て

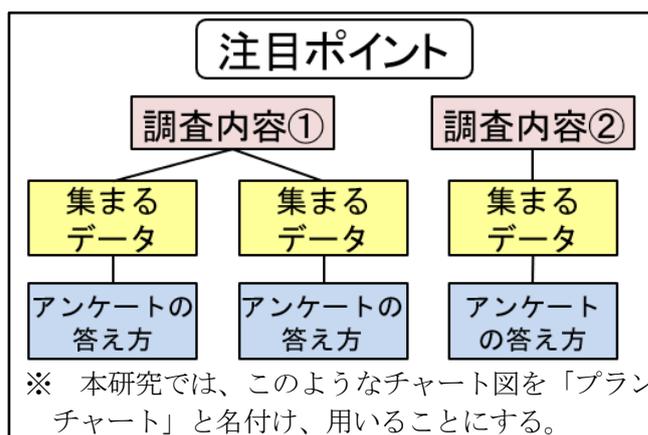
手立て 1・4 は、「エージェンシーの側面」、手立て 2・3 は、「統計的な問題解決の側面」に関わる。

【手立て 1：変革を起こすための目標設定】

切実感や改善の必要感のある日常生活や社会の問題に直面させた後、「この問題に対して、あなたが改善したいことや調べて解決したいと思ったことはありませんか」と問い掛け、問題を自分ごととして捉えさせた上で、表現する場を設けることで、変革を起こすための目標設定をしようとする態度を高めることができるようにする。

【手立て 2 計画の構造化】

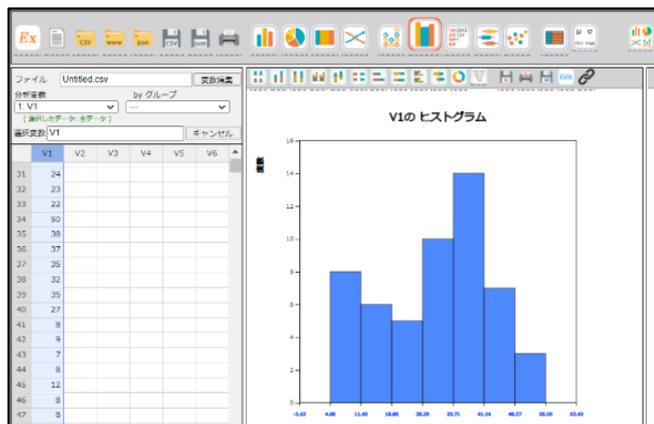
「計画」の段階で、注目した日常生活や社会の問題について、現状を明らかにするために調べること(調査内容)として、思い付くものを付箋(ピンク色)に書いて意見を出し合う。その後、分類・整理した上で、プランチャート(図 6)を活用して、集まるデータ(黄色)・アンケートの答え方(水色)について話し合わせることで、計画を構造化できるようにする。



[図 6 計画を構造化するためのプランチャート]

【手立て3：試行錯誤サイクルの発生】

「分析」の段階では、はじめに分析の目的(だれに何を伝えるために分析するのか)を明確にさせる。その後、図5の「試行錯誤サイクル」を児童に提示した上で、統計分析ソフト「eStat」(図7)を活用して、目的に合ったグラフの種類や階級幅について話し合わせることで、試行錯誤しながら多面的に考察することができるようにする。



[図7 統計分析ソフト「eStat」]

【手立て4：責任をもった判断や選択の発信】

「発信する対象(学校・地域・家庭など)」「発信する目的」について、児童に設定させる。そして、実際に発信させた後、受け手からフィードバックを受ける

[表1 単元計画「目指せSDGs『水不足～私たちが世界を変える～』」]

時数	教科領域	統計的探究プロセス	主な学習活動
1・2時 LIXIL 出前授業	総合	問題 【手立て1】	※ 本時までには、SDGsの意味と17のゴールについて調べている。本単元では、ゴール6「安全な水とトイレを世界中に」に着目して学習を進める。 企業の出前授業を通して「安全な水とトイレを世界中に」について企業の努力と世界の現状を学び、改善したいことや調べて解決したいことといった視点で、変革を起こすための目標設定をする。
3時	算数	問題	「ゴール6を達成するために関係のあることについての現状を調べ、その向上について考えよう」という統計的な問題を捉え、その注目ポイントを考える。 調べたい注目ポイントが同じ児童でグループを作り、結論を伝える対象について話し合う。
4～6時		計画 【手立て2】	注目ポイントの現状を明らかにするために調べることにについて意見を出し合い、分類・整理する。その後、プランチャートを活用して計画を構造化する。
7時		データ	アンケートを作成し、必要なデータを収集する。(ロイロノートのアンケート機能の活用) ※ 授業時間は、アンケート作成の時間とし、回答は、休み時間または宿題として行う。
8～11時		分析 【手立て3】	一つの注目ポイントを例に、資料のちらばりについて、ドットプロット・度数分布表・ヒストグラムを用いた分類・整理する方法を学ぶ。(知識・技能の習得) グループごとに「eStat」を活用して、収集したデータを分析し、結果から傾向を読み取る。
12～14時	総合	結論 【手立て4】	水不足について、自分たちの生活や社会をどのように変えていきたいか、そして、何を呼び掛けたいか結論を導き出す。日常生活や社会を変革するために呼び掛ける資料は、ロイロノートを活用して作成する。(対象への発信、フィードバックを受ける機会は後日)

機会を設けることで、責任をもった判断や選択を発信しようとする態度を高めることができるようにする。

4 検証(実践前：6月1日 実践後：9月30日)

【エージェンシーの側面】

実践前後に「日常生活や社会の問題に対して、データを活用して、改善したいことはありますか」と問い掛け、それに対する記述の変容から、日常生活や社会に変革を起こすための目標を設定し、責任をもった判断や選択を行おうとする児童が育ったか検証する。

また、実践前後に、日本財団(2019)が行った「18歳意識調査」を行い、その変容や図1の日本の結果と比較し、より広い視点で、エージェンシーの側面が育ったか検証することも行う。

【統計的な問題解決の側面】

日常生活や社会においてデータを活用して解決することができる調査問題について、実践前後の記述内容の変容から、解決に統計的探究プロセスを活用することができる児童が育ったか検証する。

IV 実践

本研究では、2022年6月に筆者が担任である小学校6年生の学級(32人)を対象にして図5に基づき、単元「目指せSDGs『水不足～私たちが世界を変える～』」(表1)を構成して実践を行った。

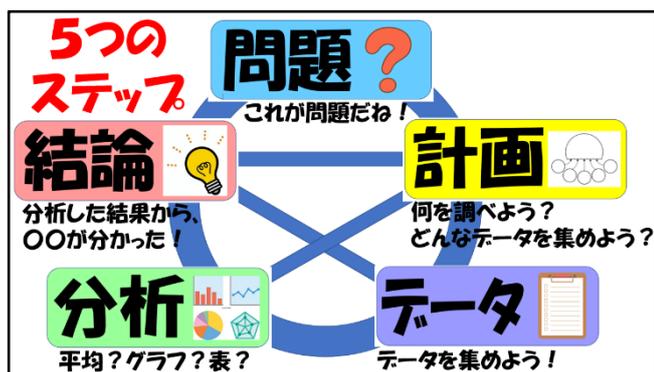
1 実践の様子

1. 1 第1・2時 問題の段階【手立て1】

本校に LIXIL の方を招き、出前授業における体験活動を通して、水に関する世界の現状や水を大切にすることの重要性、水に関して日本が恵まれていることを実感できるようにした(図8)。授業では、「世界には、きれいな水が十分に手に入らない人が多くいること」「私たちは、1日に約300Lものきれいな水を使っていること」「VWで考えると、世界の貴重な水を使い、捨てているということ」などに触れ、児童がもっている水に対する感覚に対してギャップを感じさせた。そして、SDGsの目標6「安全な水とトイレを世界中に」についての課題意識を高めていった。最後に、LIXILの方から、LIXIL名古屋ショールームに、児童が結論の段階でまとめた資料を掲示することを提案された。その後、統計的探究プロセス(図9)を提示し、単元の見通しをもつことができるようにした。図9については、単元を通して教室に掲示しておき、児童が、統計的探究プロセスを意識できるようにした。



[図8 すごろく「しずくの大冒険」に取り組む児童]



[図9 提示した統計的探究プロセス]

「手立て1」では、出前授業を受けて、改善したいと思ったことや調査して解決したいと思ったこと、水に対する思いをロイロノート上に表現させ、学級で共有できるようにした。32人中28人の児童が、課題意識をもった目標を立てることができていた。

《児童の記述内容（一部抜粋）》

- ・ 給食の残量の VW を調べて全校に知らせることで、残量を減らしたり、水を大切にすることを高めたりできるようにしたい。
- ・ 1日に300Lの水を使っていることを知って驚いた。このことをみんなに呼び掛けて、節水意識を高められるようにしたい。
- ・ トイレやシャワーの使用量の現状を調べて、節水を呼び掛けたい。
- ・ 日本は、恵まれていることが分かった。だから、世界の現状をさらに調べて、みんなに伝えることで、水を大切にできるようになってほしい。

また、記述内容を考察すると、出前授業で感じたギャップを多くの人に知らせたいという思いだけでなく、神宮寺小学校の児童や保護者から必要なデータを集め、分析したことを根拠にして、水の大切さを呼び掛けるといった統計的探究プロセスを意識した内容が32人中20人の記述に見られた。

1. 2 第3時 問題の段階

まず、児童の「水についての現状をさらに調べたい」という思いから、統計的な問題「SDGs ゴール6を達成するために関係のあることについての現状を調べ、その向上について考えよう」を捉えさせた。そして、何に注目すればよいかという「注目ポイント」を考えさせた。発表された考えについて、話し合いを通して分類・整理させると、以下のようになった。

《児童が考えた注目ポイント》

- ・ 1日に使う水の量
 - ・ トイレ
 - ・ 洗濯
 - ・ 手洗い
 - ・ 食器洗い
 - ・ 風呂とシャワー
 - ・ きれいな水が使えない国
 - ・ 給食の残量と VW
 - ・ 料理と VW
- ※ VW：バーチャルウォーター(仮想水)

その後、調べたい注目ポイントが同じ児童2～5人で1グループを作り、結論を伝える対象について話し合う時間を設けた。結果、伝える対象は、LIXIL名古屋ショールームで働く社員と客、神宮寺小学校の全校児童と教職員、保護者となった。

1. 3 第4～6時 計画の段階【手立て2】

「手立て2」では、まず、自他のグループに関係なく、それぞれの注目ポイントの現状を明らかにするために調べることにについて、思い付くものをピンク色の付箋に書かせ、全体で意見を出させた(図10)。



〔図10 全体で意見を出した後の板書〕

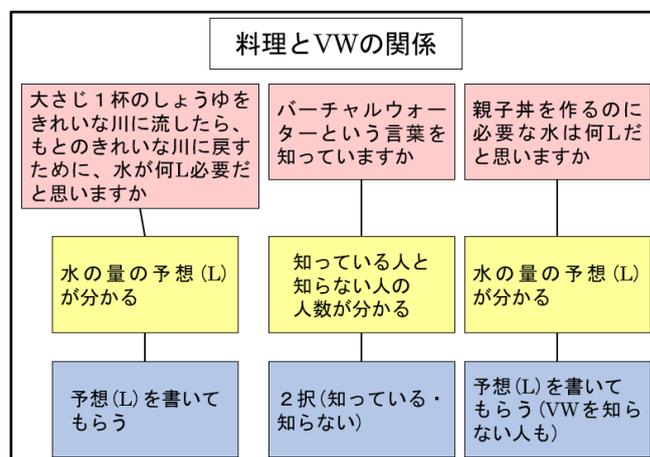
次に、グループに分かれ、集まったピンク色の付箋について、「内容が似ているものはないか」「注目ポイントに関係している内容か」「他に考えられることはないか」といった視点で話し合わせ、分類・整理・追加できるようにした(図11)。



〔図11 付箋を分類・整理・追加をする様子〕

その後、児童は、プランチャート(図6)を活用して、計画を構造化していった。その際、アンケートの答え方で戸惑うことが予想されたため、全体で、アンケートのタイプ(選択肢・数値・自由記述など)を確認・板書してから活動させた。第4時終了時のプランチャートへの記述を見ると、調べることとして「一日に洗う皿の量は何枚ですか(食器洗いグループ)」「1日に何回(何分間)トイレに行きますか(トイレグループ)」など、調査不可能な内容や効果的でない内容が見られた。そのため、次時には、第1時に出勤授業で感じた水に対するギャップの内容を振り返らせた後、「読み手に実態と現実のギャップを感じさせ、水を大切にしたいという思いを効果的にもたせることができる内容か」といった観点でプランチャートを見直す活動を設けた。

ここでは、「料理とVWの関係」に注目して活動を進めたグループを取り上げてその様子を述べることにする。このグループでは、「料理には、たくさんのVWが含まれていることを知ってもらいたい」「食べ物を大切にしてほしい」という思いを、学校や地域の人に届けることを目標に設定していた。第5時終了時のプランチャートの記述内容は図12のようになった。



〔図12 プランチャートの記述内容〕

第6時には、他のグループに回答者の立場になってもらい、「この内容だと4択よりも、数値の方が回答しやすいと思う」「食器洗いや洗濯は、親に聞いた方がよいと思う」など、客観的なアドバイスを受けることができる機会を設けた。こうしたことで、作成したプランチャートの内容を修正したり、調査対象を変更したりすることができていた。また、中には、「平均を調べる」「ヒストグラムで表す」「棒グラフで分析する」などといった分析方法についても話し合い、プランチャートに付箋を付け加えているグループも見られた。

1. 4 第7時 データの段階

ロイロノートのアナログ機能を活用し、データを収集させた(図13)。第7時の授業時間をアンケート作成の時間とした。児童は、10分程度でアンケートを作成することができた。残りの授業時間では、結論の段階でより説得力のある発信ができるように、

〔図13 アンケート画面の一部〕

グループでのやりとりを見ると、C1～4のように、グラフを容易に変更できるといった「eStat」のよさを生かして、試行錯誤しながら適切なグラフについて話し合う姿が見られた。また、C7～9のように、読み手を意識して、効果的に表現できる階級幅を話し合う姿が見られた。

さらに、他グループでは、グラフを作成した際に、回答者の入力ミスによる外れ値を発見し、試行錯誤サイクルを基に、データの見直しを行い、グラフの再作成を行う姿も見られた。

各グループ、二つ目の調査内容を分析する際、グラフを的確に選択する姿が多く見られるようになった。そこで、「グラフの種類の決め方にポイントがあるのですか」と問い掛けると、次のようなやりとりが生まれた。

《児童とのやりとり》

T：グラフの種類の決め方にポイントがあるのですか。

C10：数値で答えるタイプのアンケートは、ヒストグラムを使えばよいです。

C11：C10さんに付け足しで、何L使いますかのように、自由に数を答えてもらうアンケートのタイプのときです。

T：他のグラフはどうしますか。

C12：4択のときは、棒グラフです。人数とか量が分かりやすくなります。

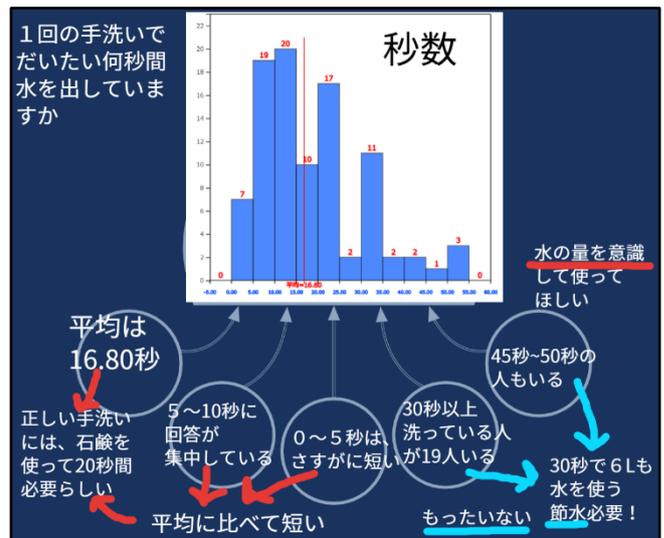
C13：2択や3択でもいいと思います。「あてはまる」「少し当てはまる」みたいな場合も棒グラフを使えばよいです。

C14：変化を表すときは、折れ線グラフだけど、今回の調査では、使いませんでした。

C15：例えば、毎日の気温の変化を調べるときには、折れ線グラフがよいと思います。

このやりとりから「eStat」を活用して、試行錯誤しながらグラフを吟味した児童は、活動を通してグラフの意味や適切な活用場面を理解できている様子がうかがえた。

その後、作成したグラフをロイロノートの共有ノートに貼り付けさせ、読み取れた特徴や傾向を、1人1台タブレットを活用して思考ツール(クラゲチャート)に書き込ませていった。考えを書き込む中で、「その考えもよいね」「この考えはどうか」といったように、交流する姿が見られた。クラゲチャートには、平均や人数が多い階級などと言った特徴や傾向だけでなく、調査結果とインターネットを活用して調べた事実を比較したことから考えた読み手に伝えたいことについての記述も見られた(図19)。



【図19 クラゲチャートの記述】

1.6 第12～14時 結論の段階【手立て4】

ここでは、分析したことから、水不足について、自分たちの生活や社会をどのように変えていきたいか、そして、何を呼び掛けたいか結論を導き出した。呼び掛けるための資料は、ロイロノートを活用してグループごとに作成させた(図20)。作成した資料は、校内だけでなく、LIXIL名古屋ショールームへ展示(2か月間)も行い、店員や訪れた客からもGoogleフォームを通して、フィードバックを受けることができたようにした。回答は、52件得ることができ、その内容は、学級で随時紹介し、導き出した結論を伝えたことによって、人々の水に対する意識を変容させるためのきっかけ作りができたことを実感できるようにした。

《Googleフォームの質問内容》

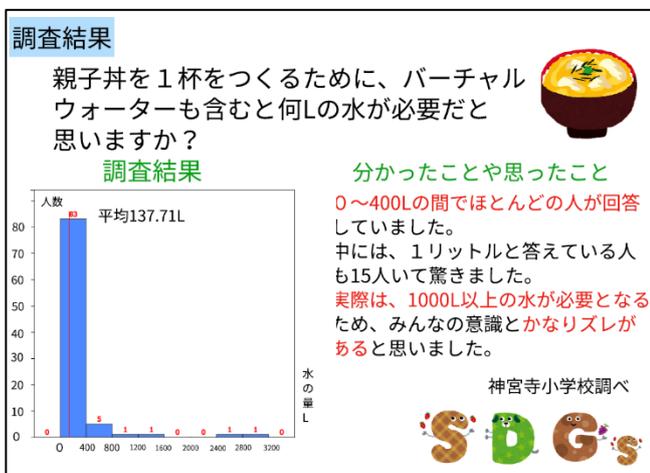
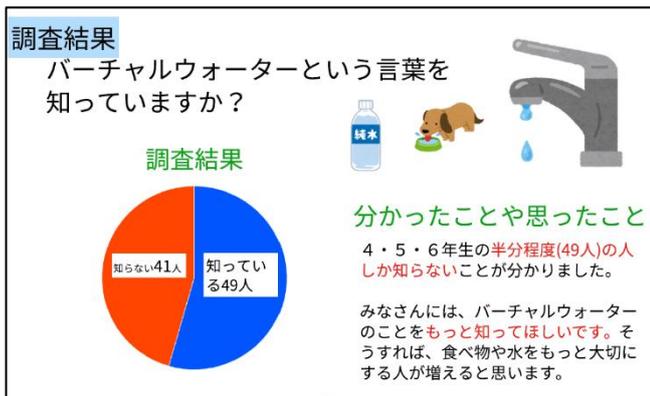
- Q1. 作品を見たことで、水を大切に(節水)しようとする気持ちが高まりましたか。
- Q2. 資料を読んだことで、水を大切に(節水)することについての新たな発見がありましたか。
- Q3. 水を大切に(節水)するために、生活を見直したり、これから何か取り組んだりしたいと思うことはできましたか。
- Q4. 資料を作成した6年生の子どもたちに、感想や思ったことなどのメッセージをお願いいたします。

【表2 Q1～Q3の回答結果(人) N=52】

質問	当てはまる	少し当てはまる	あまり当てはまらない	当てはまらない
Q1	52	0	0	0
Q2	49	3	0	0
Q3	50	1	1	0

《④の回答結果(一部抜粋)》

- ・ 水の大切さについて知ることができました。バーチャルウォーターの話は初めて聞き、ご飯1杯で、こんなにも多くの水が使われていることを知り、驚きました。
- ・ 水はつつい出しばなしにしまったり、必要以上に使ってしまったことがあります。 「今一度使い方を考えよう」と思いました。
- ・ 自分の行動がこんなにも環境に関わるのだと知り、生活を改めたいと思いました。水が当たり前に見える国で生活できるからこそ、より大切にしなければいけないことを実感できました。



伝えたいこと

私達の生活には水がたくさん必要であり、水の大切さを改めて感じました。
みなさんには、この資料をきっかけにバーチャルウォーターのことを知って、周りの人に水を大切に思う気持ちを広めてほしいと思います。
そして、みんなが節水や環境のことを考えるようになればいいなと思います。

[図20 料理とVWの関係を調べたグループが作成した資料(一部抜粋)]

読み手からのフィードバックを受けた児童からは、「節水意識を高めることができてうれしい」「私たち

でも、たくさんの人の心を変えることができると分かった」などの発言が聞かれ、社会に参画し、貢献できたことを実感している様子が見えたと。

2 検証

「エージェンシー」と「統計的な問題解決」の側面について、実態調査の結果から検証した。

【エージェンシーの側面】

実践前後に次の二つの調査を行うことで、エージェンシーの側面が育ったか検証した。

- ① 生活や社会の問題に対して、データを活用して、改善したいことはありますか」と問い掛け、それに対する記述から検証した(表3)。
- ② 日本財団(2019)が行った「18歳意識調査」を行い、その変容や図1の日本の結果と比較し、より広い視点で、エージェンシーの側面が育ったか検証した(表4)。実践前後のデータについて数値を比較することで分析した。

[表3 ①の調査結果 N=32]

	6月	9月
記述することができた	5	30
記述することができなかった	27	2

《児童の記述内容(一部抜粋)》

- ・ ながらスマホの実態と交通事故の減少
- ・ ゲームの時間と生活改善
- ・ マイクロプラスチックの現状とその対策
- ・ 命についての考え方と自殺予防

[表4 ②の調査結果 N=32(6・9月の結果のみ)]

調査内容	「はい」と回答した人数(割合)		
	6月	9月	日本財団*
Q1.自分は責任がある社会の一員だと思いますか	17 (53.1%)	24 (75.0%)	44.8%
Q2.自分で国や社会を変えられると思いますか	5 (15.6%)	10 (31.2%)	18.3%
Q3.自分の国に解決したい社会問題がありますか	11 (34.3%)	25 (78.1%)	46.4%
Q4.社会問題について、家族や友人など周りの人と積極的に議論していますか	6 (18.8%)	12 (37.5%)	27.2%

※ 図1の日本財団(2019)が行った結果

表3から、実践前後で記述できた人数に変容が見られた。記述内容を分析すると、実践前は、学校生活に関わる記述のみであった。一方、実践後は、社会生活や環境といったより広い視野についての具体的な記述が32人中23人に見られ、児童の、社会に参画し、人々、事象、および状況をよりよい方向へ進めようとする態度に変容したことがうかがえる。

また、表4から、全ての調査項目において、実践後に数値の増加が見られた。さらに、日本財団の結果と比較しても、割合が高いことが分かる。このことから、変革を起こすための目標設定をさせたり、責任をもった判断や選択を発信させたりする手立てを講じたことが、児童のエージェンシーの側面の育成に有効であったと言える。しかし、Q2・Q4においては、成果は見られたものの、十分な結果には至らなかった。これらは、日本財団の結果を見ても、数値が低くなっている。この課題を解決するために、教科・学年の枠を越えた系統的な視点で、カリキュラムをデザインし、児童が本実践のような経験を積み重ねていくことが必要であると考えられる。

【統計的な問題解決の側面】

日常生活や社会においてデータを活用して解決することができる調査問題(図21)を用いて、実践前後の記述内容について、統計的探究プロセスの各段階に着目できているか、その変容から検証した(表5・6)。合計点別のデータについては、t検定により分析した。

6年1組の学級では、1か月後に健康な生活について全校に呼びかけることになりました。全校のみんなの意識を効果的に高めるために、説得力のある呼びかけが必要です。



ひろきさんのグループでは、全校のみんなにアンケートをとるというアイデアを考えました。しかし、説得力のある呼びかけができるようにするためにどのようなアンケートをすればよいか困っています。ひろきさんのグループが説得力のある呼びかけができるように、アドバイスをあげましょう。

[図21 調査問題(実践後)]

[表5 実践前後の変容データ(合計点別) N=32]

	5	4	3	2	1	0	平均値	p値
6月	0	0	0	3	12	17	0.56	0.00
9月	15	9	4	1	2	1	3.97	**

※ 一つの段階の記述につき、1点として計算 ** : p<0.01

[表6 実践前後の変容データ(各段階別) N=32]

	問題	計画	データ	分析	結論
6月	1	0	15	2	0
9月	15	30	30	28	24

※ 児童の記述から各段階の記述を抽出

表5から、統計的な問題解決の側面に関して、実践前後で有意差があることが明らかとなった。また、表6を見ると、「計画」「分析」の段階の記述が顕著に増加している。これは、計画を構造化させたり、思考錯誤サイクルを発生させたりする手立てを講じたことの成果であると考えられる。一方で、「問題」の段階についての記述は、他の段階に比べて少ないことが分かる。このことから、児童が自ら課題を設定する場をさらに充実させたり、捉えた課題を強く意識して統計的な問題解決を進めることができるように単元構成を工夫したりする必要があると考えた。

V おわりに

本研究では、児童が、将来生き抜くVUCAな時代に必要となる力である「エージェンシー」統計的な問題解決に焦点を当てて主題を設定し、そこに迫るための具体を示すことができた。今後も児童が生き抜く未来を見据え、統計的に探究することができる児童が育つ授業の在り方を明らかにしていきたい。

《謝辞》

研究する機会を与えていただいた名古屋市教育委員会をはじめ、勤務校である名古屋市立神宮寺小学校の職員のみならず、本当に感謝しています。また、愛知教育大学准教授の青山和裕先生には、私の研究に多くの時間を費やしていただき、適切なお助言をいくつもいただきました。本当にありがとうございました。

《引用・参考文献》

- OECD(2020) : OECD Future of Education and Skills 2030 project.
- OECD(2020) : Student Agency for 2030仮訳.
- 日本財団(2019) : 18歳意識調査(第20回テーマ : 「国や社会に対する意義」).
- 白井俊(2020) : OECD Education2030 プロジェクトが描く教育の未来 エージェンシー、資質・能力とカリキュラム、ミネルヴァ書房.
- 内閣府 : Society5.0.
- 文部科学省(2017) : 小学校学習指導要領(平成29年告示)解説 算数編.
- 算数・数学ワーキンググループ(2016) : 「小・中・高等学校を通じた統計教育のイメージ」.
- 文部科学省(2019) : GIGAスクール構想の実現. 統計分析ソフト「eStat」 <http://www.estat.me/estat/eStat/>.
- 文部科学省(2017) : 小学校学習指導要領(平成29年告示)解説 総則編.
- 日本学術会議(2020) : 新学習指導要領下での算数・数学教育の円滑な実施に向けた急提言 : 統計教育の実効性の向上に焦点を当てて.

データの特徴を適切に表現する力を育む学習指導

～小学校第1学年「わかりやすくせいりしよう」の実践を通して～

白石 円 岩手大学教育学部附属小学校

岩手県盛岡市加賀野 2-6-1 siraisi@iwate-u.ac.jp

1 研究の目的

子供たちは学校教育において、棒グラフや円グラフ、ヒストグラムや箱ひげ図など、データの分類整理の仕方が多様にあることを学習する。そして、新しい統計グラフに出会う度に、整理の仕方が異なるとデータの見え方が変わることを体感していく。青山（2011）は「活用・実践力」重視の統計教育の普及を喫緊の課題としつつ、「表現力育成」も平行した指導の必要性を述べている。こうしたデータの見え方が変わる体験をもとに、目的に合った統計グラフの存在と意図した選択について学ぶことが大切である。

そのため、統計教育の入門期である小学校第1学年の時期に、同じデータでも整理の仕方を工夫することでデータの見え方が異なること、特徴が捉えやすくなることについて実感を伴って理解させたいと考えた。よって、本研究の目的を「データの特徴を適切に表現する力を育む学習指導」と設定した。

2 研究の手立て

本単元「わかりやすくせいりしよう」は、全2時間とし、単元を通じて魚釣りゲームの結果の表し方について考えていく。学習活動は以下の通りである。

第1時：ものの個数を種類ごとに整理し、分かりやすく表す方法を考える。

第2時：データを整理した簡単な絵や図を見て、個数の大小を比べたり、データの個数の特徴を読み取ったりする。

本時は第1時、魚釣りゲームの結果を見て、種類ごとの数の多さを表現するために整理の仕方を考える場面である。並べ方を改善していく中で【図1】、どの生き物がどの程度多いのかというデータの特徴を表現する際、高さを揃え均等に配置することや、絵の大きさを揃えることのよさを実感できる学びを目指した。



【図1】児童が並び変えの表現における変容

栢元（2009）は、統計的な見方・考え方として、ものの個数を集合ごと集めるなどの観点を示し、実践を行っている。その際、使用した絵カードに注目してみると、同じ種類の絵は大きさがほぼ同じものであった。本研究では、あえて大きさが不揃いの絵カードを使用する。これにより、絵カードの並べる際に、高さを揃えて配置したり、絵自体の大きさを揃えたりする必要があることを子供自身が見出す展開をねらった。

以上のことから、4つの手立てを設定した。

- (1) 「縦に並べるだけでは数の多さが把握できない」という問いを生むために、それぞれの生き物の高さを不均等に設定する。
- (2) 「困った」「こう並べたい」という思いを顕在化させるために、考えの根拠を問い返す。
- (3) 操作することを通して考えを表出できるように板書と同一の教材を準備する。
- (4) 初めの並べ方と、高さを揃えた並べ方を比較しそのよさを価値付ける。

3 実施した授業

授業実施対象：国立大学法人附属小学校 第1学年

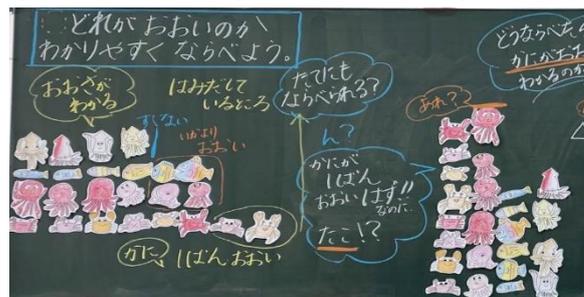
授業実施日：令和3年12月23日 1時間

授業者：白石 円

4 授業実践と考察

【手立て（1）（2）について】

子供たちは長さの学習をもとに絵を横に並べ、カニが一番多いことを結論付けた。その後、同じように縦でも並べることができるか考え、並び変えた【図2】。



【図2】児童が並び変えた絵

並べた結果を見て、「カニの数が一番多いはずなのに、なぜタコの数が多くなったのか。」という問いが生まれ

た。初めに“カニが多い”が確定できたことで、縦に並べた時に「おかしい」とズレを感じ取り、「どのように並べたらカニが多いが伝わるの」と問いが生まれたのだろう。この問いをきっかけに、“カニが多い”というデータの特徴を表現するための並べ方を追究し始めた。また、学習の振り返りにも以下のような記述【図3】がみられた。

A 児	縦だと全然最初は上手いかわなくて、「ん。」ってなったけど、縦を横にしたら、高さが揃ったし、こんな数え方知らなかったけど、今知れてよかったです。
B 児	縦でもできるんじゃないってやってみたら、カニが一番多かったのにタコが一番多く見えて、ん？ってなった。

【図3】学習の振り返り①

この子供たちの姿から手立て(1)(2)は有効に働いたと考える。

【手立て(3)(4)について】

その後、「どのように並べたらカニが多いが分かるのか」という問いの解決を始めた。各班に黒板と同一の生き物の絵を配付し並べ方を検討する活動を設定することで、子供たちは絵を動かし、試行錯誤しながら並べ方を考えていた【図4】。班の4人で考えを伝え合いながら、データの特徴を表現するためには、絵同士の隙間をあけて高さを揃えればよいことに気づき始めた。



【図4】児童による考察の様子

各グループをまわりながら、教師が「どうしてこう並べたの?」「それって、この絵のどこのことなの?」と、並べた背景にある子供の見方や考え方を表出させる声かけを行った。このことにより、「横を揃えて並べると見やすくなるんだよ。」「高さがバラバラだから、隙間をあけて、横が揃うようにするにするといい。」と、自分たちの並べ方の工夫点やよさをアウトプットし、自覚できたのだと考える。

また、「ペアをつくといい。」というまとまりを意識する子供の素朴な考えを全体で共有した。「どこがペアなの?」と問い返すことで、横に並ぶ4つのまとまりを「ペア」と表現していることを確認した。そして、「どうして4つのまとまりをつくらうと思ったの?」と着想を問うことで、「同じ数にあたる場所をそれぞれ揃えて並べることで見やすくなった。」という高さに着目して並べるよさ表出させ、全体で共有することができた。

解決を終えた子供たちの振り返りが以下【図5】

である。

C 児	初め縦はできないよねとっていたけど、同じ高さでやるとカニが2個余るといいと思いました。
D 児	カニが一番多かったのに、タコが一番多く見えて、「ん?」となって3班の人たちが4つのペアを作って、カニが多いって分かりやすくして、Fさんは、タコは横にして魚を縦にして分かりやすくしていいなっていました。

【図5】学習の振り返り②

振り返りの内容からも、手立て(3)(4)は有効に働いたといえるだろう。

一方、実践上の課題もあった。それは、絵グラフの外形にも目を向けさせるべきであったことである。「数えなくても、高さを見るだけで多い少ないが分かる」という子供が無意識に感じていることを表出させ、外形に着目していることを価値づける必要があった。高さを揃えて並べなおし比較した際に、上のかに2匹に対して「はみ出しているのこりもの」と発言した子供の考えを取り上げることができれば、外形によって数の多さの判断ができるというグラフに表すよさを実感させることができたと考える。

5 研究のまとめ

本研究を通して、小学校第1学年におけるデータの特徴を適切に表現する力を育む指導の手立てについて検討し、有効性について検証することができた。そして、子供たち自身が本質に迫る学びをつくっていくためには、学びの柱となる「問い」の質が重要であること、その「問い」を子供自身が解決したいと強く思っていること必要だということも明らかになった。

また、学習の振り返りには、以下のような記述【図5】もみられた。

E 児	1個1個足やヒレを書いているとめんどくさいので、カニは□、タコは○、魚は●、イカは△で書くことが書きやすいなと思いました。
-----	---

【図5】学習の振り返り③

絵の大きさを揃えた整理の仕方を考える中で、絵を○などの記号に置き換えて表現するという第2学年の簡単なグラフの学びにつながる子供の姿も見られた。このように、自分で解決を振り返りながら、データの特徴をよりよく表現するために追究し続ける子供の育成を今後も目指していく。

【引用・参考文献】

表現力育成に向けた統計教育の展開について

青山 和裕 愛知教育大学数学教育講座

日本科学教育学会年会論文集 Vol.35 (2011) p93-94

小学校第1学年における児童の統計的な見方・考え方

裕元 新一郎 静岡大学教育学部

日本科学教育学会年会論文集 Vol.33 (2009) p141-144

統計的問題解決力の育成を目指した指導の在り方～標本調査の実践を通して～

稲垣道子・岩手大学教育学部附属中学校

岩手県盛岡市加賀野 3 - 9 - 1

電話 019 - 623 - 4241 FAX 019 - 623 - 4243

E-mail inaka@iwate-u.ac.jp

1 はじめに

まもなく日本には、「Society 5.0=超スマート社会」が到来し、社会の在り方が劇的に変わると言われている。また、新型コロナウイルスの感染拡大など先行き不透明な予測困難な時代に突入している。複雑で変化の激しい社会において、人間には経験的に判断することだけでなく、出会った未知の状況に対してよく考え根拠や論拠を基にした判断や意思決定をすることが求められている。

そのような社会を生き抜くために、中学校学習指導要領（平成 29 年告示）解説数学編では、統計的な内容等の改善・充実を強調している。今後の人間社会において、社会生活など様々な場面で、必要なデータを収集して分析し、その傾向を踏まえて課題を解決したり、意思決定したりする力が求められている。また、渡辺（2014）は、国内外の統計教育の情勢から、統計的問題解決力育成に向けた教育の重要性について言及している。さらには、小学校からの学習の系統化についても指摘している。

本研究では、変化の激しい時代において、生徒が自ら問題を発見し、データを収集して分析し、問題を解決するような統計的問題解決力を育成する授業や指導計画の在り方について、3年標本調査の実践を通して明らかにしていく。

2 研究方法と内容

(1) 主体的に学びを実現する単元計画の作成

単元の最後に、生徒自身が統計的問題解決力の高まりを自覚することを目指す。そのために、生徒自身で PPDAC サイクル（図 1）を活用できるよう、標本調査への確かな理解を促したい。標本調査について、生徒が必要感を持ちながら主体的に学び続けることができるような計画を作成する必要がある。そ

こで、単元の導入で見いだした問いを、単元を通して解決することができるよう計画した。

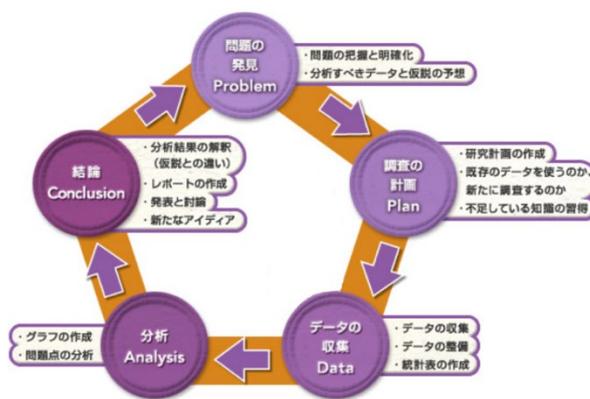


図 1 PPDAC サイクル（総務省統計局）

下の題材（図 2）は、単元の導入で扱ったもので、実際に生徒に提示した。

1936 年のアメリカ大統領選挙には、共和党からは
ランドンが、民主党からはルーズベルトが立候補し
た。その事前予想は

①リテラリーダイジェスト（雑誌）
当選予想：ランドン（調査数 237 万 6523）

②ジョージ・ギャラップ社（世論調査会社）
当選予想：ルーズベルト（調査数 5 万）

当時は、調査規模があまりにも違うため、どちらの予想が正しいのか、勝負にならないと思われていたが、
選挙の結果はルーズベルトの圧勝だった。なぜこのようなことが起きたのだろうか。

図 2 単元の導入で扱う題材

生徒が、標本調査の結果が実際の結果と異なった理由を考えることで、無作為抽出や標本の大きさなど本単元で理解すべき概念に着目し、単元の問いを見いだすことを期待した。

また、本単元では、生徒にとって身近なデータだけでなく、社会一般的なデータまで幅広く扱うこ

ととする。身近なデータを扱うことで実感を伴った理解を促し、社会一般的なデータを扱うことで、社会生活で活用することへの意識を高めたいと考えたからである。その中で、相対度数を用いて考察する場面を設け、学んだことが生かされていることを実感し、高校の学習につながる問いの生成につなげていきたい。単元の最後には、生徒自身がPPDACサイクルを意識した調査活動を行うことで、統計的問題解決力を高めるとともに、実感を伴っての理解を促したいと考えた。

次の図3は、実施した単元計画である。

時	学習内容	題材
1	単元の問いの生成 全数調査・標本調査	アメリカ大統領選挙 身のまわりの調査
2	無作為抽出	附中生がお正月に食べるもの
3	標本の大きさによる違いの考察※高校への接続	選挙の当確情報
4	母集団の傾向を推測する	工場の不良品検査 魚の捕獲再捕獲法
5	世の中の調査方法や結果を批判的に考察する	お客様満足度調査・TVの卒業ソングランキング
6.7	実際に標本調査を行う	自分たちが調べたいこと
8	交流・振り返り	

図3 単元計画

(2) 標本調査の概念形成にかかる授業開発

① 標本比率の分布に関連づけた母比率の推定

渡辺(2014)は、標本調査の概念形成について、不確実性を伴いながらの意思決定やリスク等に対する感覚を養うことの必要性を指摘している。標本比率の分布に関連付けた母比率の推定は、学習指導要領でも経験的に理解させるよう記述されている。

そこで、授業の題材として「国政における選挙の当確情報」を扱うこととした。選挙では、全て開票されなくても当選確実といった当確情報が出されることがある。これは、全ての投票数を母集団、既に開票した票数もしくは出口調査の回答数を標本として標本調査を行い、母集団の傾向を推定した結果を用いて、選挙結果において逆転することはないと判断している。選挙の投票日には、どのテレビ局においても選挙に関する報道番組を

放映しており、生徒にとってはなじみのあるものである。また、テレビ番組の開始と同時に当確が出たり、開票率がわずか10%で当確が出たりする候補者もあり、生徒が調査への疑問を持ちやすい題材であると考ええる。開票率10%で、本当に逆転は起こりえないのかという問いから、自然と標本比率の分布に着目し、母比率との関係を考察するのではないかと考えた。しかし、本来の選挙では、事前調査の情報が加味されていたり、1つの選挙区から複数の候補者が立候補していたりする。それらの考察については難易度が高いため、事象をモデルとして単純化する。母集団を1000人、立候補者はAとBの2人、無効票はないこととし、単純なモデルとして標本比率と母比率の関係に着目して考察することとした。その中で、結果の推定について触れる。

授業展開については下のとおりである(図4)。

	学習活動
導	1 選挙の開票速報で当確情報が出ている場面を見る 2 開票率が10%で当確情報が出ることに着目する
展	3 AとBの得票率が4:1(80%:20%)の母集団から実際に10%(100個)データを取り出した時のAとBの得票率がどうなるか実験する 4 開票率(標本の大きさ)を5%,1%に変えても、母集団の傾向を推測できるか実験する 5 AとBの得票率を5:4(56%:44%)に変えたら、母集団の傾向を推測するために何%開票すれば信頼できるか考える
終	6 本時の学習を振り返る 7 高校数学へのつながりを知る

図4 授業展開について

② テクノロジーの活用によるシミュレーション

選挙の当確は様々な情報をもとに多角的に判断しているため、実際の条件で調査をしたり、実験したりすることは難しい。しかし、状況を単純化することで、生徒自身でテクノロジーを活用しながらのシミュレーション活動を可能とした。使用するソフトはMicrosoft Excelである。AとBの得票率が設定してある1000個のデータの中から、標本数を自分で設定し、無作為に抽出できるようファイルを配付する。

データ	標本	10% 投票数	得票率	当選予想
データ 1	A	Aへの投票	88	0.88 ○
データ 2	A	Bへの投票	12	0.12 ×
データ 3	A	総投票数	100	
データ 4	A	開票率	10%	
データ 5	A		100人抽出	
データ 6	B			
データ 7	A			
データ 8	A			
データ 9	B			
データ 10	A			
データ 11	A			
データ 12	A			
データ 13	A			

図5 配付したファイル

そして、生徒が100個、50個、10個の票を標本として無作為に抽出し、AとBの得票率（相対度数）を調べ、標本の分布の様子を観察し、考察する場面を設定した。無作為抽出のシミュレーションを複数回行い、その結果を箱ひげ図で表すことで標本の大きさと標本の分布についての変動を捉える活動を設定した。

③学習の振り返りにみる生徒の概念形成の様相

学習の前後で生徒の認知や思考がどのように変容したかわかるように、単元の最初に考えた問いについて最後にもう一度考え、本単元の学びの価値について振り返る場面を設定した。それにより、生徒が数学的に考える資質・能力の高まりについてメタ認知的に実感できるとともに、教師側も指導の成果を検証することができる。振り返りの場面では、ロイロノート（教育支援アプリ）を使用し、スタディ・ログとして記録の蓄積を行った。

3 授業による検証

授業は、2022年1月に国立大学附属中学校において、3年生4学級（1学級35人）を対象に行った。

(1)単元における統計的な問いの生成【第1時】

まず、アメリカ大統領選挙の題材に関する生徒の問いについて、記述の内訳を図6に示した。

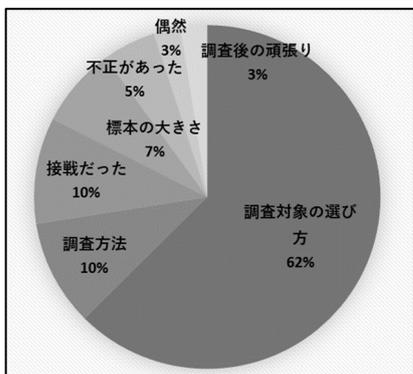


図6 生徒の記述1

本単元での学びの重要となる標本の抽出方法や、標本の大きさに着目した生徒が多かった。これら

を問いとし、単元を通して解決することとした。

(2)無作為抽出の必要性とその意味の獲得【第2時】

無作為抽出の意味について、教師が一方的に教えるのではなく、標本を抽出するときの留意点を考えるを通して、概念的に理解することを重視し、生徒が見いだすことを大切に授業を行った。

①問題把握・問題解決

「班新聞の記事で、附中生（全校生徒420人）がお正月に食べるものについて書きたい。時間がないので、アンケートは標本調査で行う。どのようにアンケートをするか。」と問い、調査人数や方法等すべて考えさせた。生徒の考えは以下の通り。

- ・クラスや学年関係なく35人
- ・各クラス代表者1名全12人
- ・1年B組、2年B組、3年B組105人
- ・各クラス全クラスから男女2名ずつ24人
- ・1年A組2年B組3年C組105人
- ・各クラスの窓際4人全48人
- ・全校からくじびき
- ・各クラス同じ男女比1：1で選ぶ。

②共有・定義

「全体を均等に」「バランスよく、偏りなく」「少ないケースの人を代表にはいけない」などの意見が出された。具体的には、座席の指定（当時の座席は窓際が男子）やクラスや学年によっては偏りがでるということである。そのように標本を抽出する必要性について、「全校の意見を知るため」「散らばっていないといけない」という意見が出され、無作為抽出の必要性を捉えていた。その後、無作為抽出の定義を確認した。

③検証・新たな問いへ

グループごとに決めた方法で標本調査を行い、事前に行っていた全数調査の結果と比較した（図7）。

総度数が異なることから、割合で比較する方がよいことや、割合を用いるのであれば円グラフを用いるとよいといった考えは生徒から出された。これまでの学習を生かして、データを整理しようとしている姿が見られた。2つの調査の比較が似たような傾向になったことに驚く生徒もいた。

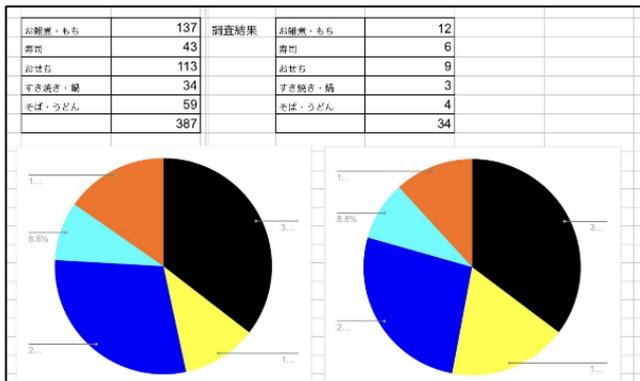


図7 全数調査（左）と標本調査（右）の結果

各学年学級の調査結果には特色があり、それらも提示すれば、より無作為抽出の重要性が認識できたと思うが、時間の都合上、本実践では提示できなかった。授業の最後に、「何人選べば妥当なのか」という疑問が出た学級もあり、次時の問題解決につながった。

(3) 標本調査による母集団の推定【第3時】

選挙の当確情報を題材とした授業である。

①問題把握

開票率10%の当確情報に対して、生徒からは「残りの90%があるのになぜ分かるのか」「逆転するかもしれない」「10%の中の割合を考えたら確実なのでは?」という意見が出された。その後、圧勝や接戦などの状況によっても変わるのではないかという考えを共有した。

②得票率4:1の母集団から、10%の標本抽出（全体での活動）

事象を単純化して実験することを確認した。生徒には、得票率が母集団において4:1であることは伝え、それを圧勝のケースとして、開票率10%の標本で1度実験した。後半の活動で、生徒が実験を進めることができるように、操作方法を確認し、全体で実験と考察を行いながら進めた。1度実験した結果である得票率0.77を見て、生徒は驚きの声を上げた。生徒は、10%での当確情報を懐疑的に見ていたことが伺える。その後、「10%で母集団の傾向を推測できるか」と問うと、10%で推定できるという意見した中には、「もう1回やってほしい。偶然かもしれないから」があった。これは、無作為抽出の不確実性を捉え、多数回試行の必要性を感じていることが伺える。この意見から、複数回試行することにした。

	10%	5%	1%
1	0.77	0.86	0.9
2	0.84	0.84	0.9
3	0.87	0.82	0.7
4	0.78	0.81	0.7
5	0.81	0.78	0.8
6	0.71	0.72	0.7
7	0.71	0.8	0.9
8	0.82	0.72	0.5
9	0.76	0.82	0.5
10	0.79	0.78	0.8
11	0.79	0.8	0.9
12	0.85	0.88	1
13	0.78	0.78	0.8
14	0.85	0.66	0.8
15	0.82	0.92	0.8
16	0.76	0.8	0.8
17	0.85	0.78	0.8
18	0.8	0.72	1
19	0.8	0.74	0.3
20	0.8	0.9	0.9
	0.798	0.7965	0.775

図8 実験結果（一番下が平均値）

20回実験をしてデータを集め（図8）、「10%で母集団の傾向を推測できるか」と問うと、生徒は平均値で比べようとした。得票率の平均値を求めると0.798となり、改めて、生徒は10%で推定が可能と判断をした。

③得票率4:1の母集団から、5%、1%の抽出による標本調査（全体での活動）

5%、1%と標本の大きさを変え、同様の実験を行った。得票率の平均値が四捨五入で0.8になるが、値にばらつきが出てしまうことに気づく生徒がいた。これは、標本の大きさに対する標本変動を捉えている表れである。散らばり具合を見るために、グラフを用いればよいことを確認し、下の箱ひげ図（図9）をつくった。

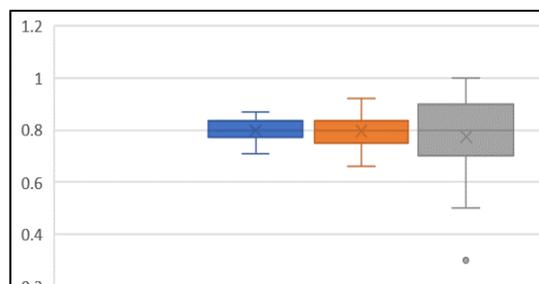


図9 箱ひげ図1（左から10%、5%、1%）

このことから、生徒は、「標本の大きさが小さくなるほど、範囲や四分位範囲が大きくなる。」「中央値や平均値は同じくらいでも範囲が異なる。」などの考察をしていた。その後、5%だと推定できそうだが1%での推定は不安定であると判断した。

④得票率5:4の母集団から標本抽出

グループごとに自由に開票率を変えて調べさせた。平均値だけではなく、範囲や四分位範囲に着目しようとする生徒が多かった。これは、標本の不確実性を理解した上で考察しようとしていること表れである。区間による推定

の必要性を感じていることが伺える。

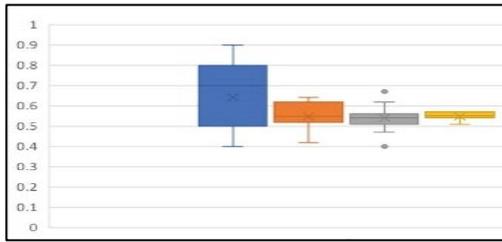


図 10 生徒が作成した箱ひげ図①

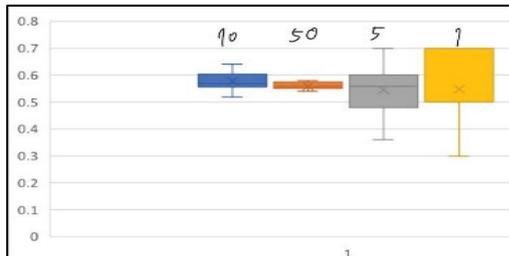


図 11 生徒が作成した箱ひげ図②

また、次の図 12 のように考えたグループもあった。このグループは、グラフは作成せず、多数回試行を行っていた。確率の考え方に基づいた実験である。開票速報は正確性が求められる。不確定な事象から正確に予測する必要性を感じていることが伺える。本時のねらいとはズレが生じてしまったが、この単元で捉えたいことについては考察していたといえる。

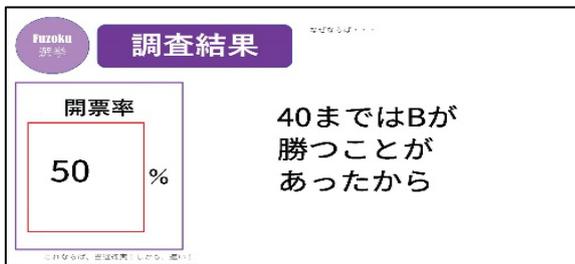


図 12 生徒が提出したシート

時間がなく、考察までたどり着かなかったグループもあった。あるグループが作成した箱ひげ図(図 13)を全体で共有した。時間があれば、20%と1%の間も調べたかったと話していた。

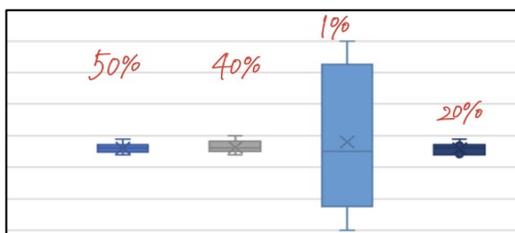


図 13 全体で共有した箱ひげ図

⑤高校数学との接続

授業の最後に、「実際にやるときに(報道で)、得票率の最大と最小がどのくらいの範囲であれば

妥当とってよいのだろうか」ということを疑問視する生徒がみられた。これは、高校で学習する区間推定の学習につながる問いである。高校で学習することを伝え、本時を終了した。本時の授業を終えた生徒の振り返りの記述から、得票率の予測の精度が標本の大きさによることを理解していたことが伺える。さらに、下のような記述も見られ、高校の学習につながるような問いを持った生徒がいた。

- ・傾向を推測するとき、AとBの比がどれくらいなら、誤差と考えて母集団の傾向を推測してよいのか難しかった。特に開票率が低い時、どのようにして判断すればいいか分からなかった。
- ・標本の大きさを全体の何%にしたら良いのかは勝手に決めて良いのか、それとも良いところがあるのかをもっとちゃんと考えていきたい。

(4)PPDAC サイクルでの調査活動【第6,7時】

単元の最後に、自分たちの疑問や問いを解決する活動を行った。次の図 14, 図 15, 図 17 は生徒が作成したレポートである。

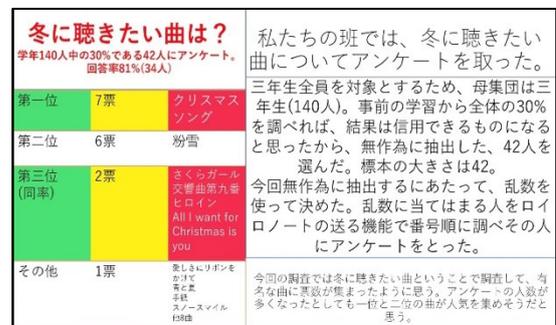


図 14 生徒が作成したレポート①

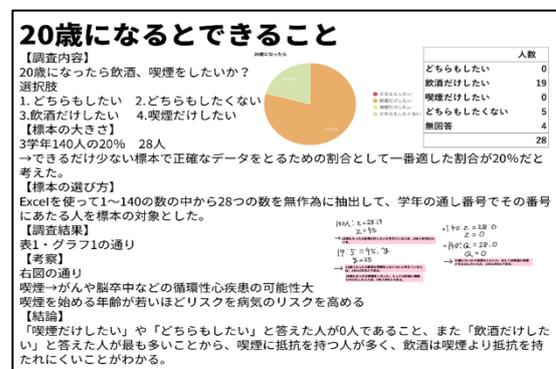


図 15 生徒が作成したレポート②

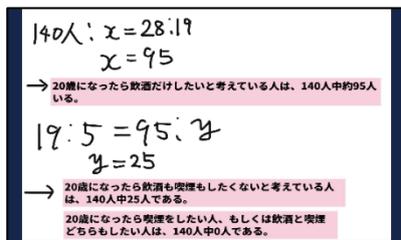


図 16 生徒の作成したレポート② (拡大)

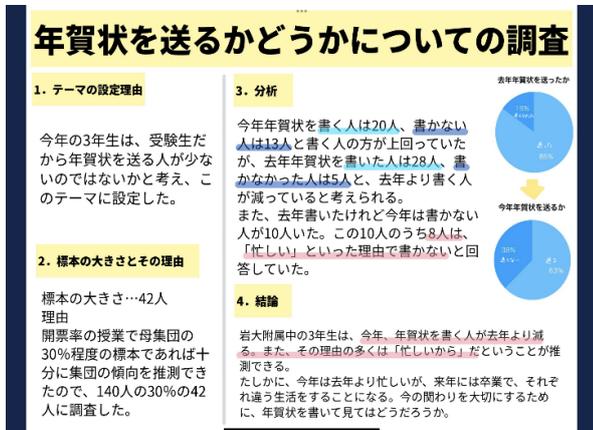


図 17 生徒の作成したレポート③

図 15 のグループは、標本調査の結果をもとにして、比例式を用いて処理し、学年全体の傾向を推測している。さらにその結果を用いて考えを主張している。また、標本の大きさを決める際、当確情報の授業時は母集団が 1000 人だったが、3 学年は 140 人で少ないことから、選挙の時より標本の大きさを大きくしたグループもあった。現段階で妥当な標本の大きさは判断することはできないが、その重要性を捉えていることが伺える。このように、標本調査を推測するための統計手法として捉え、結論を導き出しているグループがある一方、推測するという意識で調査結果をまとめようとしたか判断がつかないグループも多数あった。最初は推測統計の意識で調査を始めているが、最終的にその意識が薄れ、標本の分析結果をそのまま結論にしている可能性が伺える。小学校算数での記述統計から推測統計への切り替わりを、より意図的に行う必要があることが浮き彫りになった。

(5) 調査結果の交流【第 8 時】

8 時間目は、調査結果の発表・質疑応答・調査活動の振り返りを行った。

<抽出方法について>

- 0. 先着順は無作為なのか？早く回答できる人は時間に余裕がある人だと思う。
- 0. アンケートを取る時に、一度失敗して、このクラスの人に回答を呼び掛けていた。それを聞いていた

人が多から、標本は B 組の人の意見が多いのではないか。全体というよりは、B 組の人の意見が濃い結果だと思う。

<標本について>

Q. 以前テレビで同じアンケートを取った番組を見たが、テレビと結論が逆になっていた。どのような点はその違いを生んだと思うか。

A. テレビでは、いろいろな年代に聞いていると思うが、僕たちは中学生に聞いた。中学生はこのようなことを考えているということが分かった。

<調査方法と結論について>

0. テストで一番点数が悪かった教科を質問していた。0 票だった英語を 3 学年は英語が得意と結論づけていたが、1 番ではなかっただけで、2 番目に点数が悪かったかもしれない。

このように、この単元で学習してきたことに着目して調査方法や結果を批判的に考察する生徒がおり、全体で共有できた。また、最後に自分たちの調査結果を振り返った。視点は①他グループから学んだ視点②次に同じように標本調査をしたらどこを改善するか。「調査方法」「調査結果」「調査分析」の視点で振り返る。

①他グループから学んだ視点

標本の大きさを授業での経験をもとに設定している点は共通していたが、標本の選び方（性別ごとに分けて選ぶ・クラスごとに選ぶ）や選択肢の数を根拠に標本を設定しているグループがあり、なるほどなと思った。質問の内容によって標本の選び方（性別ごとに分けて選ぶ・クラスごとに選ぶ）を変更している点を参考にしたい。分析の結果を母集団の人数に戻して実際の標本調査が母集団の傾向を掴むのに役立っていることが実感できた。

次に同じように標本調査をしたらどこを改善しますか。「調査方法」「調査結果」「結果の分析」の視点で振り返ってみましょう。

- ・調査内容に合わせて標本の数を変える（質問の選択肢の数、母集団の規模などによって）
- ・結果を標本の範囲内で完結させず、母集団の数に還元して結論付ける
- ・性別や年代に着目して結果の内訳を見て判断をする
- ・どのグラフを使うべきか判断する

図 18 生徒 A の振り返りシート

上の生徒 A は、シートを作成した段階では、記述統計の意識で問題解決を終えている。しかし、他グループの発表を聞いたことで、母集団の傾向を推測するという目的を想起し、学びを調整することができた。また、より正確な結果を得るために、状況に応じて標本の抽出方法や標本の大きさを変える必要性を認識し、新たな視点を得ている。

(6) 単元を終えて

単元の最後に、「もしあなたが、リテラシーダイ

ジェスト社の記者なら、どのように対象者を選びますか？」という問いに対する考えを記述させた。それにより、これまでの単元の学習で得たことが表出されると考えた。

生徒の振り返りシート（図 19）をみると、単元の最初から対象者の抽出方法に着目していたが、単元の最後には、記述がより具体的なものに変容している。また、単元を通して標本の抽出方法や推測するときの留意点に着目しながら学習していることが伺える。自身の調査では、先着順での抽出方法を行っていたが、単元の最後には、住民に番号を付与し、表計算ソフトでの無作為抽出の方法を選択している。単元の中で学びを調整し、自己の変容をメタ的に認知できたのではないかと考えられる。

4 考察

(1) 生徒の記述をもとにした分析

標本調査を学んでよかったと思うかという問いに対しての回答は右の通りであった（図 20, 図 21）。

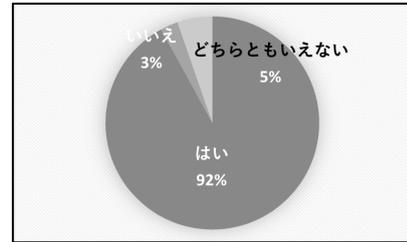


図 20 生徒の回答 1

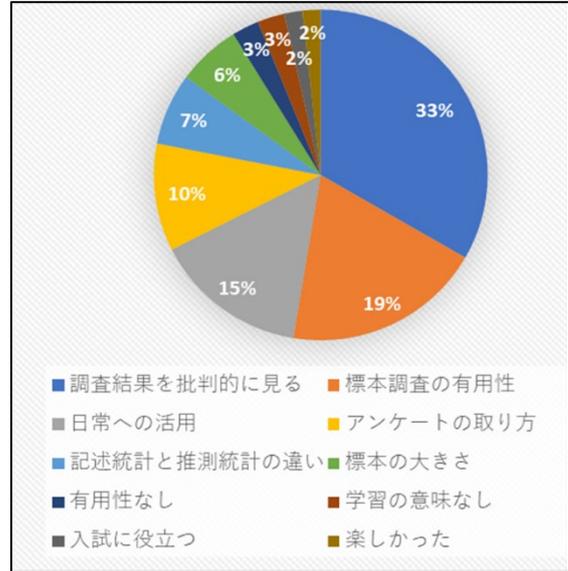


図 21 生徒の記述内訳

8章 標本調査振り返りシート 単元のテーマ：集団の傾向を推測しよう！

単元の学習の前に

1936年のアメリカ大統領選挙には、共和党からはランドンが、民主党からはルーズベルトが立候補した。その事前予想は

①リテラリーダイジェスト（雑誌） ランドン (調査数 237万6523)

②ジョージ・ギャラップ社（世論調査会社） ルーズベルト (調査数 5万)

当時は、調査規模があまりにもちがうため、どちらの予測が正しいのか、勝負にならないと思われていたが、選挙の結果はルーズベルトの圧勝だった。なぜこのようなことが起きたのだろうか。

1節 標本調査 P210～P219

全数調査と標本調査

全数調査

調査の対象となる集団全部について調査すること。答えるか答えないかは別として対象となるものに答える権利を与えること。

- 手間・時間・費用がかかる
- 公正
- 正確

• 国勢調査

• 学力調査

• いじめ調査

• 身体測定

• 体力測定

標本調査

対象となる集団全部ではなく、ある一部分を調べて集団全体の傾向を推測すること

- 調査対象が少なくて済む
- 調査に時間がかからない
- 調査に費用がかからない
- 調査結果が正確に推測できる

• YouTubeのアンケート

• 出口調査

• 製品の品質検査

• 世論調査

• テレビの視聴率の調査

実際に標本調査をやってみよう！

母集団の傾向を推定する方法

母集団の傾向を調べるには比の考えを利用することができる。それにはまず母集団から無作為に標本を取り出してそこからデータを得る。その標本とデータで表して、母集団とその母集団で得られると思われる標本をxと表した時を比で表して、その二つの比を=で等式にしてxについて解くと求められる。

もし自分がリテラリーダイジェストの記者なら、どのように事前調査を行いますか？そのように考えた理由も書いてみよう

<事前調査>

アメリカの州ごとに、そこそこの割合で無作為に標本を抽出（番号を振り分けた後、Excel等で無作為に選ぶ）。その州の代表者の意見が州全体の意見として考える。それを一つ一つの州で行いそこからランドンとルーズベルトのどちらの意見の方が多いかを求める。

<理由>

このようにすることで、地域、貧富、男女、人間関係なく抽出することができるから。また、アメリカの州それぞれで一般人が選挙人を選んで、その選挙人が直接投票を行い、州ごとにどちらの意見が多いかで大統領が決まるという方式に似ているため、できるだけ正確に調査することができると思ったから。

図 19 生徒の振り返りシート（単元後）

自由記述として、主に次の記述があった。

- 以前のアンケートなどでも結構 3-A でのアンケート

ト結果が全校も同じようになるだろうと仮定して話を展開していたりして偏りがあったと思う。次もしも自分達で新聞などのアンケートを作るときに、3-Aでのアンケート結果が果たして全校の生徒たちの意見と大体合致しているのかや、アンケートの内容やクラスの状態によって偏りが生まれてしまっていないかを確認したい。それを踏まえてアンケートの仕方や送る対象の状態をもっと吟味すべきだと思った。

- ・無作為に抽出することで都合よくバラけることが不思議だと思った。一部を調査しただけでそれを全体の結果と見做していいのか疑問だったが、選挙の開票率を変えて調査した授業で、10%ほどの開票率でもほとんど正しい結果を得られたことに驚いた。
- ・メディアで紹介されていたアンケートや企業が出したお客様満足度など、本当に正しい値なのか、偏りがあるのか、批判的に捉える観点として学ぶことができた。
- ・学校新聞のアンケートで無理に全校にアンケートをしなくても20%前後(84人)の生徒数で傾向をつかめるということがわかり、効率化や労力の削減につながると思った。
- ・総合的な学習の時間でたくさんの標本調査、全数調査を行ったが、その前に授業を行えたら、もっと信頼性のあるデータをとれたなと思った。
- ・アンケートをとるときの手段が1番悩むし、1番大変だと思った。全数調査はただ全員にアンケートをとるが、標本調査は、無作為に抽出する必要があるが、無作為の中でも男女比を同じにしたり、クラスの数人の比も同じにしたり、など調整も大変だと思った。本当にその抽出の仕方は無作為に抽出しているのかとか、標本調査をした結果から全体の傾向を読み取るためには、標本に偏りが出たはいけないので正しい結果が出やすいアンケートを作るのは大変だと思った。
- ・自分たちで調査をしてみることで、どういうところに気をつけなければいけないのかわかったので、そういう経験をするのは大切だと思いました。ほとんどの生徒がこのように記述しており、標本調査と全数調査の違いや仕組み、有用性について実感もちながら理解していることが伺える。

しかし、次のような記述がみられた。

- ・そもそも標本調査を使ったものをあまり見ないので日常と結びつけて考えることができなかつたです。(生徒C)
 - ・特に自分の生活に関わってくる訳ではないから。(生徒D)
 - ・データの的には、母集団の20%で、全体の傾向を予測できることはわかったが、やはり、全数調査をした方がいいんじゃないかと思った。(生徒E)
- 生徒C、Dは、単純に標本調査の理解が不足しており、世の中で行われる調査と、標本調査の必要性を関連付けて考えることができなかつたと考えられる。また、社会の事象を数学的に捉えることに関心がない場合もある。この場合は、学びに向かう力、人間性等を育むための手立てを考える必要がある。生徒Eは、標本調査の有用性に対して懐疑的にみているといえる。全数調査が難しい場合があることを理解させる必要があると共に、高校数学の学習をつなげられるよう、涵養的に有用性を実感させたい。

(2) 教師側の分析

①単元計画について

第1時に標本調査の失敗例を取り上げるといった工夫により、生徒は問いを持ち、単元を通して探究的に考察する生徒が多く見られた。

第1時では、調査の対象者に着目した生徒が多かったため、第2時で抽出方法に着目し、無作為に抽出する必要性やその方法について批判的に考察することができた。このとき、標本の大小関係に疑問を持ち、第3時への問いの生成を確認できた。第6時以降の標本調査を行う活動では、標本の大きさを決める際に、第3時の授業で得た事項を視点として議論していた。

第6、7時は、生徒が標本調査のよさを実感する時間となった。生徒の調査計画の時間を大切にしたが、そのことで生徒が統計的な問題解決を力強く進めることができたと考えられる。

②標本調査における概念形成

標本比率の分布に関連づけて母比率を推定することに関しては、生徒は標本比率の変動について、統計的な確率の考えに関連付けながら考察していた。その際、標本の分布に着目する必要性に気づき、標本の大きさと標本比率の分布

の関係と母比率について確認していた。また、母比率が変わることで、標本の大きさも変えて考える必要があることについて、直観的に捉えたと考えられる。範囲や四分位範囲、平均値や中央値等を用いた考察も確認できた。生徒が高校で学習する際に、本実践の内容を想起することを期待したい。

③学習の振り返りにみる生徒の概念形成

単元の最後には、「もしあなたが、リテラシーダイジェスト社の記者なら、どのように対象者を選びますか?」と問うた。図 22、図 23 の生徒は、それぞれ単元の学習後には、標本調査の意味や無作為抽出の大切さを捉えている。

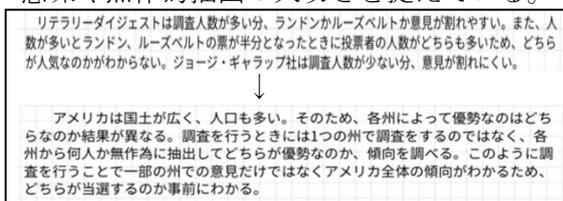


図 22 生徒の単元前後の記述 1

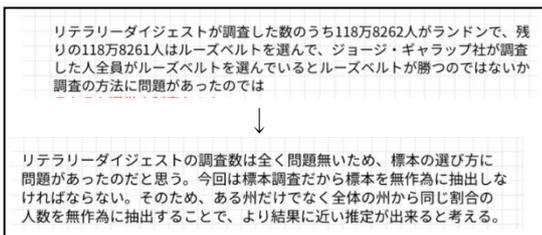


図 23 生徒の単元前後の記述 2

しかし、標本を無作為抽出する場面で、実際には無作為に抽出されているとはいえない方法を正しいと捉えていた生徒もいた。

知識を活用し、実際に問題を解決しようとする際には難しさがある。また、指導者にも確かな知識や確かな理解が求められる。概念形成へのさらなる手立てが必要であることがわかった。

5 本研究の成果と課題

(1) 成果

1つ目として、生徒が PPDAC サイクルを遂行して統計的問題解決を行う姿を確認できたことである。これまで、PPDAC サイクルを活用した統計的問題解決を行ってきたが、標本調査でも同じように解決した。単元の終盤での活動では、これまでの学びを生かして調査計画を立てる姿は、目指したものに近かった。また、調査結果の交流の場面を設定したことで、自らの学びを調整する姿や、標本調査の方法や結果を批判的に考察し表現する姿も確認できた。

グループ活動が中心だったが、個人の活動でより統計的問題解決力を高めることができると考える。

2つ目として、シミュレーションを取り入れた標本調査を体験することで、標本調査の理解を深めることができたことである。生徒の記述から、標本調査の意味や標本の大きさへの理解が深まったこと、さらなる疑問をもって単元を終えたことが伺える。

(2) 課題

1つ目として、記述統計と推測統計の緩やかな移行を、場面化していくことである。標本調査は推測統計の始まりであることを想起させ、よりよく問題解決することができるようにしたい。

2つ目として、無作為抽出することについての正しい理解である。一見無作為に見えても、そうでないことの理解は難しい。これには指導者の深い教材研究が求められる。用いた方法が適切だったかどうかや、次の問題発見へとつなげる視点等、指導者の力量が問われるだろう。

6 まとめ

本研究では、生徒が自ら問題を発見し、データを収集して分析し、問題を解決するような統計的問題解決力を育成する授業や指導計画の在り方を検証してきた。改めて小・中・高の系統性について考えていく必要があると感じた。小学校での学習の上に立ち、さらには高校への扉を小さくノックするような中学校の学びを築いていきたい。そして、生徒が生涯にわたって、統計を用いて豊かな人生を切り開いていけるよう、授業づくりや指導に精進したい。

引用・参考文献

- 1) 文部科学省 (2018) 『中学校学習指導要領 (平成 29 年告示) 解説数学編, 日本文教出版 pp. 9-13
- 2) 渡辺美智子 (2014). 「不確実性の数理と統計的問題解決力の育成-次期学習指導要領の改訂に向けて-」, 『日本数学教育学会誌』 第 96 巻, 第 1 号, 2014, p. 33-37

教師による授業への統計的探究プロセスの実装の試み

佐藤 寿仁 岩手大学教育学部

岩手県盛岡市上田三丁目 18-8 toshis@iwate-u.ac.jp

1 はじめに

小中高のすべての校種において、新しいカリキュラムとしての学習指導要領が全面実施され、新カリキュラムでの学校教育が進められている。これまでのコンテンツベースから、コンピテンスペース、つまり資質・能力の育成を中心としたカリキュラムとなった。中央教育審議会(2021)は、これまでの日本型学校教育を振り返り、一定の成果を示しつつも、今後の時代において、社会からの要請を見据えた、「令和の日本型学校教育」のあり方を示した。その中で、授業で問題解決する子供たちの伴走者となる教師の在り方などにも触れ、子供の学び方を中心に考え、資質・能力の育成を目指すため、より一層学びが充実する授業が求められている。

2 問題の所在

(1) 学校教育において求められる教師の姿

令和の日本型学校教育の推進にあたっては、子供の学び方や教師の支援や促しなどが注目されており、子供の学び方として、「個別最適な学び」と「協働的な学び」の両方が重視されている。また、GIGA (GIGA: Global and Innovation Gateway for All) スクール構想での ICT 活用や学習者の多様性の認知など、学校教育を取り巻く環境に劇的な変化が訪れているといえる。特に、児童生徒への 1 人 1 台端末の配付とその活用など、教師には新しい環境やその変化への対応、指導力やスキルの向上がますます求められるようになった。

(2) 学校教育において重視される統計教育

誰もが予期しなかった新型コロナウイルス感染症 (COVID-19) の感染拡大は、社会へ大きな打撃となったとともに、私達の生活様式に大きな変化をもたらした。私達人類は、自然現象だけでなく、人間社会の動向そのものを予測し、その事態によりよく対応するための力を重視することとし、結果、教育へ大きな期待が寄せられているといえよう。小学校学習指導要領 (平成 29 年告示) 解説算数編 (以下、「学習指導要領解説算数編」と示す) では、新領域として「データの活用」が設定された。また、中学校では、「資料の活用」とい

う領域名が「データの活用」と変更され、小学校算数と領域名が同じとなった。領域名が揃えられたことから小中の指導に接続も求められている。

(3) 学校教学における統計教育の実践

学習指導要領解説算数編では、内容だけでなく、学習者の問題解決の過程に着目したプロセスが説明されており、「統計的探究プロセス」をあげている (図 1)。

問題	・問題の把握	・問題設定
計画	・データの想定	・収集計画
データ	・データ収集	・表への整理
分析	・グラフの作成	・特徴や傾向の把握
結論	・結論付け	・振り返り

図 1 統計的探究プロセス

統計的探究プロセスとは、統計的な問題解決活動を指すものとし、「問題-計画-データ-分析-結論」という段階があるとした。これは、Wild & Pfannkuch (1999) のいう PPDAC サイクル (図 2) からきている。

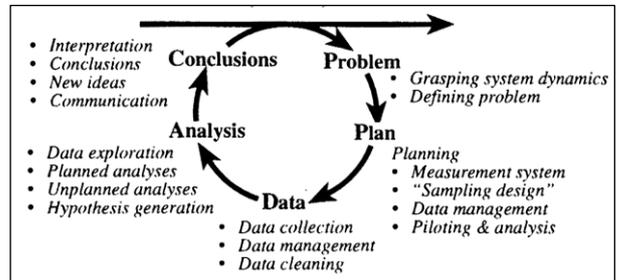


図 2 PPDAC サイクル

学習指導要領解説算数編では、高学年から統計的探究プロセスの過程を意識して、学習者が探究的に学ぶことの重要性を述べられており、実際に統計的探究プロセスについて教科書各社に掲載されている (図 3)。



図 3 教科書に掲載された PPDAC サイクル

中学校学習指導要領 (平成 29 年告示) 解説数学編 (以下、「指導要領解説数学編」と示す) においては、同様

の記載がみられるものの、図1のような具体的なものは示されていない。なお、高等学校学習指導要領（平成30年告示）解説数学編には算数と同様の記載がみられ、校種によってその示し方に差がみられる。探究過程を教師が受け止め、授業への実装が求められている。

菊野・栢元（2020）は、中学校数学での統計的探究プロセスを遂行する単元の開発及び授業の実践を試み、統計的な問題解決において、データの収集、整理するプロセスを授業者が行い、データやグラフを分析する時間を十分に確保することの大切さに加え、学習経験を積むことで、サイクルを2回にわたって回し考察し続けたことを報告している。このように先駆的な実践者が中心となって、統計教育の充実が図られてきた。

藤井（2021）は、統計的な問題解決のプロセスを授業の中に取り入れるということについて、日本はまだ初期の段階であることを指摘し、問題点として時間の確保を挙げている。また、教科横断的な学習にも可能性をつなげ、学校による統計教育のカリキュラム・マネジメントの必要性について述べている。統計的探究プロセスの授業への実装は、まだまだ開発期にあると考えられ、今後の充実期を目指すためには、より多くの教師による実践と、教材研究や指導計画の立案、指導技術の周知が必要となるだろう。

(4) 統計的探究プロセスの授業化への障壁

小中高における算数数学の学習指導要領への統計的探究プロセスの記載により、多くの教師がこのプロセスの認知に至ったと考えられる。国立教育政策研究所（2019）は、平成31年度（令和元年度）に実施した全国学力・学習状況調査中学校数学報告書において、調査結果及び分析のみならず、調査問題を活用した授業づくりのために、統計的探究プロセスを重視した授業を行うことにより、データの活用領域における授業が改善されるとし、そのサイクルでの場面や内容について、その具体例について紹介している（図4）。

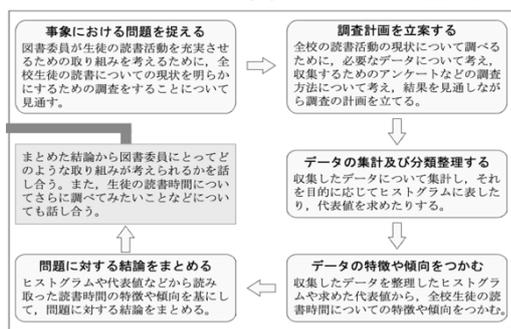


図4 調査問題における統計的探究プロセス

このことは、問題の正答率だけにフォーカスをするのではなく、問題解決する過程そのものを重視して理解を促すことをねらうものと考えられる。

また、青山（2016）は、現状での統計教育の課題の1つとして、統計的探究を通常の授業に取り入れ、児童・生徒が自ら遂行できるように指導を構成する必要性を述べている。一般的な教師が授業実践として成立させる要素について検討し、統計的探究の授業化への配慮事項について7点を事例的に明らかにした。児童生徒自身で統計的探究プロセスを遂行し、統計的な問題について解決をすることができるようになるために、指導する教師の行動は簡単なものではないことがわかる。

(5) 問題発見・解決能力と統計的探究プロセス

文部科学省（2019）は、育成を目指す資質・能力として3つ（[知識及び技能]、[思考力、判断力、表現力等]、[学びに向かう力、人間性等]）と説明しているが、小学校及び中学校学習指導要領（平成29年告示）解説総則編（2019）」において学習の基盤となる資質・能力として3つが取り上げられている。これらの資質・能力は算数・数学のみならず、教育課程一般で育成され、子供たちの学びを後押しするものとして重視されている。

[学習の基盤となる資質・能力]

○言語能力 ○情報活用能力 ○問題発見・解決能力

特に「問題発見・解決能力」は、問題解決の過程とそれを遂行する力を指し、まさに、学び方重視の傾向を示している。このことから授業に統計的探究プロセスを実装することは極めて大切であり、先駆的な実践者だけではなく、多くの教師が実践するといった裾野を広げることが必要であり、学校数学における統計教育を充実期に移行させていくべきではないだろうか。

3 研究の内容

(1) 目的

本研究の方向性となる目的について、2点あげる。この視点に沿って検討したことを報告する。

目的Ⅰ：校種を越えた教師との対話で統計的探究プロセスをカリキュラムベースでの実装化について検討し、課題を明らかにする。

目的Ⅱ：教師の授業研究コミュニティでの統計的探究プロセスの実装の可能性を検討する。

(2) 方法

教師は省察活動により、その実践力が育まれるといわれてきた。本研究においても、教師の省察活動を主

として授業開発に取り組んだ。ここでは、佐藤(2022)が試案した Lesson Plan-Sheet をもとに、国立大学法人附属小学校、同附属中学校教員が具体的に作成した Lesson Plan-Sheet を見ながら、振り返ることで省察活動を進めていく。

教師が指導技術を高めるための筋道として、より現実的なアプローチが求められる。コルトハーヘン(1985)は、教師の省察活動モデルを提唱し、それを5つの局面に分けて考える ALACT モデル(図5)で詳細に説明している。本研究での省察活動については、このモデルに沿って行い、教師どうしで行う省察として段階2、3を中心とする。そのことで、授業への統計的探究プロセスの実装の在り方を検討する。

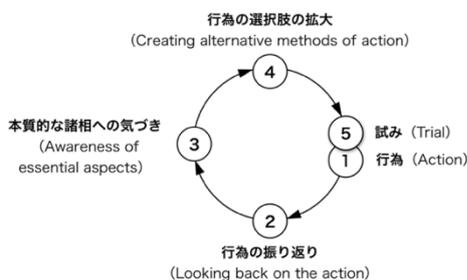


図5 ALACT モデル

国立大学法人附属小学校、同附属中学校の教員それぞれ1名ずつ、大学教員1名の3名を授業研究コミュニティと位置付け、共同・協働を軸に省察活動を行った。この3名については、数学教育を専門としており、大学教員については中学校での勤務経験がある。具体的には、次のように進めた(表1)。対話に参加した教師は、本研究において教師集団の代表とするが、附属教員ということで限定的であること、また、省察活動による捉えについては、あくまで質的なものであることとし、考察を進めていく。

表1 検討のスケジュール

	検討内容
第1回	Lesson plan : 小6の振り返り
第2回	Lesson plan : 中1の振り返り
第3回	統計的探究プロセス実装化の検討
第4回	Lesson plan の作成・検討
第5回	Lesson plan の最終検討

(3) 検討の結果

実際に作成し授業も行った Lesson Plan-Sheet (図6)について振り返り、「問題解決と探究」を検討した。



図6 省察に使用した Lesson plan-Sheet (小6)

問題解決と探究

統計的探究プロセスをどう意図し、位置付けたのかについて振り返った。さらに実装の可能性についても意見交換を行った。検討は Google スライドの共同編集機能を用いて、図7のように、横軸を PPDAC の実装の度合い、縦軸を児童生徒にとっての学習効果の程度を2軸での図に表した。

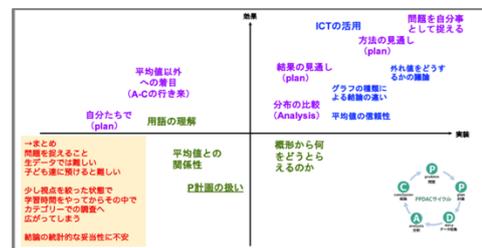


図7 問題解決と探究を協議 (小6)

小学校教師は、PPDAC サイクルを重視しようとしてはいるものの、用語の理解、平均値の意味など指導内容について、過程でそれらの指導をどのように位置付けるかについて課題とし、授業する際も困っていた。過程よりも当該学年の学習内容の習得の程度が気になっていたといえる。また、教科書に沿って授業を考えた際に、教科書の流れ(進行)と PPDAC サイクルがどう関連しているのかわからないことにも言及していた。中学校教師は、小学校教師と異なった点として、データには外れ値として扱うことができるデータに気づくなど、より専門性のある指摘がみられた。PPDAC サイクルの必要性は共有・共感ができたが、教えるべき内容の習熟や理解の程度に重きをおく傾向がみられた。

さらに、算数・数学の内容の関連を確認した(図8)。互いの内容の確認にとどまらず、校種間の接続を重視した指導法についても話題となった。ここでも専門的な知識の不足を不安視していることがわかった。

関連の実装・効果	教師の困難性
<p>指導する際、中学校ではどのような扱いがされているか(児童へ)少しづつ情報を伝える。データの領域に関わらず、同様に小学校内でもそのような指導を行っている。(児童が覚えていないことも多い)つまり効果は高くないかもしれない。</p> <ul style="list-style-type: none"> ・ドットプロット ・数学的表現など ・学年会など教師集団での 日常的に 大きい学校で <p>代表値のよさや限界に、より触れられる(深い意味理解へ)</p> <p>素地があるので、批判的な思考に入りやすい</p> <p>PPDACサイクルの問題解決に時間をとれる(経験的)</p> <p>数形や代表値など複数の視点を用いて意見を述べられる(言語的表現の深み)</p>	<ul style="list-style-type: none"> ・中学数学の内容を知らない。 ・結論を出すところまで到着? ・表現方法、用語の指導で児童も教師も手一杯? ・ヒストグラム?棒グラフ? ・微妙にデータ数が異なる教科書の事象の意味への気づき。

図8 算数・数学との関連について

学校設定科目「データサイエンス」と課題研究との接続

神戸大学附属中等教育学校数学科教諭
神戸大学数理・データサイエンスセンター客員研究員
林 兵馬(h-hayashi@people.kobe-u.ac.jp)

1. はじめに

AI戦略2019・2021に示されている通り、近年AI・DSがますます求められている。高等学校においても、令和4年度スタートした新学習指導要領においても、ますますAI・確率・統計分野の充実が求められている。

令和2年度より新学習指導要領の改訂を見越し、神戸大学附属中等教育学校の数学科として、中等4・5年（高等学校1・2年相当）の週1時間「データサイエンスI・II」を実施してきた。

多くの他校と異なる点は、「総合的な学習・探究の時間」ではなく、「数学科」の授業として設置している点である。

主な学習内容は以下

- ・データサイエンスI（中等4年・週1時間）
数学I・Bの学習項目を中心
- ・データサイエンスII（中等5年・週1時間）
情報I・IIの学習項目を中心

また勤務校では、「総合的な学習・探究の時間」をKP（Kobe Port Intelligence Project）と称して、中等1～2年：主に学年・グループ研究
中等3～6年：個人研究に取り組んでいる。

生徒は入学以来、KPや様々な教科・科目において教科の学習をはじめ、本校での様々な学びを通じて、「見つける力」、「調べる力」、「まとめる力」、「発表する力」、そしてそれらを総合する「考える力」を身につける。

学校設定科目「データサイエンスI・II」の設置目的もKPにおける個人研究に寄与することも目的としている。

2. 本報告の概要

本報告では、学校設定科目「データサイエンスI・II」を受講した生徒のKPにおける「課題研究」の内容や、指導の様子、個人研究に入るまでの中等1～2年生の指導などについて、を以下の観点を中心に報告する。

- ・個人研究においてどのようにサポートをしていくべきか
- ・中等1～2年や個人研究をスタートした3年生の段階において、課題研究の全体講義の枠組みでどのような授業内容が考えられるか
- ・学校設定科目「データサイエンス」の内容の再検討

本取り組みは、国立研究開発法人科学技術振興機構の支援を受けている。また、神戸大学数理・データサイエンスセンターの協力を得ている。

高校選択授業におけるデータサイエンス教育の実践事例

小森 晴菜・馬場 国博

慶應義塾湘南藤沢中等部・高等部

〒252-0816 藤沢市遠藤 5466

TEL (0466)49-3585 FAX (0466)47-5077

E-mail: haruna@sfc-js.keio.ac.jp

1. はじめに

本校の高校3年生を対象とした2単位分の選択必修科目(学校設定科目)「データ科学」は、R言語を用いてデータを分析する力を身につけさせることを目的に開設しており、今年度は小森(発表者)が担当し、履修者数は20名であった。

前年度同様、今年度も1人1つ以上のコンテストに参加することを必須とした。

[参加したコンテスト一覧]

A. 「第15回データビジネス創造コンテスト」(慶應義塾大学SFC研究所データビジネス創造コンソーシアム主催): 4組10名が応募

B. 「第6回和歌山県データ利活用コンペティション」(和歌山県主催): 5組10名が応募

※「統計データ分析コンペティション2022」(総務省統計局、統計センター、統計数理研究所、日本統計協会の共催)には今年度不出場

Aでは最終審査会(高校生チーム2組、大学生チーム9組)に1組3名が進み、高校生部門賞を受賞することができた。Bでは最終審査会(高校生チーム7組、大学生チーム7組)に1組2名が進み、NTT西日本賞を受賞することができた。

授業の前半(4月~7月)は、重回帰分析、主成分分析、判別分析、クラスター分析といった多変量解析の手法を扱った。授業の後半(9月~11月)は、それぞれが参加するコンテストへ向けたデータ分析や提出資料作りに取り組み、お互いの発表を見て相互評価を得た。

コンテストへの参加を通して、データ解析やその結果を的確に表現する方法を実践的に学ぶことができた。

2. コンテストに向けての取組み

(1) 授業展開

「データ科学」の履修対象である本校の高校3年は

4~11月が通常授業の期間であり、12月以降に授業の開講はない。そのため、授業時間数は1回あたり2時間連続(100分間)×14回である。

4~5月の4回の授業では、Rの基本文法、基本統計量の算出、グラフ表現(ヒストグラム、散布図など)、単回帰分析、データのRへの取り込み方を扱った。

6月の4回の授業は、重回帰分析、主成分分析、判別分析、クラスター分析の学習の時間にそれぞれ1回ずつ充てた。その後、7月の1回の授業で出場するコンテストの確認とテーマの方針決定を行った。

夏休みおよび2学期の動きは、出場したコンテストによって大きく異なった。9月上旬締切の「データビジネス創造コンテスト」に出場したチームは、夏休み中にデータ分析や発表資料作りを進めて完成させ、9月の授業1回分を使って研究発表を行った。10月中旬締切の「和歌山県データ利活用コンペティション」に出場したチームは、夏休み中は特に作業を進めず、2学期が始まってから取りかかり始めた様子であった。9~10月の授業2回分を使って作業に取り組みさせて完成させ、11月の授業1回分を使って研究発表を行った。

(2) 「データビジネス創造コンテスト」の出場報告

このコンテストは、産・学が力を合わせてデジタルの力を使って世の中を変えるための政策課題にチャレンジすることを目指し、高校生・大学生がデータを広く深く分析し、新たな知の抽出や価値の創出を競うものである。

第15回のテーマは「寿命100歳時代を生き抜く知恵~疾病予防と病後の幸せな生活~」で、企業から提供された生活習慣病患者のレセプトデータを多角的に分析し、疾病を予防でき、病気にかかったとしてもその後の生活を豊かに送るためのアイデアを提案した。

[4組の応募者の発表タイトル]

- ・福やく-服薬習慣を定着させるアプリー（高校生部門賞受賞）
- ・クラスター分析を用いた保険プランの提案
- ・糖尿病の予防と発症後のためのアプリケーションの提案
- ・生活様式の変化による運動不足の改善の提案

(3)「和歌山県データ利活用コンペティション」の出場報告

このコンテストは、和歌山県が総務省統計局や独立行政法人統計センターとともに、次世代の日本を担うデータサイエンス人材を育成すべく、行政課題に対するデータを利活用した解決アイデアを競うものである。

第6回のテーマは「人口減少社会における人や企業をひきつける地域づくり」で、活力が弱まっている地方についてのデータを自ら集め、地域活性化につなげていくためのアイデアを提案した。注目する都道府県市区町村を決めてデータを集めることが要求される。

[5組の応募者の発表タイトル]

※ []内は、取り上げた市区町村

- ・谷戸地域を鬱病療養地に[神奈川県横須賀市]（NTT西日本賞受賞）
- ・安定した暮らしを見つけよう！ワーキング@設楽[愛知県設楽町]
- ・限界集落・自治体の飛び地化の提案[高知県土佐清水市]
- ・繋げて広げる中津愛政策！[大分県中津市]
- ・複合型商業施設の提案[岡山県倉敷市]

最終審査会に出場した高校生チームのほとんどが現在の居住地または学校の所在地をテーマとして取り上げていたが、本校では自分が生まれ育った場所やゆかりの地をテーマとしているチームが多かったのは興味深い。その分、指導教員には幅広い知識が必要とされるわけだが、データサイエンス教育に必要な教科横断をした総合的な知識の必要性を教員が感じとる良い機会でもあるといえる。

(3) 反省点

コンテストへの参加について、履修者からは以下のような感想が寄せられた。

- ・Rの使い方を理解することができないままコンペに臨むことになり、やりたいデータ処理（主成分分析、クラスター分析）ができなかった
- ・夏休みに先生に質問をしながら取り組みばよかった
- ・最終審査会で他のチームがクラスター分析を使った発表をしていた、ああいったものができればよかったと思う

これらの反省点は、夏休み前の授業で一通り多変量解析の手法のインプットをさせたいあまり、いわゆる「詰め込み型」の授業となってしまったことに起因していると考えられる。

また、異なる種類のコンテストに参加する履修者たちを同時に指導することにも課題が残った。締切時期が異なるため、授業内で与えられる作業時間に差ができてしまったこと、さらには予選を通過したチームと通過しなかったチームでも足並みが揃わなくなることである。こちらは、授業担当者のマネジメント不足をひどく痛感した。なお一方で、発表者の研究発表をそのコンテストに参加していない履修者が聞くことで、データの見せ方に対する素朴な疑問が飛び交ったり、重大な論理の欠落に気がついたりするというメリットもあった。

3. 今後の課題

今年度は「統計データ分析コンペティション」に出場した生徒が一人もいなかった。これは、論文での発表を遠ざけ、慣れ親しんだプレゼンテーション形式での発表を好む傾向にあったことが考えられる。その為、今後は4月の段階で過去の受賞論文を紹介し、データ分析が論文へ仕上がるまでの過程を見せたい。

加えて、今後は「中高生スポーツデータ解析コンペティション」への参加を視野に入りたい。この授業で何のデータを分析したいか、という質問を初回授業で履修者へ投げかけたところ、多くの生徒がスポーツを挙げた（野球、サッカー、テニス、バスケットボールなど）。しかし、1月末締切のスポーツデータ解析コンペティションに「データ科学」の授業の一環として取り組ませるのは時期の問題で不可能である。そこで、現行のカリキュラムで本校生徒にスポーツデータ解析をさせるには、高校1,2年生の段階でコンペの存在を周知させることが必要だと考える。そこで面白さを感じた生徒が半年間データ科学を履修し、自主的に高校3年時に再挑戦する流れが望ましい。

最後に、コンテストにこだわらず好きなデータについて好きなだけ分析する時間を確保することも重要だと感じる。今年度は任意課題に取り組むことを推奨したところ、2名の履修者が分析結果を発表した。

[任意課題テーマ]

- ・音楽配信アプリのヒットチャートトップ30曲のクラスター分析
- ・プロ野球選手のホームラン数を決定づける打撃記録の要因分析

これらの発表では、実に有意義な質疑応答が行われた。発表者と同じ趣味を持つ者は熱心に分析結果の解釈を考え、何より発表者本人が楽しそうに話していたのが印象的である。データサイエンスの第一歩を踏み出させる場において、コンテストには参加しながら「詰め込み型」中心の授業から「自らスキルを獲得していく」授業へとうまくシフトさせていくことが目指したい授業形式である。

(注) 本授業実践報告は慶應義塾学事振興資金による共同研究「初等中等教育段階におけるデータサイエンスプログラムの開発」の助成を受けて行った。

探究への連携を見据えた数学の授業としてのデータサイエンス教育の試み

中田 雅之(神戸大学附属中等教育学校)

連絡先(mnakada@people.kobe-u.ac.jp)

1 本実践の背景

近年、機械学習の発展とともに、収集された大量のデータを処理することが以前に比べて容易になってきており、私たちは簡単にデータを収集し、分析することが可能になった。それは同時に、私たちがいかにデータを見るかという、統計的な技術やリテラシーがより求められるようになったことを意味している。実際、今年度より試行されている、高等学校の新学習指導要領においても、統計・データサイエンスへの理解は重要視され、従前のものに比べて内容の拡充が測られている [3, 4]。例えば、数学科においては、

社会生活などの様々な場面において、必要なデータを収集して分析し、その傾向を踏まえて課題を解決したり意思決定をしたりすることが求められており、そのような脂質・能力を育成するため、統計的な内容の改善・充実を測った [4]

とある。生徒は、統計的な理解を深め、数学の問題としてだけでなく、現実の問題の解決に利用するような活動が期待されている。

そのような状況を鑑み、発表者の勤務校でも、数学科として統計教育に本腰を入れることになった。勤務校は2020年度よりSSHに指定されている。このSSHのプロジェクトの目玉として、勤務校では独自科目「データサイエンスⅠ・Ⅱ」を設置することになった。発表者の勤務校は中高一貫校であり、その4年次(高校1年次相当)、5年次(同2年次相当)にそれぞれ1単位ずつ統計の科目が開講されることとなる。

発表者の勤務校では、かねてより探究が盛んである。勤務校では、総合的な探究の時間をKP(Kobe Port Intelligence Project)と称している。そこで、生徒は入学以来さまざまな探究の手法を学び、卒業論文を提出して卒業する。探究の過程では、実験やアンケートの結果の分析などで、統計的手法を用いる生徒も多い。勤務校の科目「データサイエンスⅠ・Ⅱ」は、生徒たちが探究を進めるうえで、(中等生なりの)正しい統計的理解を身につけるためのサポートとしての役割も担っている。

発表者も早速この講座を担当することになった。そうは言っても、自身も15年くらい前の学生時代に統計学の入

門の講義を受けて以来のことであった。統計を専門的に学習したこともなければ、プログラミングにそこまで親しんできたと言えるわけではない。そのため、教材研究と授業準備にはかなりの時間とエネルギーを費やすことになった。今となっては、統計の分野を勉強するいい機会になったと思っている。

2 本実践の概要

本実践では、同一の集団に対して、「データサイエンスⅠ・Ⅱ」における2年間の継続した指導を行った。2年間の指導の中で、扱った内容は以下の通りである(内容の選定は、「情報Ⅱ」程度までを参考にしている [5])。

4年生春学期 「記述統計」

- 代表値
- 箱ひげ図
- 分散, 標準偏差
- 散布図と相関, 共分散, 相関係数

4年生秋学期 「推測統計」

- 二項分布と推定, 検定の考え方
- 数値計算と区分求積法
- 連続型確率分布の例と性質
- 正規分布とその性質, 大数の法則
- 正規分布による区間推定
- 仮説検定
 1. 帰無仮説と対立仮説, 検定の過誤, 検出力
 2. 二項検定, Z検定, t検定, χ^2 検定

5年生春学期 「教師あり学習」

- 数学的基礎づけⅠ(ベクトルと行列, 偏微分)
- 線型回帰分析(単回帰分析と重回帰分析)
- ロジスティック回帰分析
- 決定木

5年生秋学期 「教師なし学習」

- 数学的基礎づけⅡ(固有値, 対角化)
- 主成分分析
- クラスタリング
- PBL演習

統計に限ったことではないが、学んだ内容の定着のためには、実際に学んだことを使って何かを調べてみるものが

有効である。そのため、それぞれの項目において、学んだ内容を確認するための課題を設定した。また、上の一覧には記載しなかったが、その時々に応じて、生徒が小集団や個人で自由にテーマを設定し、データの収集、分析を経て相互発表を行うという活動の時間も用意した。

指導にあたっては、表計算ソフトのみならず、プログラミング言語である Python [6] を用いることにした。これらのツールを、生徒間で実行環境になるべく差が出ないように、勤務校で一括して提供している Google Workspace for Education [1] の中で行った。授業の中で Python を利用する場合は、(動作環境が安定する限りは)このサービスに含まれる Google Colaboratory を利用するのが便利であると思う。区間推定や仮説検定の学習を最終目的とするのであれば、ツールを表計算ソフトに限ることも考えられる。ここで敢えて統計分析のツールとして Python を利用するねらいは、線型回帰分析以外の機械学習の手法を授業で扱うこと、そして KP での探究の場面でこれらの手法を生徒が利用する可能性を見越してのことである。

なお、本実践の裏では、生徒たちは情報科の授業においても、Python を用いた実習を行っていた。情報科の授業では、Python を利用して、条件分岐やループなどの基本的な書法を学んだようである。数学の授業において Python を用いた実践例は、現在のところまだそれほど多くないように感じている。Python は統計計算のモジュールが充実しているが、他に整数分野の探究に親和性がありそうである [2]。本実践を経て、数学科の授業として Python をどのように利用すべきかについては、議論があつてよいと思っている。

さらに、本実践においては、取り扱う分析手法の背後にある数学的な原理の解説も(高校生に扱えると思われる範囲で)行った。これは、次のような考えからである；

高度に進んだ統計的手法が広く扱われるようになった要因として、その手法の基礎となる数学的背景をアルゴリズムに委ねてブラックボックス化しているところにある。しかし、数学的背景を全く見ないまま、道具としての統計的分析のツールを用いるのでは、統計に対する深い理解には結びつかない。本実践は、数学科としての統計教育の試みであるから、統計的分析手法を支える数学に触れる機会を提供することも本実践の使命である。

統計の授業を授業で行う上で悩ましいのは、統計の分析に用いられている数学が中等教育を受けている最中の生

徒にとっては、高度なものも多く含まれている点である。学習の円滑さを考慮すれば、高度に難解な内容には一切触れずに結果を認めることにして済ませることも可能であった。しかし、それによって省かれるものこそが、数学が「役に立つ」ものであることの証でもある。中等教育課程を修了すれば、生徒たちはともすれば数学が科学技術に直接結びつく身近で華やかな例に対して、無自覚なままその先の専門分化した世界に進んでいくかも知れない。それはとてももったいないことだと、発表者は考える。そのため、本実践では、かなり数学的な内容にも踏み込んだ教材を提供している。

本実践は、このまま他の学校で統計や機械学習に係る数学の授業実践として参考にしていただけるような形にはなっていない。扱う内容は、授業の対象とする生徒によって注意深く考えるべきであると思う。将来的には、本実践の内容を整理する必要があると、発表者は考えている。しかし、その際には、生徒がより進んだ手法を自習したいときに、必要な情報にアクセスできるようなサポートとともに、そこに至るための橋渡しとなるような形で情報提供できることを目指したい。

本実践は勤務校の SSH のプロジェクトとして、国立研究開発法人科学技術振興機構の支援を受けている。

参考文献

- [1] Google Workspace for Education, https://edu.google.com/intl/ALL_jp/workspace-for-education/editions/overview/, 2023/02/23 閲覧。
- [2] 松本昌也, 清水克彦, Google Colaboratory を用いた実験数学教材の開発: Python で完全数・メルセンヌ素数を探究する(数学ソフトウェアとその効果的教育利用に関する研究), 京都大学数理解析研究所講究録, 第 2208 巻, pp.20-30.
- [3] 文部科学省, 高等学校学習指導要領解説 情報編, 2018.
- [4] 文部科学省, 高等学校学習指導要領解説 数学編理数編, 2018.
- [5] 文部科学省, 高等学校情報科「情報Ⅱ」教員研修用教材, 2020.
- [6] Python Japan, <https://www.python.jp/>, 2023/02/23 閲覧。

「情報Ⅰ」統計ポスターを用いたデータサイエンス教育実践 ～データの活用×情報デザインの体系化～

雲雀丘学園中学校・高等学校 林 宏樹
連絡先 (Email h-hayashi@hibari.ed.jp)

1. はじめに

2022 年度より高等学校において新たな学習指導要領のもと、教育活動が行われている。教科「情報」では共通必修科目「情報Ⅰ」と選択科目「情報Ⅱ」が設定され、プログラミングとデータサイエンスに関する内容が充実したことが特徴的な点である。一方、教科「情報」において「問題解決」は重要なキーワードであり、コンピュータを用いて問題解決のプロセスを実践することが求められている。このことを踏まえ、問題解決のプロセス「問題発見・明確化→データの整理と変換→データ分析と可視化→結果発表と解釈→分析の評価」を情報Ⅰで行うカリキュラムと授業実践が依然少ないことを懸念している。

以上より、本稿は高等学校「情報Ⅰ」において統計ポスターを用いたデータサイエンス教育実践を紹介する。特徴としては、データの活用、プログラミング、情報デザインを体系的に組み立て、問題解決につながる探究活動を行い、統計ポスターの作成を行う展開である。

2. 「情報Ⅰ」年間計画

高等学校学習指導要領情報編では「情報Ⅰ」は、問題の発見・解決に向けて、事象を情報とその結び付きの視点から捉え、情報技術を適切かつ効果的に活用する力を全ての生徒に育む科目としている。内容は、(1)情報社会の問題解決、(2)コミュニケーションと情報デザイン、(3)コンピュータとプログラミング、(4)情報通信ネットワークとデータの活用とである。本実践の年間計画を表1に示す。

前提として、作成した統計ポスターは兵庫県統計グラフコンクールへ提出する。その提出期限が9月初旬である。取組1では「データの活用」を実施し、夏季休業中の課題として統計ポスター(作品①)を作成する。取組2では、「プログラミング」の基礎基本を学ぶ。取組3では、「情報デザイン」の内容を学び、統計ポスタ

表1 「情報Ⅰ」授業年間計画

	時期	授業内容
取組1	4月～7月	(1)情報社会の問題解決 (4)データの活用
	8月	統計ポスターの作成(作品①)
取組2	9月～12月	(3)コンピュータとプログラム (4)情報通信ネットワーク
	1月～3月	(2)コミュニケーションと情報デザイン
	3月	統計ポスターの作成(作品②)

ー(作品②)を作成する。1年次の9月上旬に作品①、2年次の9月上旬に作品②を統計グラフコンクールに提出する計画を立てた。本実践は、雲雀丘学園高等学校の高校1年生185名の「情報Ⅰ」(2単位)の教育実践である。

3. 取組1の内容

「データの活用」に関する講義・演習を行った。データサイエンス教育に直接的につながる分野である。1回50分の授業を全19回実施し、授業内容を表2に示す。段階的な指導展開として、基礎学習、探究1、探究2、探究3と段階的に指導した。なお、授業時数の都合により探究1の動画視聴、技能習得を行い、演習を行うことができないクラスもあった。探究3は、第10回・第16回に外部講師の講座を行った。

取組1の特徴は、統計的探究プロセスPPDACサイクルを段階的に指導したことと、動画教材を用いた指導を行ったことである。

3.1 PPDACの段階的な指導

取組1では、PPDACを段階的に指導した。PPDACサイクルとは、「問題(Problem)－計画(Plan)－データ(Data)－分析(Analysis)－結論(Conclusion)」の5つの段階からなる統計的な探究活動を行う上での問題解決プロセスである。

表 2 取組 1 の具体的な授業内容

基礎学習 第 1 回～第 5 回
情報と情報社会, 問題解決の考え方, 文書作成 表計算, 発表ソフト, 表計算の技能習得 グラフの描画方法, 統計量の理解, 度数分布表 ヒストグラム, 箱ひげ図, 分散, 標準偏差, 尺度水準, 単回帰分析, 相関係数
探究 1 第 6 回～第 9 回
動画教材アの視聴, e-stat, 気象庁のデータ収集 家計調査と気温との相関分析演習
探究 2 第 12～第 15 回
動画教材イの視聴, 標準化に関する指導 標準化したデータを分析する演習
探究 3 第 10 回・第 11 回・第 16 回～第 19 回
PPDAC サイクルの理解, ポスター構成についての講義, 外部講師による統計ポスター講座, 統計ポスター演習

基礎学習では DA の知識・技能を習得させる。探究 1 では DA に焦点を当てた演習を行う。探究 2 では DAC に焦点を当てた演習を行う。探究 3 では PPDAC サイクルによる統計ポスター作成を行う。この指導手順によって、データを処理するだけでなく、出力された結果を考察することが重要であることを理解させ、PPDAC サイクルによる問題解決につながる統計ポスターを作成させた。

3.2 動画教材を用いた指導

探究 3 において生徒自身が興味関心をもった内容に対して問題解決に向けた統計ポスターを作成する前に、具体的な事例からデータ収集方法、整理整形、分析、考察について学ぶために動画教材を活用した。動画教材は経済産業省「未来の教室」の STEAM ライブラリーの動画コンテンツを活用した。

4. 取組 2 について

「コンピュータとプログラム」と「情報通信ネットワーク」に関する講義・演習を行った。「コンピュータとプログラム」は 13 回、「情報通信ネットワーク」はクラスにより多少異なるが約 7 回行った。

このうち、プログラミングの 1 回の授業において、総務省の「家計調査報告」からと都市別消費金額データを取得する方法を行った。データサイエンス教育につながる内容として、プログラミングを用いてデータ収集する方法を習得させた。

5. 取組 3 について

「コミュニケーションと情報デザイン」に関する講義・演習を行った。授業の目的は、情報を整理したり、目的や意図をもった情報を受け手に対して分かりやすく伝達するためのデザインの基礎知識や表現方法を理解し表現する技能を身に付けることとした。

まず、情報デザインに関する理解、ピクトグラムやシグニファイア等の理解を行った。次に、学んだ情報デザインの観点から統計ポスター（作品①）の振り返りを行う。そして、原則、作品①を題材にし、その内容を伝える受け手の属性などを定めた統計ポスター（作品②）を作成する演習を実施した。なお、作品①を題材とせず、それ以外の内容で作品②を作成してもよいし、新たにデータ収集、分析、考察を加えてもよい。データ収集ではプログラミングを活用したデータ取得をしてもよい。

さらに、情報デザインの発展的な内容として企業のデザイン部門の専門家の講義や、大学で実施しているデザイン思考を学ぶ講座を実施した。これらの講義によって、理論的に受け手の背景を読み解き、デザイン構成に取り入れたポスター作成を期待した。

このように、生徒自ら作品①を評価・改善を行い、作品②を作成する。作成者である生徒が定めた受け手の属性に対してわかりやすく伝わるデザインに構成することを目的とした演習である。

6. おわりに

本実践では、統計ポスターを題材に「情報 I」の授業内容を組み立て、統計ポスターを 2 回作成することができる展開である。取組 1 で統計ポスターを作り、取組 2 で習得したプログラミングを活用してデータ収集の技能を身に付け、取組 3 で受け手に伝わりやすいデザインを考える統計ポスターを作成する。「情報 I」の内容を体系的に結びつけ、問題解決につながる探究活動が実施できた。1 年間の 1 科目の授業内で問題発見・分析・まとめ・評価・改善・次なるアクションを行うことができるサイクルである。

課題としては、取組 1 の探究回数工夫、段階的な指導の内容改善、プログラミングの活用事例の工夫、情報デザインの内容検討が挙げられる。

子どものワクワクを育てる「数式抜きのデータサイエンス」

～身近なデータから世界を推測～

都丸 希和 (名古屋大学教育学部附属中・高等学校)
 佐藤 健太 (名古屋大学教育学部附属中・高等学校)
 愛知県名古屋市千種区不老町 (052) 789-2680
 tomaru@highschl.educa.nagoya-u.ac.jp

1. 本校の課題研究について

本校では、文理を越えた科学的思考力の育成目指し、「探究し続ける力」、「他者と協同する力」、「新たな価値を生み出す力」、「科学的思考と活用力」の4つの力の育成を目標としている。この目標に向かい、高校3年間を通じ課題研究の基礎から実践までを行う。高校1年においては、高校2・3年生での個人探究の基礎となる科目として、アカデミックライティング(AW)とデータサイエンス(DS)を設定した。

表1. 総合的な探究の時間 STEAM 年間計画

・アカデミックライティング (AW)	
前期	仮説検証の考え方 資料の探し方・小論文の書き方 倫理的な考え方・情報の収集方法
後期	課題の設定とその解決方法 課題の分割・クリティカルリーディング 個人テーマの設定・カウンセリング
・データサイエンス (DS)	
前期	定量的な評価の理論・演習
後期	研究計画の立て方・進め方 Problem Based Learning (PBL)

2. 授業の目的

高校1年生で実施するDSでは、問題解決のためのフレームワークであるPPDAC (Problem Plan Data Analysis Conclusion) サイクルのうち、前期は、Data (データ), Analysis (分析), Conclusion (結論) に焦点を当てた。取り扱うデータに対し、適切な統計分析方法を選択し、結果を適切に解釈する素養を身につけることを目的としている。また後期は、PBLを用いたグループ研究を行った。Problem (問題), Plan (計

画), Data (データ), に焦点を当て、問題や仮説の設定をした後に実験計画を立て、適切なデータを収集する方法を理解することを目的とした。本発表では、主にDS前期の実践について報告する。

3. 実践内容

◆ 概要

DS前期では、名古屋大学全学教育科目となる「データ科学基礎(石井秀宗, 名古屋大学教育学部教授)」の一部を活用し、数学と他教科の教員(当該年度においては体育)がTT(チームティーチング)で実践した。履修する内容は「情報I(問題解決)」と「数学I(データの分析)」の範囲を中心としている。第1~3回までは教室で行い、4回目以降は実践演習を踏まえた。

表2. DS前期計画

回	内容
1	基礎 データの種類・構造
2	理論 基礎 統計図表・代表値
3	基礎 散布度
4	演習 相関関係
5	基礎 回帰分析・時系列分析
6	基礎 仮説検定
7	演習 t 検定
8	基礎 クロス集計表・適合度検定
9	応用 データ収集・集計
10	応用 統計ポスター作成
夏	応用 統計ポスター作成・提出
11	基礎 分析・レポート作成
12	基礎 レポート作成時の注意
13	応用 分析・レポート作成2回目
14	応用 分析・レポート完成・まとめ

◆ 初年度からの変更点

本校でのDSの実践は2年目となる。初年度は統計理論を構造的に理解することでデータを正しく読み取ることを目指し、数式とあわせた解説をしたところ、数式を用いたことで、数学が得意な生徒にとっては理解が深まり、腑に落ちる様子で授業を受けていたことが印象的であった。一方、反省点として「言葉の意味を理解するのが難しかった」など、統計用語や独特な理論に抵抗感を示す生徒がいた。数学が苦手な生徒にとっては、馴染みのない統計用語や理論に加えて数式が用いられることで拒否反応を示したようである。このような生徒も演習を通して分析の楽しさに気づくことを狙ったが、十分と言える成果には至らなかった。また、「Excelの使い方がすごく難しかった」と答えた生徒がおり、操作に慣れる時間が十分でなかった可能性がある。

初年度の実践をふまえ今年度は、データ分析を楽しみながら、データの読み取り方を身につけることを目的とし、「数式抜きでのDS」を実践した。これは、「データリテラシーの基礎の習得とともに、それらのスキルの価値や必要性、また、面白さについて気が付かせる」という当初からの教材開発の目的達成に向けた方向転換である。また、新学習指導要領に則し、情報Iの授業と連携した。「問題解決」の内容を一部DSで、「Excel操作の演習」を情報で行うことでそれぞれの時間を確保した。

◆ 具体的な授業内容

教材の作成においては「生徒自身の気づき」を大切にしたい。具体例の提示→結果の読み取り→概念の大きな理解の順で伝えることを意識した。個人で見つけた「気づき」は適宜共有し、クラスでデータ分析を楽しむ授業を心がけた。

【第1回】データの種類・構造

- ・導入 データには種類があるのではないかな？
- ・展開1 次のデータを2種類に分けてみよう！
グループ分けを発表
- ・展開2 質的変数・量的変数の解説
- ・展開3 自分で質的変数や量的変数を見つけよう！

工夫した点は、変数のグループ分けをまず生徒自身が行うことで、「自分たちでグループを発見した」という気づきを持ってもらうことを意識した。

【第4回】相関分析演習

- ・導入 相関分析復習
- ・展開1 相関関係のペアを調べよう！
見つけたペアの特徴を発表
- ・展開2 散布図を作成してみよう！
相関係数と散布図からわかったことを発表

工夫した点は、データの貼付けのみで簡単に相関分析とt検定ができるExcelシートを作成し、データの読み取りや考察に時間を割けるようにした。

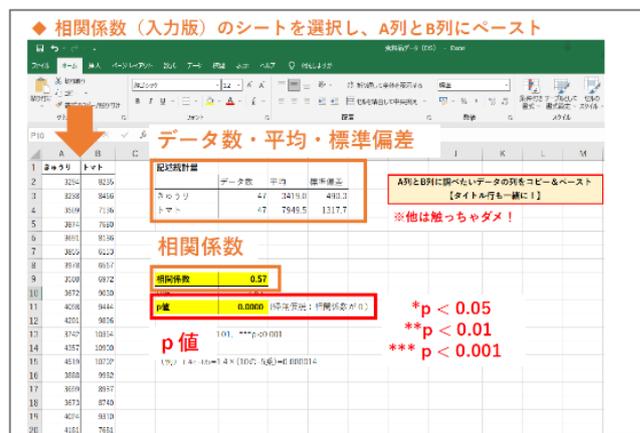


図1. 実際に使用したExcelシート

【第10回】統計ポスターを作ろう

- ・導入 ポスターの作り方の紹介
- ・展開 統計ポスターを作ろう！

成果は、第66回愛知県統計グラフコンクール（第5部、パソコン統計グラフの部）へ各10点ずつ出品した。第5部は金賞3点、銀賞4点、銅賞3点、パソコン統計グラフの部では、金賞1点、銀賞1点、銅賞1点を受賞した。

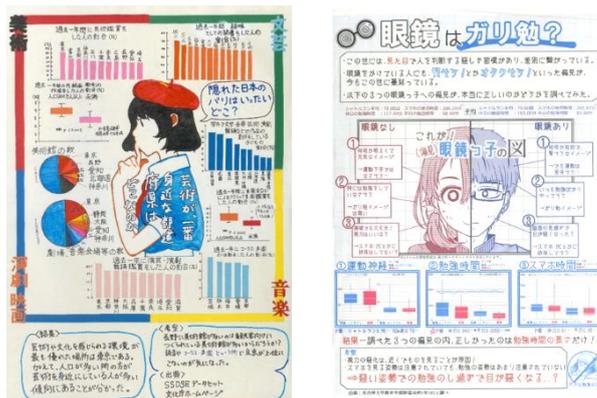


図2. 生徒の作品

【第12回】 レポート作成時の注意

- ・ 導入 見本レポートの間違い探しをしよう！
- ・ 展開1 レポートの修正点を発表
- ・ 展開2 自分のレポートの書き方を考えよう！

工夫した点は、間違い探しを通して分析結果の解釈を深めるきっかけになるようにした。さらに、自分のレポートはどう書いたら良いかを考えさせた。

◆ 評価方法

- ・ 基礎知識の確認として、100点満点の試験を実施
- ・ ポスターやレポート課題はルーブリック評価

表3. DSで使用しているルーブリック

	目標	つけたい力	
①	仮説を設定して研究計画を作成し、多様な解決法で探究し続ける力を身につける	課題分割された適切な仮説を設定する力	課題解決に向かう適切な仮説を、課題分割によって設定することができる
		扱うデータに対し適切な分析方法を導き出す力	扱うデータの特性を理解し、適切な分析方法を導くことができる
②	科学的に思考・吟味して客観的データを評価し、現象と原因の関係を説明する力をつける	科学的に思考・吟味し、客観的に評価する力	得られた結果について客観的な根拠をもった定量評価ができる
		他者へ論理的に説明する力	得られた結果の背景にある因果関係や、データ評価の根拠について論理的な説明ができる
③	多様な価値観を持つ他者と協同して思考を深め、課題の本質に迫る力をつける	多様な価値観を持つ他者と協同して思考を深める力	他者の研究を理解し論理的な助言をすることができ、相手の意見を聞き入れるなど、自身の思考の枠組みを柔軟に変化させることができる
		知識を結びつけて課題の本質を理解する力	得られた結果について、先行研究や他者の意見など新たに得た知識と既存知識を結びつけ、課題の本質に迫ることができる
④	これまで学んだ知識を結びつけて、新たな価値を生み出す力をつける	新たな価値を生み出す力	データを可視化や統計分析によって論理的に解釈し、新たな付加価値を生み出すことができる

4. 成果と課題

2022年4月と9月に実施したアンケート結果は次の通りである。アンケート項目はすべて、「⑤よくあてはまる ④あてはまる ③どちらともいえない ②あまりあてはまらない①あてはまらない」とした。

表4. アンケート結果 (2022年)

質問項目	4月	9月	
現在「数学」は好きですか	3.40	3.93	***
現在「理科」は好きですか	3.24	3.73	**
現在「統計」「データ分析」は好きですか	2.83	3.23	*
現在「統計」「データ分析」は得意ですか	2.54	2.73	
グラフや表を適切に読むことが出来ますか	3.43	3.23	
実験や観察、アンケートなど自分でデータを収集することが出来ますか	3.06	3.26	
データの種類からどんな分析方法が適切かを判断できますか	2.85	3.07	
先行研究などを調査して、課題を発見することが出来ますか	2.83	3.17	*
データを加工・整理することが出来ますか	2.88	3.27	**
データ解析から結論を導くことが出来ますか	2.97	3.33	**
複数のデータを組み合わせることで結論を導くことが出来ますか	3.09	3.29	
自分の研究を客観的に他者に説明することが出来ますか	3.26	3.41	
他の人の研究にアドバイスを出来ますか	2.88	3.03	
他の人の意見を柔軟に取り入れて、考えを修正することが出来ますか	3.59	3.86	*
あなた自身にとって、DSの授業はどの程度必要性を感じますか	2.87	4.14	***

(n=120, 有意水準5%未満, * $p<.05$, ** $p<.01$, *** $p<.001$)

【授業の感想 (自由記述)】

- ・ 様々な場面に使える統計の技術を学び、とても有意義な時間になりました。
- ・ DSの授業を受けるようになってから、普段目にするデータなどをよく観察するようになりました。
- ・ 難しかったけど楽しかったです！！
- ・ テストの個票の標準偏差が理解できたので、自分の点数が全体でどのくらいの位置にいるのかが分かるようになった！

前期の実践を通じて、数学や統計・データ分析が好きであるという生徒が有意に増加した。また、データを加工整理し、結論を導くことが出来るようになった

と感じる生徒も有意に増加した。これは昨年度の実践では見られなかった傾向である。

また、2022年度の授業の感想 ($n=67$) についてテキストマイニングを用いて分析したところ、最も回数が多い形容詞は「楽しい」で13語、次は「難しい」で11語であった。共起キーワードとしては「楽しい」と「難しい」が関連しており、「難しかったけど、楽しかった」というコメントが多いと考えられる。2021年度の授業の感想 ($n=57$) で最も回数が多いのは「難しい」で26語、次は「楽しい」で12語であった。ここからも、今年度の実践ではより「データ分析の楽しさ」を伝えることができたと予想される。

一方で、生徒が正しく結果を読み取ることが出来ているかについてはまだ練習が必要だと感じている。DS後期では実際に自分たちでデータを取得し、その結果を班で共有しているが、その声から「 p 値が小さければ説得力が増す」「対応がある・ないデータの区別」などまだ理解せずに使っている様子が伺える。馴染みがない統計用語であるので、正しく使うためには繰り返し学習することが必要であると考えた。

そこで、今年度は前期DSの授業内容の補足教材として、石井の監修のもと、「THE FIRST STEP DATA SCIENCE」というテキストを作成した。生徒が自分で予習・復習をしやすいように、また、高2・3年生での課題研究の際に利用することを目的とした。次年度はこのテキストを用いて授業を行う予定である。

5. 参考資料・参考文献

- 1) 名古屋大学全学教育科目「データ科学基礎」スライド
- 2) 石井秀宗. (2014). 人間科学のための統計分析. 医歯薬出版株式会社.
- 3) 酒井淳平, & 稲葉芳成. (2018). 高等学校における「データの分析」その後の統計教育実践の一事例. 統計数理, 66(1), 135-151.
- 4) 青山和裕. (2018). ニュージーランドの統計教育から得られる示唆—カリキュラム, 教材・授業事例, 評価制度の観点から—. 日本数学教育学会誌, 100(7), 11.
- 5) Tobías-Lara, M. G., & Gómez-Blancarte, A. L. (2019). Assessment of informal and formal inferential reasoning: A critical research review. *Statistics Education Research Journal*, 18(1), 8-25.
- 6) Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8(1), 82-105.
- 7) 伊川美保, & 楠見孝. (2020). 統計リテラシー自己効力感尺度日本語版の作成—統計教育の効果測定—. 心理学研究, 91(2), 133-141.
- 8) 阿部真人. (2021). データ分析に必須の知識・考え方 統計学入門. ソシム株式会社.
- 9) 江崎貴裕. (2021). 分析者のためのデータ解釈学入門. ソシム株式会社.
- 10) 理数探究基礎. (2020). 数研出版株式会社.
- 11) 理数探究基礎 未来に向かって. (2020). 新興出版社啓林館.
- 12) 岡本尚也. (2021). 課題研究メソッド 2nd Edition. 啓林館.
- 13) 後藤芳文, 伊藤史織, & 登本洋子. (2015). 学びの技 14歳からの探究・論文プレゼンテーション. 玉川大学出版部.
- 14) 涌井良幸. (2015). 統計カクイズ. 実務教育出版.
- 15) 浦上昌則, & 脇田貴文. (2020). 心理学・社会科学研究のための調査系論文の読み方 改訂版. 東京図書株式会社.
- 16) 石村光資郎, & 石村友二郎. (2014). 卒論・修論のためのアンケート調査と統計処理. 東京図書株式会社.

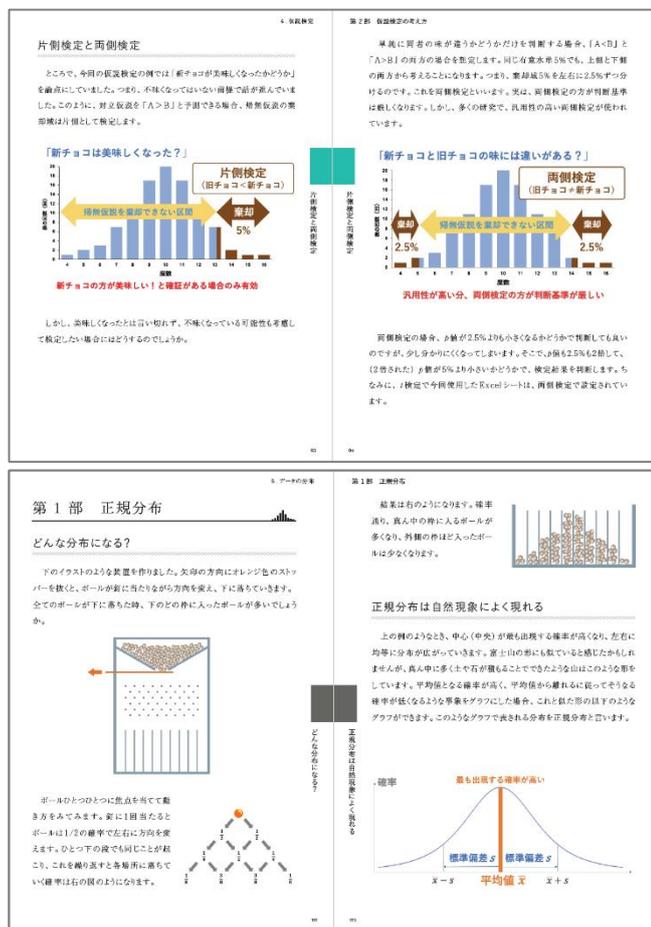


図3. 「THE FIRST STEP DATA SCIENCE」

「仮説検定」への接続を目指した「仮説検定の考え方」の教材開発とその実践

光永 文彦 西大和学園中学校・高等学校・大和大学
大浦 弘樹 東京理科大学
吉川 遼 名古屋文理大学
伏木田 稚子 東京都立大学

1. 研究背景

2022 年度より施行された現行の高等学校学習指導要領(文部科学省 2018)では、約 30 年ぶりに数学 B「統計的な推測」単元に「仮説検定」が復活し、数学 I に「仮説検定の考え方」が新設された。特に「仮説検定の考え方」においては、「具体的な事象における仮説検定の考え方を理解すること」(文部科学省 2018)とされ、主張の妥当性について、実験などを通して判断することや、批判的に考察することができるようになることを求めている。ここでいう「妥当性」とは、ある標本比率からその母比率に関する仮説を立て、二項分布を基にした実験を行い、結果の相対度数の分布を提示した上でその結果が起こりやすいか否かの判断基準として、実験結果の平均値から標準偏差いくつ分離れているのかという判断のための「客観性」を意味しており、これにより「無作為性だけでは判断できる可能性が限りなく低いこと」から「起こりえないこと」としての判断の根拠を指している。例えば丸山(2019・2021)はこのような確率分布を経ない「実験結果をもとに確率的に推測する活動」を、数学的モデルに基づく形式的な統計的手法を介さない統計的推測という意味で「インフォーマルな方法による統計的推測」と表現し、標本変動性に着目した教材を開発し、その効果を検証している。

一方で、「仮説検定の考え方」導入にあたって「実際的な場面を考慮し、具体例を通して「仮説検定の考え方」を直観的に捉えさせるようにした」(文部科学省 2018)ことがあるにしても、この「考え方」の先には仮説検定における「帰無仮説」があり、「判断の基準」の先には仮説検定における「有意水準」がある。例えば 1982 年度施行学習指導要領「確率・統計」では「危険率」とも表記され、「帰無仮説が真であるときに誤って棄却する(第 1 種の誤り)確率」を指したが、現行の学習指導要領 数学 B の「仮説検定」においても同様の内容が求められているため、「仮説検定の考え方」をどのように「仮説検定」に接続していくかが、現行の学習指導要領では非常に重要となる。

2. 「仮説検定の考え方」と「仮説検定」の前提

前述の通り、数学 B「仮説検定」を学習はその直前に学習する二項分布や正規分布といった確率分布の内容が前提となるが、数学 I「仮説検定の考え方」においてはどのような知識が前提となるかについては学校によって状況が異なる。現行の学習指導要領において、一般的に高等学校 1 年次に「数学 I」(標準 3 単位・必修)と数学 A(標準 2 単位・選択)が設置され通年で並行履修を行うことが多い。

【 数学 I (必修) 】

- データの分析
分散, 標準偏差, 散布図, 相関係数, 分散, 標準偏差
具体的な事象の仮説検定(仮説検定の考え方)

【 数学 A (選択) 】

- 場合の数と確率
集合, 確率
- 数学と人間の活動
論理的確率, 頻度確率

【 数学 B (選択) 】

- 統計的な推測
標本調査, 確率変数, 確率分布, 二項分布, 正規分布
正規分布を用いた区間推定, 仮説検定
- 数学と社会生活
事象の数学化, 数理的問題解決

そのため、数学 A「場合の数と確率」の学習を前提とした組み立てをすべての学校で行うことはできない。つまり、中学校数学における確率の学習内容が前提となる。高等学校学習指導要領解説(文部科学省 2018)ではこの前提を「問題を解決するために、この結果が偶然に起こりえた可能性はどのくらいあるのかを、コイン等を使った実験を多数回繰り返して考察する」と表現している。本研究では、この中の「偶然」に着目していく(なお、文部科学省 2018 においてはこの部分しか「偶然」の記述はない)。

中学校学習指導要領解説(文部科学省 2017)では「偶然」の記述が全部で4か所ある(強調は著者によるもの)。

中学校数学科の内容の骨子(P.39)

④不確定な事象

数学で考察する対象は多様であり、確定した事象だけではなく、集団においてばらつきのある事象や偶然に左右される事象、全体を把握することが困難な事象など不確定な事象も考察の対象とする。

指導内容の概観(P.55)

ア データの分布と確率についての基礎的な概念や性質を理解すること

この領域においては、集団においてばらつきのある事象や偶然に左右される事象、全体を把握することが困難な事象など不確定な事象も数学の考察の対象であることを理解することが大切である。その上で、不確定な事象のばらつきを表すために、データの分布や確率についての基礎的な概念や性質を理解することが大切である。

イ データを収集して分析したり、確率を求めたりできるようにすること

偶然に左右される事象は、確率を用いて考察することができ、全体を把握することが困難な事象は、確率の性質を用いて標本調査を行い、その結果を考察することができる。

第1学年の目標及び内容(P.93)

不確定な事象の起こりやすさの傾向を読み取り表現すること

日常生活や社会において、偶然に左右される不確定な事象は数多くある。多数の観察や多数回の試行の結果を基にすることにより、不確定な事象の起こりやすさの傾向を読み取ることができる。多数の観察や多数回の試行の結果を基に不確定な事象について考察する際には、相対度数を確率とみなして用いることが考えられる。

これらの記述からもわかる通り、中学校において相対度数を確率としてみなす頻度確率が「仮説検定の考え方」の前提となっていることがわかる。

では、生徒は「偶然に左右される不確定な事象の確率」をどのように理解しているのだろうか。仮説検定の考え方の授業前に予備調査を行い、次の実践を計画した。

3. 授業実践

【対象】 高校1年生3クラス(各47名 計141名)

【単元】 数学I「仮説検定の考え方」

【形態】 普通教室における一斉授業(個人・班別)

【日時】 2023年2月25日:予備調査

2月28日:実践① ゲーム実践

3月13日:実践② グループワーク

【教材】 数学I教科書, Chromebook, 統計ゲーム

【予備調査(一部抜粋)】

- ・1枚のコインを投げたところ、表面が上になりました。この結果は偶然といえますか。
- ・1枚のコインを投げて表面が上になるか、裏面が上になるかを記録することを30回繰り返したとき、18回表面が上になりました。この結果は偶然といえますか。
- ・1枚のコインを投げて表面が上になるか、裏面が上になるかを記録することを30回繰り返すことを10セット行ったとき、10セット連続でちょうど18回表面が上になりました。この結果は偶然といえますか。

【ゲーム教材】

仮説検定の考え方をういてどのようにして「妥当性」を判断するのかが感得しやすいノベライズ型統計ゲームを開発し、実際にプレイしてワークシートにまとめている。

4. まとめ

現在、授業実践を行っており、今回は仮説検定の考え方を学習する前の生徒の理解度と開発した教材について、当日に資料とともに発表する予定である。

参考文献

- [1] 丸山達法(2019) 高校数学の「仮説検定の考え方」におけるより良い推測のための標本の変動性に関する一考察, 数学教育学会誌, 60(1・2), 61-70.
- [2] 丸山達法(2021) 高校数学「仮説検定の考え方」における標本変動性に関する研究, 数学教育学会誌, 62(1・2), 1-19.
- [3] 文部科学省(2017)【数学編】中学校学習指導要領(平成29年告示)解説
- [4] 文部科学省(2018)【数学編・理数編】高等学校学習指導要領(平成30年告示)解説

付記

本研究は、科学研究費助成事業 課題番号18K02815・21K20277の助成を受けて行った。

高等学校における仮説検定にかかわる指導について

—統計学習と探究活動の迫間をどう埋めるか—

立命館宇治中学校・高等学校 数学科 稲葉芳成

〒611-0031 京都府宇治市広野町八軒屋谷 33-1

Email : inava.y2011@gmail.com

1. はじめに

2021 年度の本ワークショップにて「高等学校における仮説検定の扱いにかかわる期待と懸念」と題して発表を行った。数学 I 「データの分析」における「仮説検定の考え方」の扱いが開始され、2022 年度からは数学 B の「統計的な推測」が扱われている。新しい数学 I の教科書と同様、数学 B の教科書においても「統計的な推測」にかかわる記述に各社の個性が見られる。また一部には、記述の正確さにやや欠けるものも見られ、実際の指導では留意を要するものもある。学習指導要領の改定に際しては、情報科におけるデータサイエンスの扱いの中で統計的な手法が紹介されるなど、教科での教授内容の重複も見られる。さらには「総合的な探究の時間」などでの探究活動の実践上、統計リテラシーの向上への期待が高まる情勢下にある。一方で統計学習は探究活動にかかわる能力の向上に直結するものではない。ここでは、新学習指導要領下での統計教育と探究活動のかかわりについて課題意識の一端を記す。

2. 新学習指導要領下での教育課程の一例

学習指導要領の改定に伴い、数学 B、数学 C の再編があった。また情報科とのかかわりも無視できない。統計教育の充実を実現するために筆者の勤務校では数学科で議論を重ねた結果下表のような教育課程を採用している。勤務校には、国際バカロレアの IB コース、1 年間の留学を含む IM コース、そして一般的な IG コースがある。表は IG コースのものである。

年次	文系	理系
1 年		数学 I③ 数学 A②
2 年		数学 II④ 数学 C③ [選択*]
3 年	数学 B② 統計学③ [選択**]	数学 III④ 数学 B② 統計学③ [選択**]

* : ** : 文系・理系共通の選択科目

第 2 学年まではゆるい文理選択という感じとなり、将来の文系理系を問わず第 2 学年で数学 C を選択することができる。数学 B は全員が履修して卒業することを考え、あえ

て第 3 学年に設置している。第 3 学年の「統計学」は 3 単位の学校設定科目である。数学 B の内容とは一部で重複するが大部での重複を避ける工夫として、2 標本の検定など統計学の一般的な話題を扱う。

3. 学校設定科目「統計学」における取り組み

高校生として身につけておくべき学習指導要領の範囲を超える知識については時代の要請によるところが大きいと思われる。近年では探究的な活動が注視され、エビデンスベースでの議論が重視されている。統計学がそこに果たす役割は大きいと思われる一方で、データをとりあえず処理すれば結果が出力されるという、ブラックボックス統計の弊害は避けている。近年は、①仮説検定は Z 検定を超えて t 検定、カイニ乗検定（適合度検定）を扱う。ただし F 検定は扱わない。②線形近似または多項式近似を扱う③2X2 の分割表及び関連する疫学関連事項を扱う④Python プログラミングを扱う⑤学術論文に現れる統計手法とその理解について扱う⑥探究活動に結びつくデータ収集と分析をセットとした課題学習を扱う、などの方向にある。

4. 統計を学ぶこと・教えること、「問い」を解決することの難しさ

教科書でも扱われるようになった仮説検定も含めた統計手法の理解や実際にデータを扱う経験によって、道具として統計的な手法・技法を使うスキルをある程度身につけることができる。これは生徒に課したレポートの内容などから評価できるものである。しかしそれだけで、統計的な手法・技法の使い方に慣れることは直結しない。教科書の例題の理解や類題の練習問題に取り組むことは難しくない。あたかも料理の材料もレシピも用意された状況から料理をつくるようなものである。翻って、後者は、目の前の課題に対して、実際にどんなデータを集めて、どのように手法を用いることで統計的に何が言えそうか、というデザインの上で立って実際の分析を行った結果として何らかの成果が得られるものである。この点では、「このような料理をつくる」という目標だけが与えられた段階で、どのような材料を用意しどのようなプロセスを経て何を料理するかは自分で考える、というようなものである。

このように統計を学び統計リテラシーが向上することと探究リテラシーが向上することは必ずしも一致しない。探究リテラシーに結びつく統計リテラシーは意識的かつ系統的に学ぶ・教える必要がある。そのためには①実際の探究の場面での統計的な手法の用例を見ること・見せることや、②探究のプロセスを意識した課題学習を行い。それを実際に体験すること、の2つが必要であろう。

4-2 「統計の型と流儀」と「定期考査における採点」

ひとつの問い「ある硬貨を1枚続けて120回投げたところ、表が70回出た。この硬貨は表が出やすいように歪んだ硬貨であると言えるか」に対する回答としてどのようなものが許容されるであろうか。①素朴に表と裏の回数が $70 - 50 = 20$ 回もあるという事実から、②表の出る回数 X について二項分布 $B(120, 1/2)$ の $P(X \geq 70) = 0.0412$ から、③二項分布 $B(120, 1/2)$ の正規分布近似 $N(60, 30)$ の $P(X \geq 70) = 0.034$ から、④標本比率 $70/120$ 、母比率 $1/2$ としたときの母比率の有意水準5%の片側検定での検定統計量 $Z = 1.85$ から、⑤あるいは同有意水準5%の両側検定から、などいくつかの方法が考えられるが、そのどれが正しく、誤りであるのかについて、感覚的に正しい、統計的に正しいという判断、という枠組みで考えることは容易であろう。一方で、ではなぜ有意水準は5%でなければならないか、片側検定であるべきところを両側検定では不味いのか、帰無仮説は明示する必要があるか、などの疑問に立ち返るとき、統計の型を教え(学び)、統計の流儀で綴ることの意義は、即ち定期考査におけるこの1問の解答に対する採点の結果に等しい。

4-3 「統計の型と流儀」の学習と現実の場面のギャップ

筆者が過去(2017)に看護学生を対象に、統計的検定の基礎までを終えた学生が正しく問題解決ができるかどうかについて行った質問紙による調査では、学習した6つの選択肢〔①母平均の推定②母比率の推定③母平均の検定④母比率の検定⑤適合度の検定⑥t検定による母平均の検定〕のうち適当なものを選ぶものであった。問い「ある町で新製品のモニター調査をしたところ80名中54名の「良い」評価を得た。この町に住む約3万人の評価がどれほどか知りたいとき必要なのは?」に母比率の推定と回答したのは38名中11名であった。この学生たちがこの調査の近後の考査で出題して仮説検定のある問いの正答率は74%であったが、この問いには「平均寿命が1500時間であるという主張は正しいと言えるか。有意水準5%で検定しなさい。」という指示文が含まれていた。

4-4 探究的な問いの解決と統計的な型の習熟

統計的な探究プロセスについては総務省が2016年に刊行

した「生徒のための統計活用～基礎編～」の中で、「身の回りの現象への気づきや疑問を問題としてとらえ、統計を通して解決するこの一連のプロセスを統計的な探究プロセスと言います。」との言葉に続き「統計的な探究プロセスは以下のステップで構成されます。」として[テーマ][とらえる][計画する][集める][まとめる][読み取る][活かす]としている。続く高校生向きに書かれたものではよく知られたPPDACサイクルが冒頭から紹介されている。これらは統計学習(教育)と探究を結ぶ貴重な資料である。

教科書と「探究」を結ぶ取り組みのひとつは統計が使われている現場を知ることである。実際に統計的な分析が行われている学術論文[3]を紹介し、何が何のためにどのように使われているかを知ること、統計の使い方を知ることができる。教科書と「探究」を結ぶ取り組みのもうひとつは限定的な課題学習である。PPDACサイクルを実践するには準備と時間が必要であるが、生徒の興味関心に応じた単純な疑問を「問い」として、その問いが正しいと言えるか否かを統計的な手法によって結論づける。単純な「問い・データ・分析・結論」のQDAC過程である。これは従前からよく取り組まれている事例でもあろうが、自分で分析に必要なデータを集めることや、いくつかの手法の中から適した手法を選ぶことなどを通じて一般的な探究リテラシーの向上に資することを目的にする。勤務校での先の「統計学」では毎年レポート課題としてこのような課題に取り組ませている。「食品の内容量の表示が正しいかを母平均の検定によって確かめる」などの様々なレポートが提出されるが、そのプロセスを通じて、誤った認識が修正されたり、実際にデータを集める際の気づきを体験する。

5. まとめ

指導要領の改訂により統計教育の内容は充実したが、それを利活用して探究活動に活かす能力の向上のためにはそれを目的とした訓練が必要である。教科書の理解だけでは必ずしも探究リテラシーの向上には結びつかないという意識を持った指導が求められる。

参考文献

- [1]総務省政策統括官「生徒のための統計活用～基礎編～」2016
- [2]総務省政策統括官「高校からの統計・データサイエンス活用～上級編～」2017
https://www.soumu.go.jp/toukei_toukatsu/info/guide/stkankyo.htm
- [3]金衿佳, 森川慧子ら「遠隔授業と対面授業, その課題と可能性」Asphodel, 56, pp.77-107(2021), 同志社女子大学

ICT を活用した推測統計の指導とその評価

橋本三嗣・広島大学附属中・高等学校
〒734-0005 広島市南区翠一丁目1番1号
TEL.082-251-0192 FAX.082-252-0725
E-mail : mhashimo@hiroshima-u.ac.jp

1. はじめに

コロナウイルス感染拡大を契機に、教育界においても ICT を活用した指導の需要が増大した。臨時休校が続いた時期には、学校の授業をどのようにして生徒に届けるかに教員の関心が向く傾向が見られたが、結果として従来の指導カリキュラムの見直しの契機となる実践知が集まった。また中学校、高等学校の学習指導要領に各教科で育成すべき資質・能力が明確に記述され、限られた授業時間の中で、どのように生徒の学びを実現させるかという工夫が、教員に求められている。

ICT を活用した指導は小学校から進められており、最近では PC 等の操作に慣れた生徒たちが中学校、高等学校に入学してくる状況である。文部科学省の進める GIGA スクール構想の実現に向けて、各学校に Wi-Fi 環境も整いつつあり、生徒一人1台の端末使用が可能になっている。学校現場では、ICT をどう活用するかが課題になっている。

今年度から実施の高等学校学習指導要領では、数学科、情報科に統計内容の指導が期待されており、ICT を活用した統計授業の展開は、今後の教育の在り方を考える上で重要であるといえる。

本稿は、スーパーサイエンスハイスクール事業による学校設定科目「GS 社会と統計」の授業実践から、ICT を活用した推測統計の指導とその評価について報告する。

2. 学校設定科目「GS 社会と統計」(高2:1単位)

(1) 目的

統計的な見方や考え方をいながら ICT を適切に活用することを通して、主体的な問題解決のために必要な資質と能力を育成する。

(2) 目標

- ①統計的な見方・考え方、概念を理解することで、データやデータの分析を批判的に思考する力を育成する。

- ②社会科学における具体的な実例を通して統計的にデータを処理する方法を学び、目的に適合する方法を課題研究に活用することができる。

(3) 評価の方法

レポート、授業中の学習状況等を総合的に判断し、評価する。

(4) 単元・学習内容・時数

単元	学習内容	時数
「統計」の見方・考え方	1. いろいろな「統計」 2. 相関 3. 回帰直線・回帰分析とその利用	4
確率分布とその性質	1. 確率変数と確率分布 2. 二項分布と正規分布 3. 分布表の使い方	4
標本と標本分布	1. 母集団と標本 2. 標本平均の分布	6
推定の考え方	1. 大数の法則と中心極限定理 2. 点推定と区間推定 3. 信頼区間の作成と意味 4. 母平均の区間推定 5. t 分布による区間推定 6. 母比率の区間推定	8
検定の考え方	1. 検定の概要 2. 仮説の設定と検定 3. 標準正規分布や t 分布を利用した検定 4. 平均値の差の検定 5. χ^2 分布を利用した適合度、独立性の検定	8
社会における統計の利用	1. いろいろな統計調査 2. 課題研究への活用 3. まとめ	5

3. 指導の実際（第31時～第35時）

単元「社会における統計の利用」にて、「センサス@スクール」を利用した課題レポートづくりの活動を実施した。

- (1) 実施時期：令和5年1月～3月
- (2) 対象：高校2年生42名
- (3) 授業計画

- ①第31時：課題の提示，データの測定
- ②第32時：注目するデータの決定・分析
- ③第33時：課題レポートの作成
- ④第34時：生徒間の相互評価と自己評価
- ⑤第35時：課題レポートの推敲

(4) 課題

「センサス@スクールのデータを活用して、『社会と統計』の授業で学習した χ^2 検定および平均値の差の検定を用いて分析せよ。」

4. 課題レポート（高2生徒が作成）【第33時】

学校設定科目『社会と統計』課題レポート

「センサス@スクールの利用」

II年1組 () 番名前 ()

1. 課題
「センサス@スクールのデータを活用して、『社会と統計』の授業で学習した χ^2 検定および平均値の差の検定を用いて分析せよ。」
2. 調べたいこと・注目したデータ
起きた時刻、または通学時間により朝食の有無には関係について調べた。
起きた時刻が早い、または通学時間が長い人ほど朝食を食べていない人が多いのではないかと仮定した。

3. 分析結果

(表1) 抽出したデータ

Q13.起きた時刻	Q24.通学時間(分)	朝食の有無	度数
-5:00	15	有	4.5
-5:00	30	無	4.5
-5:00	30	有	4.5
5:00	90	有	5
6:00	30	有	6
6:00	50	有	6
6:30	30	有	6.5
6:30	15	有	6.5
6:45	5	有	6.75
6:45	20	有	6.75
7:00	90	有	7
7:00	30	無	7
7:00	30	有	7
7:00	50	有	7
7:15	15	有	7.25
7:30	5	無	7.5
7:30	40	有	7.5
8:30	30	有	8.5
9:00	45	有	9
9:30	20	有	9.5

(表2) 観測度数、期待度数

観測度数	朝食あり			朝食なし			合計	
	観測	期待	割合	観測	期待	割合		
4.5	2	1	0.3	3	4.5	1.75	0.1	1.85
4.75	0	0	0	0	4.75	1.75	0.2	1.85
5	1	0	1	1	5	1.5	0.1	1.6
5.25	0	0	0	0	5.25	1	0.25	1.25
5.5	0	0	0	0	5.5	1	0.25	1.25
5.75	0	0	0	0	5.75	1	0.25	1.25
6	2	0	2	2	6	0.75	0.5	1.25
6.25	0	0	0	0	6.25	0.5	0.5	1
6.5	2	0	2	2	6.5	0.5	0.75	1.25
6.75	2	0	2	2	6.75	0.25	0.75	1
7	3	1	4	4	7	0.25	1	1.25
7.25	1	0	1	1	7.25	0.1	1	1.1
7.5	1	1	2	2	7.5	0.1	1	1.1
合計	14	3	17	合計	10.45	6.55	17	

- ・「起きた時刻」「通学時間」「朝食」の3つに注目したが、データを表に起こしたところ、通学時間にはばらつきがあり、関係が得られそうにないと判断した。また、今回は朝食の内容までは焦点を当てる必要を感じなかったため、「起きた時刻」と「朝食の有無」に絞って調査した。(表1)
- ・縦軸を「起きた時刻」とし、5時以前のデータは4:30に統一した。また、4:55分刻みで縦軸を設定した。横軸を朝食の有無とした。今回は学校に行く日の生活に絞って調査したかったため、本校の登校時刻8:30に絶対に間に合わない8:30以降に起きたとするデータは除外した。(表2)

(表3) χ^2 値

χ^2 値	朝食あり	朝食なし	合計
4.5	0.03571429	8.1	8.13571429
4.75	1.75	0.1	1.85
5	0.16666667	0.1	0.26666667
5.25	1	0.25	1.25
5.5	1	0.25	1.25
5.75	1	0.25	1.25
6	2.08333333	0.5	2.58333333
6.25	0.5	0.5	1
6.5	4.5	0.75	5.25
6.75	12.25	0.75	13
7	56.25	0	56.25
7.25	8.1	0	8.1
7.5	0.1	1	1.1
合計	88.7357143	12.55	101.2857143

(表4) 結果

自由度	12
有意水準0.05の境界値	21.02606982
観測度数のような確率	0.000000000001

4. 考察

- ・起きた時刻と朝食の有無には有意な関係はみられなかった。
- ・しかし、分析の仕方に反省点が多いため正確な結果とは言い難い。具体的には、
 - i. サンプル数がかなり少なかったため期待度数の設定に難渋した。
 - ii. 朝食の有無に「起きた時刻」と「通学時間」のどちらがより関係するか分からず、とりあえず両方のデータを抽出してしまっただけでデータ数の少なさに響いてしまったかもしれない。
 - iii. 度数によって横軸の値に傾斜をつくってしまったことが χ^2 検定を用いる際には不適切であった可能性が高い。
 - iv. 朝食を登校中や登校後に食べる人が存在する想定ができていなかったという点である。
- 5. 今後調べてみたいこと
 - ・多いデータ数でかつきちんと仮説を立てて調べて直してみたい。
 - ・朝食を食べるタイミングや登校手段なども考慮した、詳しいデータを用いてやりたい。

5. リフレクションシート【第34時】

「社会と統計」課題レポート

令和5年2月28日(火)

リフレクションシート

II年()組 ()番名前()

1. 他人の人の課題レポートを読んで、「よいところ」「工夫すべきところ」を書きなさい。

出席番号	よいところ	工夫すべきところ
1	好きなスポーツに着目したのがおもしろいと思いました。	観測度数の単位は(人)だと思いますが、それをつけたらいいと思います。
2	1つの仮説を検証して、そこから更に発展させて読んでくれたのがいいと思います。	27日の検証の「Z値」や有意水準を載せてほしいです。
3	私が調査したテーマの続きの内容なので、この結果が気になります。	頑張って検証して下さい。

2. 選出された課題レポートを読んで、「よいところ」「工夫すべきところ」を書きなさい。

出席番号	よいところ	工夫すべきところ
1	細かくて表にしたのが分かりやすい。調査が細かくて簡潔でよくできている。	表2に単位があつたらいいと思います。
2	不要な情報がなく、洗練されている。	調査した値の単位が欲しいです。
3	考察が深くて良い。テーマを最初に書いてくれるのが理解の助けになりました。	棄却息→棄却域だと思います。

3. 自分の課題レポートをよりよいものにするために、何をどう修正すればよいと思いますか。

・今回の「身長と内腕の長さの差」を検証するだけでなく、そこから1つの仮説を立てて実行してみる。
・字体を整え、表の単位をちゃんと揃える。
・考察をもっと深める。

6. 生徒の感想

- ①第31時：課題の提示，データの測定
 - ・自分で選んだデータを使うことで興味がわいた。
- ②第32時：注目するデータの決定・分析
 - ・不十分なデータを削除したらデータ数が少なくなった。自分でデータを選んでどのように分析したらいいのかを考えるのが難しかった。
- ③第33時：課題レポートの作成
 - ・私たちの身近なことについても、『社会と統計』で習った検定を使って分析できると分かったので、興味のあることで分析可能なものは、検定を行ってみたい。

参考文献

- (1) 熊原啓作, 渡辺美智子 (2012) 『身近な統計』, NHK 出版。
- (2) 山田剛史, 金森保智 (2022) 『エピソードで学ぶ統計リテラシー-高校から大学, 社会へとつながるデータサイエンス入門』, 北大路書房。

情報 I におけるデータサイエンス教育の前倒し実践と今後の改善に向けて

荒川智浩・国際基督教大学高等学校

〒184-8503 東京都小金井市東町 1-1-1

TEL:0422-33-3401 FAX:0422-33-3376

Email t.arakawa.m@icu-h.ed.jp

1 はじめに

1.1 所属高校の紹介

国際基督教大学高等学校(以下、本校)は東京都にある私立高校であり、3 学年合わせた生徒数は約 720 名、そのうち帰国生徒は約 480 名で 2/3 を占めているところ。

本校は以前 SGH(スーパーグローバルハイスクール)に指定されていたことはあり、帰国生徒が多いといった特色はある一方、データサイエンス教育に時間を割いている(力を入れている)ことはなく、ごく一般的な高校である。

1.2 背景

高等学校(以下、高校)における統計・データサイエンスに関する授業実践については、SSH(スーパーサイエンスハイスクール)等の学校を中心に広がりを見せているところと認識している。

そして 2022 年度からの新カリキュラム移行に伴い、高校の「情報 I」においてデータサイエンス分野が本格的に学習分野として導入されたことにより、これまで授業実践経験のない高校等においても実践していかねばならない状況であり、本校も同様であった。

なお、本校ではこれまで「情報と社会」を 3 年生で週 2 コマ実施していたところを、「情報 I」では 1 年生で週 2 コマ実施としている。

本稿 2 以降に記す授業実践を行った動機は、本校のように、データサイエンス学習に関して、ごく一般的な高校(データサイエンス分野について、これまでほとんど授業をしてきたことがない、学校全体でまだ理解が得られていない、授業を行うマンパワーが足りない)高校において、横展開でき、一助となる授業事例を実現したいと考えたところにある。

情報 I において、データサイエンスが取り上げられるようになってはきたものの、それを究めてきている教員はまだ少ないと思われること、ハードルの低い参入事例が増える

ことにより、取り組もうとする情報科等の教員が少しでも増え、データサイエンス学習の裾野が拡大していくことを祈念し、本授業を計画・実践することとした。

2 情報 I での授業実践

2.1 授業計画

データサイエンス分野については、次の 2 点を重視し、本校では 1 年生の 1 学期(主に夏期休暇前の 6 月)に実施することとした。

・情報 I の他分野と比較して、他教科や学校生活にも利活用されうる分野であること。

・アウトプットを外部コンクール(本稿 5 参照)に出すことで、生徒に学習で培ったデータ分析技術を多少なりとも使用させる方が効果的ではないかと考えたが、その多くは夏期休暇中の実施が多かったこと。

特に 2 点目のアウトプットを外部コンクールに出すことについては、コンクールに関する情報共有や、必要な力を伝えるところまでができれば、教員としての負担は最小限で進められるとともに、生徒もチャレンジする場があることで得るものも大きいと考えたからである。

以上、上記を踏まえ、授業を次のとおり計画・実践した。

表計算ソフトを使用した経験のある生徒がほとんどいないことから、電子教材(ライフイズテック社)を用いて、ゆっくり進めること、演習時間を多めにとることを重視した。

【基礎授業(全 7 回:授業・演習 各 25 分程度)】

第 1 回 データの見方

第 2 回 データの集計(合計)

第 3 回 データの種類(データの尺度水準)

第 4 回 データの可視化(グラフ)

第 5 回 データの分析(統計量)

第 6 回 データの散らばり(分散・偏差値)

第 7 回 データの相関

【夏課題説明&準備(全体説明&グループ別作業)】

第 8 回 各種コンクール等の説明・検討

第 9 回 コンクールの決定、資料準備等の注意

第 10 回 夏作業の検討、ポスター作成等の注意

【夏期休暇以降】

生徒は事前にグループ内で確認したスケジュールをもとに、各種コンクールにチャレンジしていくことになった。

また、夏期休暇中は、オフィスアワーを事前にアナウンスし、zoom も含めて対応することとした。

そして、学校には 9 月早々に各コンクールに提出した媒体(締め切りがまだのコンクールについてはひとまず完成させた媒体)をコンクールの定めに応じて、紙あるいは電子媒体で提出させ、提出状況を見て、相互評価(コメント)を実施することとした。

またデータサイエンスに関する成績評価は、1 学期に関しては、授業で課した提出物の状況と定期試験、2 学期以降に関しては、外部コンクールの作品提出有無(もし内容を評価された場合には加点)と相互コメントを実施しているかどうか、ということで成績に反映した。

なお、生徒の ICT 環境を紹介しておく、BYOD(Bring Your Own Device)を基本とし、Google 環境が使用できれば 2022 年度はパソコンでもタブレットでも可としているが、貸し出し用 chromebook も授業ごとに用意し、必要に応じて、貸し出しを行っている。

2. 2 実践前に想定された効果と懸念

効果としては、以下の点を想定していた。

・今後の他教科への応用機会提供。(理科科目における実験データをまとめる機会、1 年生後半から探求学習が始まること等)

・アウトプットを外部コンクールに出すことで、応用機会が広がり、生徒にも緊張感等をもって取り組む機会が与えられるとともに、これまでデータサイエンスにあまり興味をもっていなかった生徒の興味を掘り起こせる可能性もある。

・同じコンクールに参加した生徒間で相互コメントを行うことで、新たな視点を獲得等の効果。

一方で懸念としては、以下の点を想定していた。

・40 人以上の HR クラス単位の授業における一斉指導であり、指導が行き届かないまま終わる可能性がある。

・理解力やモチベーションに差があること。

・夏期休暇における課題とするので、当該期間に指導を行うことが困難であること。

・評価のうち、相互コメントの仕組みがうまく機能するか。

ただし、懸念点として挙げる部分の多くについては、全員を対象にした授業・課題出題という時点で避けられない部分であることから、グループでの演習機会を増やすなどして、少しでもモチベーションを維持して学習し続けられることに重きを置くよう心掛けた。

2. 3 実践によって確認された効果及び反省

2. 3. 1 1 学期の授業実践

少し時間をかけすぎた感もあるが、生徒がついてこられるようにグループごとにフォローさせて実施したことで、毎回の演習についても取り組むことはよくできていた。

一方で、カリキュラムや扱う材料を検討すれば、全 5 回くらいでも十分という感もあるので、その分の時間を、データを用いた演習・分析作業、データの見方・まとめ方といった時間に転用すべきだったと思案している。

また、後半 3 回の授業で夏期休暇中の課題の説明・導入を行った。どのコンクールが紙媒体もしくは電子媒体か、難易度、まとめ方(ポスター・文章・動画など簡単に提出形態等の説明)を伝え、1~5 人でグループごとに挑戦するコンクールを決めるところまでは、集中して取り組んでいた。

しかしながら、一部コンクールの手続きやデータ提供がまだ先だったこと、生徒の定期試験 1 週間前の時期になり、彼らにとって喫緊の課題として認識されづらくなったこともあり、作業の進め方やグループ内における作業の検討が曖昧なままになってしまった。

2. 3. 2 夏期休暇中の実践

夏期休暇中の対応として、オフィスアワーを 8 月に週 1 回程度計 4 回は事前にアナウンスしたが、利用は 2 件であり、ほとんど活用はされなかった。

夏期休暇中ということが最大の理由と思われるが、コンクールによっては、8 月上中旬に締め切りがあったものもあり、その対応は満足にできなかったところである。

一方で 8 月最終週になると、9 月初回の授業提出に向けて、学校への課題の提出方法など些末な件をメールで問い合わせようとする者が目立った。

2.3.3 2学期以降の実践

9月の初週に紙媒体指定のコンクール(東京都統計グラフコンクール)以外のコンクール参加者は、電子媒体を提出してもらい、図1のようにGoogle Driveにコンクールごとのフォルダを用意し、図2のように各フォルダ内に当該電子媒体を共有して見られるようにした(東京都統計グラフコンクールは紙媒体を9月初日にすぐ収集して提出先に発送。電子媒体の共有が可能な生徒のみ格納)。

生徒には同一コンクールにエントリーした、他の生徒グループを図3のように2グループずつ割り当てることにした。

図1 生徒へのコンクール別作品共有フォルダ等一覧

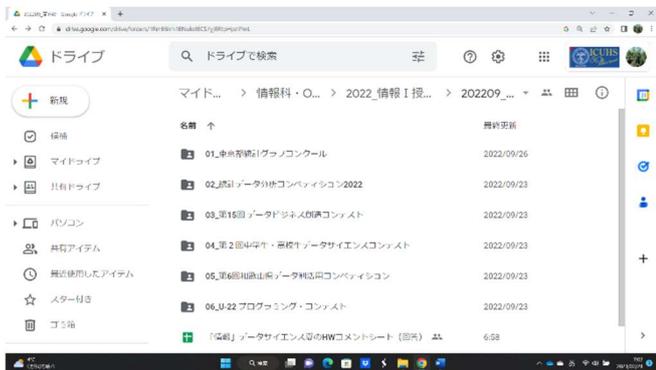


図2 電子媒体の各生徒への共有ファイル例

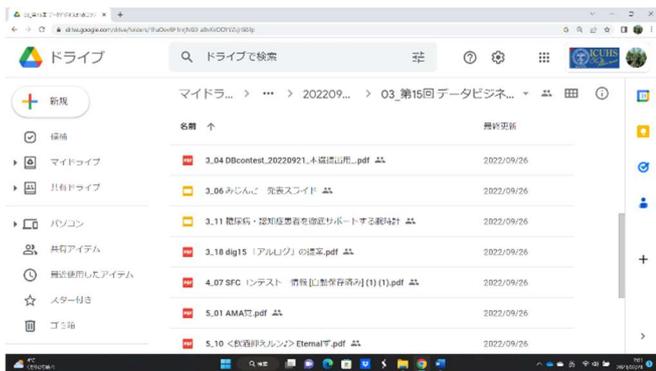


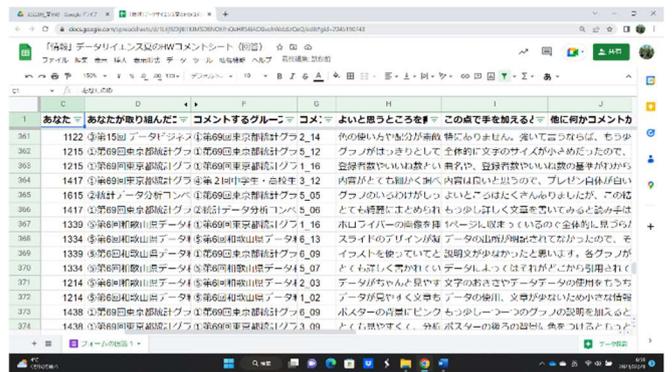
図3 各生徒への相互コメント先指示

ID	コンクール名	グループ	must1	must2
1101	⑤第6回和歌山県		1 5 02	3 14
1102	⑤第6回和歌山県		2 5 04	3 16
1103	⑤第6回和歌山県		1 5 07	4 03
1104	⑤第6回和歌山県		2 5 08	4 12
1105	⑤第6回和歌山県		2 5 09	5 02
1106	⑤第6回和歌山県		3 5 15	5 04
1107	⑥U-22 プログラ		4 2 05	5 14
1108	③第15回 データ		5 1 07	5 10
1109	⑤第6回和歌山県		6 6 01	5 07
1110	⑤第6回和歌山県		1 6 02	5 08
1111	③第15回 データ		7 1 05	5 11
1112	③第15回 データ		5 1 09	6 04
1113	⑥U-22 プログラ		4 2 10	2 05
1114	①第69回東京都		8 13 02	5 03

生徒は図3の割り当て指示をもとに、Google Formを用いて、「よいと思うところ」「手を加えらるともっとよくなると思うところ」をフォルダ内で作品を確認し、それぞれ当該グループに対し、コメントすることを求めた。

当該コメントについては、ある程度期間を置いてから、図4のとおり全員が見られるように共有を行い、自身のグループの作品に関して、どのように他者は見ているかを確認し、ふり返ってもらった。

図4 各生徒の相互コメント結果一覧(生徒共有画面)



なお、東京都統計グラフコンクール参加生徒については、紙媒体の返却が12月中旬になったことから、1月に校内のスペースに掲示し、図5のように授業時間外に確認してもらい、1月に相互コメントを実施した。

図5 東京都統計グラフコンクールの作品確認の様子



2学期以降の授業実践における問題点としては、次の点が挙げられる。

- ・紙媒体と電子媒体のコンクールが混在し、教員側の指導及び管理は煩雑なものとなってしまった。
- ・紙媒体のコンクール資料の返却・相互コメントが2学期中に終わらず、成績の反映が3学期になった。

いずれもコンクールの時期を想定しきれていなかった教員側の責任であるが、留意しておきたいところである。

3 考察

本授業実践は、全生徒に一度はデータ分析を経験させ、チャレンジさせることで、関心を持った生徒も一定数出てきたと見受けられる。また、紙媒体のコンクールの作品結果を相互コメントするために、当該作品をすべて貼りだしていたところ、他の教員や他学年の生徒から、「情報科ってこういうことをやっているのですね。」「意外にしっかり分析しているものも多く、新鮮。」といった反応が得られた。

情報 I の授業ということで、電子媒体で作品を集約し、相互コメントを実施することで問題はないが、校内での認知度向上・理解者の獲得・データサイエンスの機運向上といった観点では、紙媒体での展示の必要性も再認識したところである。

また、今回実施した相互コメントについては、自身の作品についてのコメントを自分でアクセスしないと読めないため、受動的な生徒はコメントを読んでいない可能性がある。

相互コメントを行って、そのコメントを受けた側の生徒自らが確認をすることが一番望ましいが、自身が携わった作品について振り返る時間を取ることを 2 学期のしかるべき時期に取ることに教育的意義があったのではないかと考えている。

4 まとめと今後の課題

23 年度以降の授業においても、1 学期にデータサイエンスの授業に取り組み、夏期休暇において、外部コンクールを念頭に置いた課題を出したいと考えている。

ただし、22 年度の授業実践においては、資料のまとめ方や学習した内容を活かしきれていない作品も多かったことから、次の点には特に留意して、授業計画は立てることとする。

- ・グループで提供データを分析する等して、授業計画を組み、一度は分析結果を簡単にまとめる作業を行うこと。
- ・夏期休暇中は質問が来ない前提で、1 学期の授業を組み、伝達すること。
- ・分析結果を出す段階で、それまでに学習したデータサイエンスの事項を活かせるようにすること。
- ・相互コメントにおいても、感想に留まらず、データサイエンスの観点からコメントをすることを促していくこと。

・相互コメントを踏まえたうえで、自身の携わった作品を振り返る時間を設けること。

・どの場においてアウトプットを出していくか、外部コンクールに参加することを課するのであれば、時期を踏まえて指定をするのか等を確認すること。

また、授業に関する部分以外として、校内の他教員へのデータサイエンス分野の授業実践を共有し、他教科でも積極的に情報 I で生徒が学習した内容が活かせる環境づくりに努めていくことが本校あるいは一般的な高校において必要であり、理解者を増やしていくことで、情報科の授業の幅は広がっていくことと考える。

5 謝辞(参考 URL)

本授業実践にあたって、生徒に対して適切な指導が足りない中、コンクールに応募した作品もあった。ご多忙中、ご対応いただいた各コンクールのご担当者様には貴重な機会をご提供いただき、感謝申し上げたい。

・第 69 回東京都統計グラフコンクール(東京都)

<https://www.toukei.metro.tokyo.lg.jp/kidsc/kc-2gcindex.htm#download>

・統計データ分析コンペティション 2022(総務省統計局)

<https://www.nstac.go.jp/statcompe/call.html>

・第 15 回データビジネス創造コンテスト

～Digital Innovators Grand Prix (DIG) 15～
(慶應義塾大学 SFC 研究所データビジネス創造・ラボ)

<https://dmc-lab.sfc.keio.ac.jp/dig15/>

・第2回中学生・高校生データサイエンスコンテスト

(神戸大学数理・データサイエンスセンター)

[http://www.cmds.kobe-](http://www.cmds.kobe-u.ac.jp/events/2022/2022_school_student_ds_contest/index.html)

[u.ac.jp/events/2022/2022_school_student_ds_contest/index.html](http://www.cmds.kobe-u.ac.jp/events/2022/2022_school_student_ds_contest/index.html)

・第 6 回和歌山県データ利活用コンペティション

(和歌山県)

https://datarikatsuyou.pref.wakayama.jp/jinzai/competition/2022_6/

医療メッシュ統計を活用したデザイン力の育成

和泉志津恵* ・滋賀大学データサイエンス学部

佐藤 彰洋 ・横浜市立大学データサイエンス学部

* 〒522-8522 滋賀県彦根市馬場1丁目1番1号, Email: shizue-izumi@biwako.shiga-u.ac.jp

1. はじめに

医療系のデータサイエンティストの育成について滋賀大学への期待は高い。滋賀大学は、2016年に「生物統計家育成支援事業」（主催 日本医療研究開発機構）に竹村と和泉らが中心となり応募した。日本初のデータサイエンス（DS）学部が開設された2017年に、経済学研究科プロフェッショナル・コース データサイエンスモデルにて「医療統計学基礎特講」を担当した。2018年にDS学部自主ゼミ「医療統計・防災」を担当した。このように医療統計学の教育環境を整えた。

統計数理研究所との連携活動として、2019年に「大学生のための医療統計学」の教育プログラムを共同開発した[6-9]。これは、統計数理研究所医療健康データ科学研究センター（研究所，学）の医療統計家，第一三共株式会社や味の素株式会社（企業，産）のデータサイエンティスト，そして和泉研究室（大学，学）の協働活動であった。この教育プログラムでは，大学生らが統計検定2級程度の知識を使い，課題解決に至るまでのPPDACサイクルの過程を体験することを目的にした。

データサイエンス教育の滋賀大学モデルでは，理系のスキルを持つとともに文系的なマインドも持つデータサイエンティストを育成する[2]。理系の長所として，論理的思考，問題解決力，発想力，検証力，推理力，数理処理能力があげられる[4]。一方，文系の長所として，語彙力，コミュニケーション力，協調性の高さ，感受性の豊かさ，柔軟な考え方，ユーモアのセンスがあげられる[5]。文理融合型のデータサイエンス教育では，統計系や情報系の授業だけでなく，課題解決型（Project Based Learning, PBL）演習が不可欠である[1]。PBL演習では，現実の社会を意識して，実データを用いてデータ分析から問題解決に取り組む。そこは，授業において学んだ内容の実践の場となる。演習をとおしてデザイン力が育成されていく。PBL演習の実施は，各種の連携機関が支援している。

本稿では，まず，「大学生のための医療統計学」の教育プログラムを解説する。次に，その教育プログラムにおける医療メッシュ統計を活用したデザイン力の育成について授業デザインを提案する。最後に，実践内容を紹介する。

2. 「大学生のための医療統計学」の教育プログラムの概要

授業は，対面での1年間である。参加対象は医療系DSに興味がある3・4年次生である。表1に2022年度関連科目と教育プログラムの内容を示す[10]。DS学部生は，1・2年次に記述統計や推測統計，統計解析ソフトウェア R の演習など医療統計学の基礎を学ぶ。3年次に医療統計論や機械学習など医療統計学の発展的な内容を学ぶ。教育プログラムのデザインは過去の実施状況をふまえ改善を繰り返した。

理論編では参加者が医療統計学の実践的な内容をゲスト講師や和泉から学んだ。

応用編では参加者が母集団100万人の医療データベースを用いて，(1) データに基づき参加者自身でデータ解析をデザインする，(2) (1)の計画に沿ってデータ研磨やデータ解析を行う，(3) (2)の結果を医療関係者へ説明する。参加者自身のノートPCを学内WiFi（または学外から大学へのVPN接続）から専用データサーバーへリモート デスクトップ接続し，医療データベースを利用した。患者の診療・治療・健診データを組み合わせ，解析用データを加工した。医療データベースは教育用の利用に限られ研究発表は不可であった。

学外アドバイザーは，4名（伊藤氏，立森氏，小山氏，徳山氏）であった。彼らはゲスト講義の実施，参加者への助言，教育プログラムの共同開発と改善を行った。和泉研究室の院生らは参加者のメンターの役割を担った。

コロナ禍での授業改革として，対面とZOOMオンラインの併用，LMS(Moodle)とMS-Streamを用いて講義のオンデマンド配信，SNSを用いてグルー

プ活動の対話支援，MS・Power Automate と LMS(Moodle)を用いて参加者の学習状況を時系列的に解析し，参加者の学習支援に役立てた．このようにニューノーマル時代に向けて統計教育のDX化も進んでいる．

表 1. 2022 年度関連科目とプログラム内容

学年	科目名
1	データサイエンス入門、基礎データ分析、統計学要論
2	回帰分析、多変量解析入門、統計数学、AI・情報理論、基礎統計活用演習A・B、データベース、社会調査法
3	医療統計論、医療統計演習、質的データ解析入門、AI・機械学習入門、社会調査実践演習
教育プログラムの内容	
理論編	医療統計学入門、統計的有意性とP値、臨床研究デザインと統計学、レセプトデータを用いた有害事象発現リスクの評価方法、企業での分析事例、医薬品開発の流れ、変化係数モデル入門、MESHSTATS の分析事例
応用編	医療統計学の行動基準に関する演習、医療データベースを用いた演習、構造化抄録と研究レポートの執筆・査読

2019 年度は，和泉研究室のゼミ生 8 名・経済学部生 1 名が高血圧症患者への処方薬に関する課題に取り組んだ．処方薬ディオバンに関わる臨床研究論文不正事件の影響を調べた学生もいた．国立精神・神経医療研究センターにて，データ解析結果をもとに課題解決策を提案し医療関係者と意見交換した．2020 年度は，ゼミ生 17 名・経済学部生 1 名が高血圧症や糖尿病の患者への処方薬に関する課題に取り組んだ．2021 年度は，ゼミ生 14 名が糖尿病や高脂血症の患者への処方薬に関する課題に取り組んだ．統計数理研究所医療健康データ科学研究センターからの 2021 年度最優秀賞が「高脂血症患者のメバロチンの処方」(3 年次)と「2 型糖尿病患者に対するグリメピリド処方の実態調査」(4 年次)の発表者に授与された．2020 年度は，ゼミ生 17 名・経済学部生 1 名が高血圧症や糖尿病の患者への処方薬に関する課題に取り組んだ．2022 年度は，ゼミ生 11 名が高脂血症の患者への処方薬に関する課題や学生定期健康診断の測定や問診データのクリーニングに取り組んだ．ZOOM オンラインにて，データ解析結果をもとに課題解決策を提案し医療関係者と意見交換した．教育プログラム修了生 4 名がデータサイエンス研究科に入学した．



図 1. 2021 年度成果報告会の対面参加者

3. デザイン力の育成

3.1 授業デザイン

本研究のきっかけは，MESHSTATS の理論と応用についてのゲスト講義であった(2022 年 1 月，横浜市立大学データサイエンス学部・佐藤彰洋氏による)．その際，学生たちは，位置と時間の付いた MESHSTATS を用いたストーリーづくりに取り組んだ[11]．

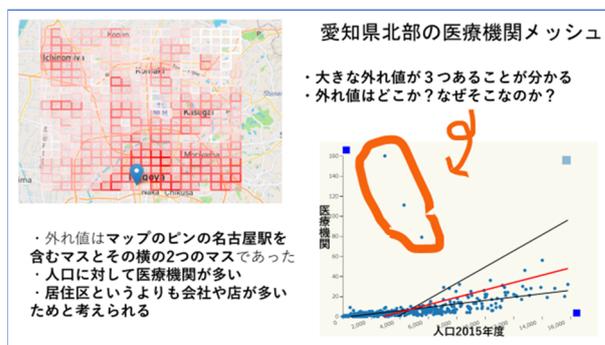


図 2. 医療機関メッシュを用いた例

学部3・4年生 15 名程度を対象にした，1 セメスター(15コマ)の医療統計 PBL 演習をデザインする．学習目標は，PPDAC サイクルを活用して，(1) 医療データに潜む課題に気づく，(2) データに基づきデータ解析をデザインする，(3) (2)の計画に沿って解析用データを加工・研磨する，(4) データ解析の結果を表や統計グラフにまとめる，(5) データを根拠にした提言を医療関係者へ説明することである．

人材(2022 年度)は，授業担当が和泉，ゲスト講師が 5 名(横浜市立大学 佐藤彰洋氏，北海道大学病院 伊藤氏，慶應義塾大学 立森氏，京都大学 佐藤俊哉氏，県立広島大学 富田氏)，アドバイザーが 4 名(伊藤氏，立森氏，小山氏，徳山氏)，オブザーバーが 2 名(滋賀大学保健管理センター 山本氏，DS・AI イノベーション研究推進センター 池之上氏)，TA が 1 名(DS 研究科 安田氏)である．

環境(2022 年度)は，学生自身のノートPCを持参する BYOD(Bring your own device)を推奨し，LMS(Moodle)による授業運営を行い，MESHSTATS モニター申込や

サポート会員申込を行い、EXCEL, Rstudio, Python をデータ解析に用い、ZOOM を学外講師とのコミュニケーションに用い、MURAL をオンラインワークボードに用いる。

授業計画(2022 年度春学期)は、オンライン講義(5 コマ)、データ解析演習(8 コマ):医療統計家の行動基準、課題のアイデア出し、研究計画書作り、MESHSTATS からデータの抽出・結合、データ解析/提案内容の根拠を視覚化、レポート作成、報告会(2コマ):最終発表を配分する。MESHSTATS を用いた PBL 演習をととしてデザイン力を育成することを目的とする。学生らは、4 年生 2 名と 3 年生 1 名の計 3 名からなる班を構成する。この PBL 演習の活動結果を、「MESHSTATS アプリケーションアイデアソン 2022」(図 3)ならびに「サマーデザインワークショップ 2022」(図 4)の 2 つの学外活動にて発表する。

MESHSTATS アプリケーションアイデアソン 2022 は、一般社団法人世界メッシュ研究所が主催し、MESHSTATS へ搭載が可能なアプリケーションアイデアを募集することが目的である[13]。テーマは、SDGs 内の 4 つの目標である、目標 7:エネルギーをみんなにそしてクリーンに、目標 8:働きがいも経済成長も、目標 9:産業と技術革新の基盤をつくろう、目標 11:住み続けられるまちづくりを、が選ばれ、SDGs×MESHSTATS で都市の社会課題解決を目指す。



図 3. アプリケーション アイデアソンポスター



図 4. サマーデザイン ワークショップポスター

「サマーデザインワークショップ 2022」は、横浜市立大学が主催し、学生らがデータ分析レポートをインプットとしたワークショップ (WS) を企画し参加することで体験学習することが目的である[14]。このデータ駆動型デザインワークショッププログラムの設計と実践は、2019 年度より文部科学省超スマート社会の実現に向けたデータサイエンティスト育成事業「YOKOHAMA D-STEP」の必

修科目「デザイン思考」、2020 年度より横浜市立大学大学院データサイエンス研究科データサイエンス専攻の必修科目「デザイン思考特論」の講義設計を原型としている[12]。企画者(ファンリテータ)は、PBL およびデザイン思考の手法を用いた WS 企画書(データ分析レポート、プログラム企画書)を作成する。一方、参加者は、グループ作業、データ利活用、デザイン思考や問題解決フレームワークを体験学習する。運営委員会が、応募された WS 企画書の中から最大 12 件を採択する。対面形式・オンライン形式で 1 日 WS を実施し、1グループ 5 名までの参加者を募集する。WS 最終日に各 WS の成果を発表会にて説明する。得られた知見は参加者で共有し、WS 成果を HP で公表するオープンイノベーション型をとる。なお、企画提案資格や参加資格は、基本的に、主催、協賛、後援組織に所属する者、または、それら組織からの推薦を受けた者に限られている。表 2 にサマーデザインワークショップの組織分布を示す。期間中の延べ参加者数は 150 名余りであった。

表 2. サマーデザインワークショップの組織分布
(1)企画者 (2)参加者

所属組織	人数	所属組織	人数
横浜市立大学	6	横浜市立大学	32
総務省	2	総務省	1
一般社団法人世界メッシュ研究所	1	一般社団法人世界メッシュ研究所	1
株式会社丹青社	1	JST 未来社会創造事業	1
滋賀大学	1	JTB 総合研究所	1
鎌倉市役所	1	滋賀大学	5
		その他	1
合計	12	合計	42

授業デザインを組み立てる際に、横浜市立大学 DS 研究科「デザイン思考特論」2022 年度シラバスを参考にした[15]。学習到達目標は、1) デザイン思考と WS の基本的な仕組みを理解する、2) デザイン思考の基本的な方法論を理解する、3) デザイン思考の基本的な方法論を組み合わせた WS を設計できるようになる、4) データ分析レポートを利用したデータデザイン WS の事例を学ぶ、である。講義概要には、1) デザイン思考に対する知識習得を目的として、参加型デザインの方法論、データ分析レポートの作成方法、デザイン思考の方法論とそれらで利用可能な手法を取り扱う、2) 参加型デザインの方法論として、チームビルディングの方法、ブレインストーミングの方法を取り扱う、3) データ分析レポートの作成方法を公的統計調査を利用して学習する、4) サービスデザインの観点から UML について説明する、5) ユーザーエクスペリエンスの観点からペルソナ的设计、カスタマージャーニーマップの描き方を実際の実例を通じて学修する、6) プロトタイピングの方法と得られた結果のレビュー方法について実際のグループワー

クを通じ体験的な学修を行う、7) データ分析の結果をデザイン思考に有効に活用するために、計算論的思考法とデザイン思考との関係について論じる、である。この中の、参加型デザインの方法論、データ分析レポートの作成方法、デザイン思考の方法論、チームビルディング、ブレインストーミング、ペルソナ、ジャーニーマップ、プロトタイピングに着目して、デザイン力の育成を図る。

3.2 実践内容

ゼミの活動記録(2022年度春学期)を表3に示す[16, 17]. 前半の4~5月は、WEB上のMESHSTATSの操作に慣れることに重きをおき、MESHSTATS上に新たに追加したいアプリのアイデアを構想した。学生らは、目標11に取り組み、自身のテーマで「ユースケース」、「リーンキャンパス」、「アウトプット例」を作成した。班ごとの提案を「MESHSTATS アプリケーションアイデアソン 2022」にて発表した。後半の6~7月は、医療メッシュ統計の解析計画を立て、WEB上のMESHSTATSから得られたメッシュデータを結合・加工し、データ解析し、レポートを作成した。その際、「ブレインストーミング」の方法を学び、「ペルソナデザイン」を設計し、「ジャーニーマップ」を描く演習に参加した。また、オンラインワークボードのMURALの使い方にも慣れた。学生らの企画書は採択され1日WSを実施した。その際、企画書の応募やWSへの参加について一般社団法人世界メッシュ研究所から推薦いただいた。

表3. 活動記録(2022年4~7月)

ゼミ		アイデアソン2022	
・オリエンテーション、MESHSTATS利用登録、アイデアソン申込、演習「アプリのアイデア」	4月12日	・お知らせ	4月初旬
・講義「医療統計とは何か」、演習「医療統計家の行動基準」		・参加申込締め切り	4月15日
・講義「臨床研究のデザインと統計学」、「レセプトデータを用いたリスクの評価」	4月19日	・インプットセッション	4月26日
・講義「MESHSTATSを使った統計的解析方法と応用」、演習「MESHSTATS」			
・演習「ユースケース」、「リーンキャンパス」、「アウトプット例」、「Web上の配置図」	5月10日	・成果物提出締め切り	5月13日
・成果物PPTXを作成			
・自主ゼミ「医療統計ことはじめ: その1」の説明動画MP4を作成	5月17日	・発表会	5月21日
・発表会の準備: 3チーム出場			
・講義「分析設計」、演習「MESHSTATSを用いた研究計画書」	5月24日		
ゼミ		サマーデザインWS2022	
・演習「MESHSTATSを用いたデータ分析」、WS企画を話し合い	6月7日	・お知らせ	6月初旬
・講義「医学・薬学における質的データ解析入門」、「製薬企業における帰帰分析の活用」			
・データ分析レポート発表会(1)	6月21日		
・自主ゼミ「医療統計ことはじめ」第1回			
・データ分析レポート発表会(2)、WS企画書の選抜	6月28日		
・講義「検査データの解釈: 検査の陽性、陰性」			
・「デザイン思考WS2022」の見学		・企画申込締め切り	7月1日
・WSに向けた準備、リハーサル(1)	7月12日	・企画書の採択通知	7月11日
・演習「ブレインストーミング」、「ペルソナデザイン」、「ジャーニーマップ」		・WS参加申し込み開始	7月15日
・WSに向けた準備、リハーサル(2)	7月19日	・ファシリテータ会議	8月8日
・演習「MURAL」		・WSの開催	8月22日
・自主ゼミ「医療統計ことはじめ」第2回		・発表会	8月24日

表4. 企画書の例

提案者氏名	滋賀大学データサイエンス学部 崎岡 小畑 小林 田中
ワークショップの概	コロナ禍で人口当たりの医療機関数、病床数が大きな話題となり、緊急事態宣言等の発令基準となったこともあった。そこで、今回高齢人口も多いと予測される京都で人口当たりの医療機関数が適切であるのかについてMESHSTATSを用いて調査したい。そこで、本ワークショップでは新たな医療機関の建設要件について議論する。特に、ホテル数やコンビニ数が多い地域では医療機関を必要とする患者が多いと考えられるため、ホテルやコンビニの分布に着目することで新たな医療機関の建設要件を提案する。
ワークショップの目的	今回高齢人口も多いと予測される京都での人口当たりの医療機関数が適切であるのかについてMESHSTATSを用いて調査していく。そこで、京都の中で地域特性、医療機関の偏りもあるかと思う。そこから見えてくる課題を精査したい。また、未来の開業医に向けて、開院するためのおすすめ地域を診療科別に提案出来たらよいと思う。外れ値を見つけることで、医療機関を建設したいシチュエーションを柔軟に検討する。さらに、それらのシチュエーションを診療科別で分類する。ペルソナデザインや、ジャーニーマップにより、医療機関を建設したい側を明確に可視化する。
取り扱う課題	地域ごとの人口に対する適切な医療機関数の調査。医療機関を必要とする人に医療を提供できるように、医療機関を建設したい側をペルソナデザインにより理解する。また、ジャーニーマップを描いてシチュエーションの具体化を目指す。そして、医療機関の建設要件を議論する。
キーワード	医療機関、建設、観光(ホテル)、経済(コンビニ)
データとデータ源	国土数値情報 医療機関データH26年度 2010年総務省統計局国勢調査3次メッシュデータ(日本) 2015年総務省統計局国勢調査3次メッシュデータ MESHSTATSのオープンストリートマップ 内閣府。(2022). 令和4年版高齢社会白書(概要版)。
デザイン思考の方法	ブレインストーミング、ペルソナ、ジャーニーマップ
期待される成果	観光、経済、人口分布を考慮した新たな医療機関の建設要件を提案する
プログラム*	10:00 開始 10:00-10:10 アイスブレイキング(自己紹介) 10:10-10:30 データ分析レポートの説明 10:30-11:30 Rを用いた相関分析、結果の可視化 11:30-12:00 結果の共有、解釈 12:00-12:50 お昼休み 12:50-13:20 ブレインストーミング 13:20-13:50 ペルソナデザイン 13:50-14:20 ジャーニーマップ 14:20-14:50 新たな医療機関の建設要件の提案 14:50-15:00 小休憩 15:00-16:00 試作発表会資料作成 16:00-16:45 発表会(フィードバック、リフレクション) 16:45 終了

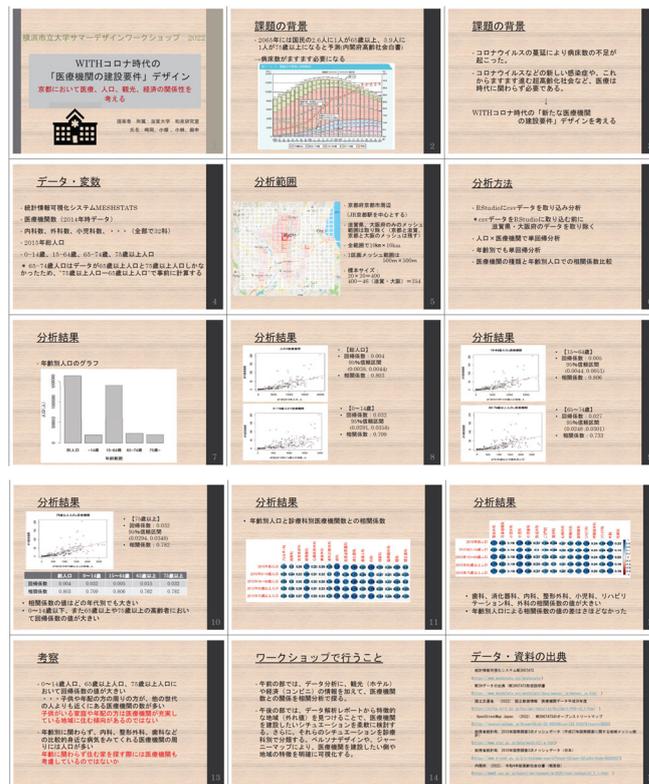


図5. 企画書スライドの例

次に、「WITH コロナ時代の医療機関の建設要件デザイン」というタイトルで学生らは1日WSを実施した。企画者は学部3年生1名と4年生3名からなるグループであった。春学期の間は、就職活動中である4年生は、3年生の活動をサポートする役割を担った。彼らのWS企画書を表4と図5に示す。

WSでは、JR京都駅を中心にした10km四方のメッシュデータ、1平方km単位のメッシュ、OSM由来メッシュ統計、医療機関：国土交通省国土数値情報由来メッシュ統計を利用した。項目は、2015年0~14歳人口総数(人)、2015年15~64歳人口総数(人)、2015年65歳以上人口総数(人)、2015年75歳以上人口総数(人)、夜間光強度(平均値)、医療機関、アレルギー科、麻酔科、…、外科、ホテルなどの観光関連施設、コンビニエンスストアなどの商業施設などを利用した。統計ソフトウェアRstudioを用いて、これらのメッシュ統計を位置情報に関して結合しデータ解析用に加工した。

データ分析では、医療機関数を従属変数、人口・観光指標・経済指標を説明変数に用いて重回帰分析を行った。ステップワイズ法を用いた変数選択や、VIFを用いた多重共線性の確認、残差plotを用いて特徴的な観測地点の割り出しを行った(図6)。

4.まとめ

本稿では、「大学生のための医療統計学」の教育プログラムを解説し、その教育プログラムにおける医療メッシュ統計を活用したデザイン力の育成について授業デザインを提案し、実践内容を紹介した。

医療系データは匿名化やデータの管理に関して制約があり、教育用に活用する場が限られる。一方、医療メッシュデータは、既に匿名化されたデータであるため教育用の利用が容易である。また、相談に応じてメッシュ統計データの追加も可能である。

デザイン力の育成には、コミュニケーション力と論理的思考の両方の育成が大事である。加えて、問題解決につながりそうな想像(妄想)を学生に促すことも大切である。自分の意図を相手に分かりやすく伝えることは、ボディー・ランゲージ(body language)が伝わりにくいオンラインの方が対面より難しいと、学生らの印象に残った。ニューノーマル時代では、その伝わりにくさをどのように補っていくのが新たな課題と考えられる。

デザイン力の育成では、エビデンスに基づいた統計的思考の育成を重要視する。また、その教育活動の振り返りにおいてもエビデンスをもとに振り返ることが大切だと考える。

謝辞

本研究は、統計数理研究所・共同研究利用・重点型研究「データサイエンティスト育成に向けたカリキュラム・教材に関する研究」(29-共研-4306, 30-共研-4105, 2019-統数研-一般研究2-00209, 2020-統数研-一般研究2-00064, 代表者 和泉志津恵, 受入教員 川崎能典氏), 共同研究利用・重点型研究「医療における時空間メッシュデータの利活用についての研究」(2021-ISMCRP-41101, 2022-ISMCRP-4302, 代表者 和泉志津恵, 受入教員 村上大輔氏), および共同研究集会「統計教育の方法とその基礎的研究に関する研究集会」(2022-統数研-共同研究集会-5007, 代表者 末永勝征, 受入教員 椿広計氏), 滋賀大学・研究ユニット事業(30-研究助成-U01, 2019-研究助成-U01, 2022-研究助成-U01, 代表者 和泉志津恵)の助成を受けた。また、次の指導・助言者(五十音順)に感謝を申し上げる。池之上辰義氏・山本祐二氏(滋賀大学), 伊藤陽一氏(北海道大学病院, 元 統計数理研究所), 小山暢之氏(第一三共株式会社), 佐藤恵子氏・佐藤俊哉氏・福岡真悟氏(京都大学), 立森久照氏(慶応義塾大学), 徳山健斗氏(中外製薬株式会社, 元 味



図6. 成果報告会スライドの例

の素株式会社), 富田哲治氏 (県立広島大学), 中村治雅氏 (国立精神・神経医療研究センター病院), 松井茂之氏 (統計数理研究所医療健康データ科学研究センター). サンプルデータの整備については, 滋賀大学データサイエンス学部や大学院データサイエンス研究科の学生たちに協力いただいた. ここに謝意を表す.

【参考文献・参考 URL】

1. 和泉志津恵, 櫻井尚子, 深澤弘美. (2016). 大学の統計教育でのインタラクティブな授業のデザインとその評価方法. 統計数理研究所共同研究レポート 362 「統計教育実践研究」 8:5-10.
2. 竹村彰通, 和泉志津恵, 齋藤邦彦, 姫野哲人, 松井秀俊, 伊達平和. (2018). データサイエンス教育の滋賀大学モデル. 統計数理 特集「統計教育の新展開」 66(1): 63-78.
<https://www.ism.ac.jp/editsec/toukei/pdf/66-1-001.pdf>
3. Aki-Hiro Sato (Ed.) (2019). Applications of Data-Centric Science to Social Design: Qualitative and Quantitative Understanding of Collective Human Behavior. Springer, Singapore. ISBN-13: 978-9811071935.
4. 小学館. (2020). 「理系脳」はどう作る? AI 時代に求められる子どもの思考力を鍛える方法.
<https://dime.jp/genre/922316/>
5. 小学館. (2020). 時代が求める「理系脳」の特徴と思考パターン. <https://dime.jp/genre/930941/>
6. 滋賀大学 DS 教育研究センター. (2020). ゼミ紹介 医療, 防災, 行政の実データの特徴を探索する. Data Science View, 4 : 24.
<https://www.ds.shiga-u.ac.jp/dscenter/about/#report>
7. 滋賀大学 DS 教育研究センター. (2021). 「学生による企業での成果発表: ①医療データベースを用いた課題研究の成果報告. Data Science View, 5 : 38.
8. 滋賀大学. (2021). きらきらしがだい人 医療統計学. 滋賀大学 広報誌「しがだい」, 54: 13.
https://www.shiga-u.ac.jp/information/about_public/guidebook/#backnumber
9. 滋賀大学. (2022). データサイエンス学部の学び: Student's Voice. 「大学案内 2023」, P.27.
https://www.shiga-u.ac.jp/information/publish/info_publish-guidebook/
10. 滋賀大学 DS・AI イノベーション研究推進センター. (2022). 「大学生のための医療統計学」2021 年度報告. Data Science View, 6: 50.
11. 和泉 志津恵, 伊藤 陽一, 松井 茂之, 佐藤 彰洋. (2022). 「大学生のための医療統計学」教育プログラムとメッシュ統計の活用 次年度の活動計画. 統計数理研究所 2021 年度共同研究集会重点型共同利用研究「地図・メッシュ・位置情報データのデータベース作成・統合と高度利用」.
<https://www.fttsus.org/webconf/20220322-20220323/#speaker12>
12. 佐藤彰洋, 水村由美. (2022). デザインワークショッププログラムの設計と実践. 2022 年度 統計関連学会連合大会 大会予稿集.
13. 一般社団法人世界メッシュ研究所. (2022). MESHSTATS アプリケーションアイデアソン 2022.
<https://www.fttsus.org/webconf/meshstatsapp2022/>
14. 横浜市立大学. (2022). 横浜市立大学サマーデザインワークショップ (WS)2022 .
<https://www.fttsus.jp/jstmirai/events/ycusdw2022/>
15. 横浜市立大学データサイエンス研究科 データサイエンス専攻. (2022). 「デザイン思考特論」2022 年度シラバス.
[https://cmweb.yokohama-cu.ac.jp/campusweb/slbsbdr.do?value\(risyunen\)=2022&value\(semekikn\)=1&value\(kougicd\)=YDDM02400&value\(crclumcd\)=2017515200](https://cmweb.yokohama-cu.ac.jp/campusweb/slbsbdr.do?value(risyunen)=2022&value(semekikn)=1&value(kougicd)=YDDM02400&value(crclumcd)=2017515200)
16. 和泉志津恵, 佐藤彰洋. (2023). 「大学生のための医療統計学」教育プログラムにおける医療メッシュ統計の活用. 統計数理研究所 2022 年度共同研究集会重点型共同利用研究「地図・メッシュ・位置情報データのデータベース作成・統合と高度利用」.
<https://www.fttsus.org/webconf/20230220-1/>
17. 滋賀大学データサイエンス学部. (2022). 「大学生のための医療統計学」成果報告 2019 年度～2022 年度.
<https://www.ds.shiga-u.ac.jp/news-faculty/p8163/>
<https://www.ds.shiga-u.ac.jp/news-faculty/p6932/>
<https://www.ds.shiga-u.ac.jp/news-faculty/p6070/>
<https://www.ds.shiga-u.ac.jp/news-faculty/p5035/>

公開型アダプティブオンラインテストのアナリティクス： 大学の基礎数学での COVID-19 以前と以降

廣瀬 英雄（久留米大学バイオ統計センター，中央大学研究開発機構）

hirose_hideo@kurume-u.ac.jp, hhideo001@g.chuo-u.ac.jp

1 はじめに

今日まで，項目反応理論（IRT, Item Response Theory）は，TOEFL を含む様々な場面で，受験者の能力を正確かつ公正に評価するために使用されてきている（例えば，[1-4]）。また，IRT は問題項目の難易度を同時に推定することも可能である。ただし，厳密な評価試験では常に問題や解答例を伏せておくことが多い。これらの利点から，例えば，国立大学で出題される問題が公開されていない状況下でも，受験者の能力の年次変化を測定することが可能になる [5,6]。受験者をユーザー，問題項目をアイテムとも呼ぶこともある。

この IRT は，評価に限らず，学習効果を高めるための練習にも利用できる。例えば，受験者の能力と問題の難易度を自動的にマッチングさせるアダプティブオンラインテストは，そのような使い方として効率的で効果的な方法となり得る。アダプティブテストでは問題の解説とその解答を公開することで学習を促進させることも多いからである。ただし，普通，アダプティブテストでは，通常の IRT と異なり，受験者を行とし問題を列とする完全な項目応答マトリクスを構築できないため，項目の難易度は常に初期値として使用し，受験者の能力のみを推定している。受験者の能力に調整された問題の困難度を設定するには，一定期間利用した段階での問題へのキャリブレーションが有効になる。

広島工業大学では，上のような学習評価と学習促進の両方の目的達成がなされるように，大学学部生を対象としたオンラインテストシステムを開発した [7]。そのための数学科目の問題を事前に多数作成し，問題を，厳密な評価試験に使用するものと，学習促進のための演習に使用するものと分離させておき，システムはそれぞれ評価用と演習用を分けて出題している [8]。

また，上記の，大学が運営するテストシステムとは別に，学部の教科書の付録ツールとして，線形代数，微積分，基礎解析，確率・統計，常微分方程式，基礎物理などの学部向け Web 支援型アダプティブオンラインテストシステムを開発し，学生がいつでも自由に利用できるようにしている [9]。このシステムでは，大学での演習用アダプティブテストと同様に，問題に対する解説とその解答は公開されている。また，このオンライン演習は，教科書を持たなくても，誰でも自由にアクセスが可能である。本稿では，このオンライン演習による学習効果について述べる。特に，2019 年度

と 2020 年度をまたいだところで COVID-19 による大学での学習形態が大きく変わったため，COVID-19 以前（2019 年度まで）と COVID-19 以降（2020 年度から）で学習効果に違いが見られるか，両者の比較について述べる。

項目応答マトリクスの要素すべてに回答が埋められている完全マトリクスを用いた厳密なアセスメントテストの場合には，評価結果の分析研究は多く報告されている（例えば，[5,6]）が，上記のアダプティブオンラインテストの場合，問題や解答およびその解説が開示されているため，受験者の能力を分析することは難しいと考えられる。更に，大規模テストにおける項目開示の有用性についての研究は限られており，継続的かつ慎重な検討が必要であるとの指摘もある [10]。こういった課題に対し，本稿では問題と解答例や解説を開示した条件下でのアダプティブオンラインテストの利用における学習効果を調査することとした。

オンラインテストによる学習効果については，限られてはいるが，以下のような研究成果は報告されている。Bock らは，College Board Physics Achievement Test のデータにおいて，10 年間にわたる項目難易度パラメータの差分線形ドリフトが実証されたと主張している [11]。しかし，Feinberg らは，このような結果に対して，資格試験において同一形式の繰り返しは不当な優位性を生むことはない，反対の見解を示している [12]。また，Wagner-Menghin らは，テスト項目を再利用しても個人の得点は上昇しないが，項目の難易度の変化に関する情報が不足しており，再利用項目の平均難易度の低下が観察されたが，重要ではなかったと述べている [13]。また，Tang らは，同じ項目に遭遇することによる利点は限られていると述べている [14]。これに対して，Raymond らは，2 回目の受験のスコアを基にした推論の有効性を報告した [15]。さらに，Selvi は，初回使用時の項目難易度は，繰り返し使用時の項目難易度より有意に低いことを述べている [16]。また，Ferreira らは，人間の記憶特性については，因子分析により，視空間記憶と言語-数的記憶が異なる変数であるが，相関があることが示されていると述べている [17]。このような扱いの難しさについては，文献 [18] に，記憶の長所と短所が項目タイプによって異なることが示されている。同様に，開示の効果は開示項目の性質に依存することが，Gilmer らによって報告されている [19]。

このように、能力パラメータの信頼性、難易度パラメータの推定値へのドリフト現象、記憶優位性については、項目開示の下では依然として不明確である。特に、アダプティブオンラインテストでは、関連する文献を見つけることができなかつた。したがって、現在でも、問題・解答が開示されている条件下で、アダプティブオンラインテストの学習効果を調査することは、意味のあることだと思われる。ただし、難易度パラメータなど IRT におけるパラメータの推定値をそれだけで分析することは困難と思われるので、本稿では学習の有効性を論じるために問題への正答率 (CAR, Correct Answer Rate) を取り入れ、主にそれについての評価を行った。その際、受験者の能力推定値 (ability) や、問題項目の困難度との関係も考慮している。ここで扱うデータは、文献 [9] に示す、Web 支援型アダプティブオンラインテストシステムから得られたものである。

の正否に従って問題のレベルを上げ下げする。このとき、問題レベルと受験者能力値とがマッチするような問題をコンピュータシステムによって自動的にデータベースから探して、受験者に与える。ここでは、一連のテストでの問題数を 5 問 (主に COVID-19 以前) とか、7 問 (主に COVID-19 以降) とかに設定している。この様子を図 1 に示す。

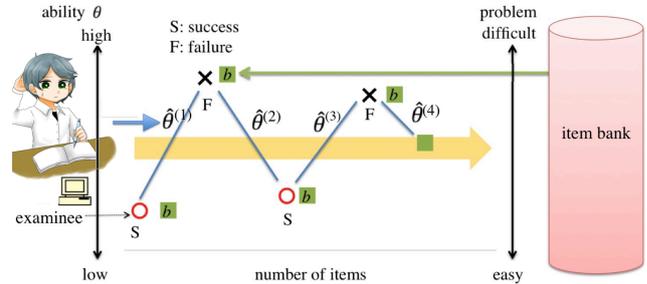


図 1: オンラインアダプティブテストシステム

2 Web アシストオンラインアダプティブテスト

2.1 項目反応理論

項目反応理論では、受験者 i が問題 j に正答する確率を次の P_{ij}

$$P_{ij}(\theta_i; a_j, b_j) = \frac{1}{1 + \exp\{-1.7a_j(\theta_i - b_j)\}},$$

$$= 1 - Q_{ij}(\theta_i; a_j, b_j) \quad (1)$$

で仮定している。ここに、 θ_i は受験者 i の能力パラメータ、 a_j, b_j は問題 j の特徴を表すパラメータで、それぞれ、識別度、困難度と呼ばれている。 Q_{ij} は受験者 i が問題 j に正答できない確率になる。このとき、最尤推定法を用いれば、パラメータ θ_i, a_j, b_j の最尤推定値 $\hat{\theta}_i$ と \hat{a}_j, \hat{b}_j は、次の尤度関数 L を最大にすることで求められる。

$$L = \prod_{i=1}^m \prod_{j=1}^n (P_{ij}^{\delta_{ij}} \times Q_{ij}^{1-\delta_{ij}}) \quad (2)$$

ここに、 m と n は、それぞれ、受験者数と問題数で、 δ_{ij} は、受験者 i が問題 j に正答できたとき $\delta_{ij} = 1$ 、できなかったとき $\delta_{ij} = 0$ と定義するインジケータ関数である。

2.2 オンラインテストシステム

本稿ではアダプティブテストを取り扱う。このとき、通常、IRT の問題の識別度、困難度パラメータは事前に与えられているため、受験者の能力値は簡単に求められる。最初に出題される問題のレベルは、事前に受験者の能力値の記録があればそれを、なければ中程度のレベルに設定され、次の問題は、受験者が前の問題に正答したときにはレベルを上げ、正答しなかったときには下げるというように解答

アダプティブテストシステムでは、受験後、問題とその解答および解説の開示を学習者から求められることがある。テストシステムを、評価というより、演習として使う際に、学習効果を高めるには問題と解答の開示は有効であると考えられるので、このシステムではそれらを開示している。図 2 に問題と解答解説の開示例を示す。

図 2: 問題と解答解説の例

2.3 データベースサイズ

ここで取り扱っている Web アシストオンラインアダプティブテストシステムは、大学の基礎数学の演習を補う意

図で作られたものであり、2015年1月14日、最初に教科書「線形代数」[20]出版時に利用が開始された。それに続いて、「確率・統計」[21],「微分積分」[22],「基礎物理学」[23],「1変数の微積分」[24],「常微分方程式」[25]などの教科書が出版されると同時にオンラインシステムも稼働し始めた。表1に、それぞれのテストシステムの開始時期、2023年2月2日までに蓄積されたデータベースでの、受験者数、問題数、アクセス数を示す。

表 1: Web アシストオンラインアダプティブテストシステムの科目とデータベースサイズ

科目	開始時期	受験者数	問題数	アクセス数
線形代数	Jan. 2015	2,253	860	72,192
確率・統計	Jul. 2015	93	172	1,797
微分積分	Feb. 2017	886	636	38,006
基礎物理学	Feb. 2017	68	133	1,261
1変数の微積分	Apr. 2020	247	945	1,458
常微分方程式	Dec. 2021	34	191	265

2015年1月14日から2023年2月2日まで

3 COVID-19以前とCOVID-19以降の学習効果の比較

3.1 正答率, CAR の定義

記録されたログファイルをもとに、回答行為を行った回数に対して、問題に正答した割合を「正答率」CAR, Correct Answer Rate とよび、次で定義する。

$$CAR = \frac{\text{正答数}}{\text{回答数}} \quad (3)$$

一連のアダプティブテストには k 個の問題が与えられるので、1回のアダプティブテスト受験での回答数は k となる。

CAR には、一人一人の受験者に対する受験者ベース CAR と、1問1問の問題に対する問題ベース CAR とがある。これらを区別するには、例えば、 $CAR_{\text{受験者}}$ とか $CAR_{\text{問題}}$ とか記法が必要になるが、ここでは、後者のことを主に取り扱うので CAR を $CAR_{\text{問題}}$ と同一視する。

ログファイルには、次の3つの場合が記録されている: 1) 回答を試みて正答できたとき, 2) 回答を試みて正答できなかったとき, 3) 回答を試みようとしたが途中で断念したとき, である。この3)を回答行為数に入れるか(ケース1)入れないか(ケース2)で、線形代数と微分積分の場合について比較したものが図3と図4である。両図でケース1(右側)とケース2(左側)での CAR の値を比較すると、2017年度を除いて、両者にほとんど違いがみられない。そこで、

本稿では、ケース2の場合での(3)式から計算された数値を CAR の値とする。参考までに、図中の“sample size”は、左側がケース2の場合の回答数、右側がケース1の場合の回答数を示す。

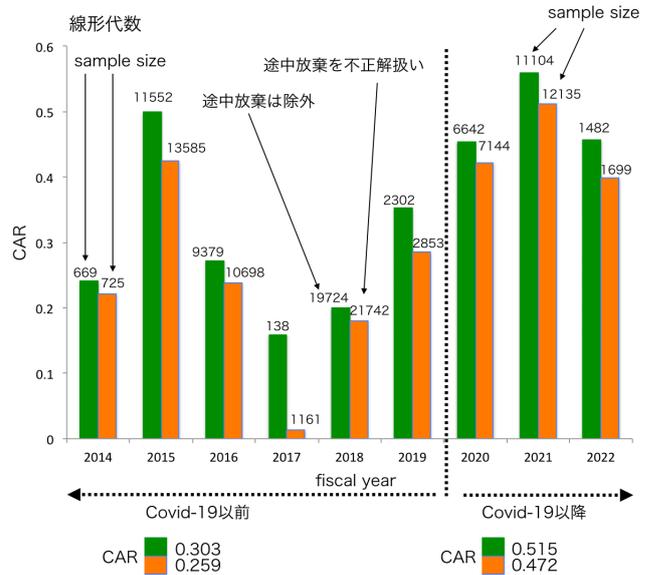


図 3: 2つのケースの CAR の年トレンド (線形代数)

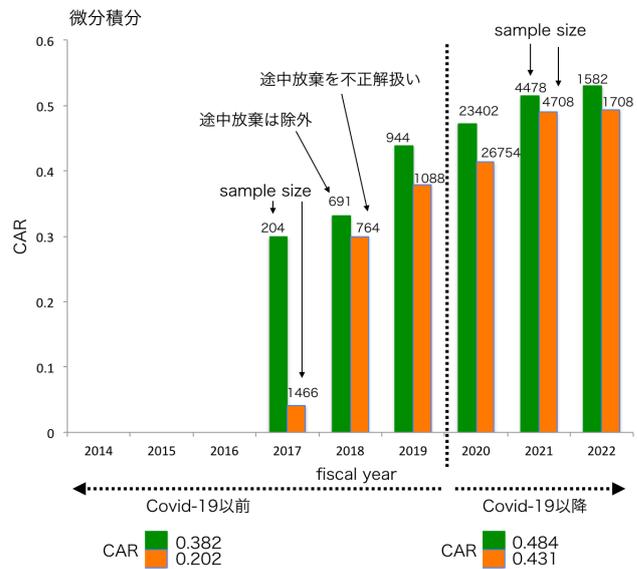


図 4: 2つのケースの CAR の年トレンド (微分積分)

図3と図4の CAR の年トレンドをおおまかに見ると、正答率 CAR は年々上がっているように見え、一見、学習効果が上がっているように思われる。

3.2 オッズ, odds の定義

CAR は回答数に対する正答数の比であるが、ここで、オッズについても(4)式で定義しておく。オッズは改善度を表

現するのに用いられている。

$$\text{オッズ} = \frac{\text{正答数}}{\text{誤答数}} \quad (4)$$

表 2 に COVID-19 以前と COVID-19 以降の CAR とオッズの比較を示す。CAR については、COVID-19 以前では 0.5 よりも小さいが COVID-19 以降では 0.5 付近の値をとっており、COVID-19 以前よりも COVID-19 以降の方が数値が大きい。また、オッズについても、COVID-19 以前よりも COVID-19 以降の方が数値が大きい。ただし、確率的な変動によって数値は変わるので続く 3.3 で統計的検定を用いて確認する。

表 2: COVID-19 前後での CAR とオッズの比較

科目	COVID-19 以前			
	正答数	誤答数	CAR	オッズ
線形代数	13,244	30,520	0.303	0.434
確率・統計	156	734	0.175	0.213
微分積分	691	1,118	0.382	0.618
基礎物理学	203	499	0.289	0.407
1変数の微積分				
常微分方程式				
科目	COVID-19 以降			
	正答数	誤答数	CAR	オッズ
線形代数	9,906	9,322	0.515	1.063
確率・統計	39	56	0.411	0.695
微分積分	14,913	15,886	0.484	0.939
基礎物理学	67	76	0.469	0.882
1変数の微積分	668	571	0.539	1.170
常微分方程式	62	181	0.255	0.343

3.3 CAR 比とオッズ比による COVID-19 前後での学習効果

問題の難易度と受験者の能力値の組み合わせによって、得られる CAR とオッズは変動する可能性があるが、ここでは、COVID-19 以前でのそれらの組み合わせと COVID-19 以降でのそれらの組み合わせの間にそれほど相違はないと仮定して、検定を行ってみた。

2次元コンティンジェンシー表によって検定を行うため、次の、2つの式によって、CAR 比とオッズ比を定義した。

$$\text{CAR 比} = \frac{\text{COVID-19 以降の CAR}}{\text{COVID-19 以前の CAR}} \quad (5)$$

$$\text{オッズ比} = \frac{\text{COVID-19 以降のオッズ}}{\text{COVID-19 以前のオッズ}} \quad (6)$$

CAR 比とオッズ比の推定値と下側 95%信頼限界の値を、表 3 に示す。これらの比が 1 より大きければ、COVID-19 以前よりも COVID-19 以降の方が受験者はより問題に正答していることになり、学習効果が上がっていることが期待される。表からは、下側 95%信頼限界値はすべて 1 よりも大きいので、有意に学習効果が上がっていると考えられる。しかしながら、実際には問題の難易度と受験者の能力値はさまざまであるため、更に詳細に検討を行って考察する必要がある。

表 3: CAR 比とオッズ比の推定値と下側 95%信頼限界

subject	CAR 比	95%下側信頼限界
線形代数	1.702	1.669
確率・統計	2.342	1.718
微分積分	1.268	1.195
基礎物理学	1.620	1.290
subject	オッズ比	95%下側信頼限界
線形代数	2.449	2.365
確率・統計	3.277	2.102
微分積分	1.519	1.378
基礎物理学	2.167	1.502

3.4 問題毎の CAR 値とオッズ値の変化

これまでの解析から、問題すべてにわたる全体的なトレンドを見ると、Web アシスト演習の学習効果は平均的には上がっているように思われる。しかし、問題バンクには難しい問題と易しい問題が混在しているので、問題毎でどうなっているか、そのトレンドも見る必要がある。ここでは、データの量が最も多かった線形代数の科目の場合について述べる。

図 5 に、線形代数における、2016 年度から 2019 年度まで (COVID-19 以前) の問題毎の CAR 平均値と 2020 年度から 2022 年度まで (COVID-19 以降) のそれと比較したプロット図を示す。正確には、CAR_{問題} の比較になっている。2014 年度と 2015 年度のデータは問題の困難度のキャリブレーションを行うことに使ったため、CAR 値の比較からは除外した。また、問題によっては回答数が少ない問題もあるため、そのような問題も除外している。例えば、線形代数のデータベース上の問題総数は表 1 に示すとおり 860 であるが、このうち一部はシステム始動時のテスト用の問題も含まれ、また、その中の一部は問題へのアクセス数が十分でないため問題困難度のキャリブレーションを行っていない。860 から、この始動時テスト用の問題数 388 と、キャリブレーション対象外の問題数 145 を引いた問題数 327 が、ここで学習効果の改善度比較に用いた問題数である。

図から、全体的には CAR 値は COVID-19 以前よりも COVID-19 以降の方が大きいことがわかる。多くの点は原点から伸びる傾き 1.7 の直線のまわりに分布しており、この傾きの値は表 3 の CAR 比 1.702 と合致している。

図 6 に、線形代数における、2016 年度から 2019 年度まで (COVID-19 以前) の問題毎の対数オッズと 2020 年度から 2022 年度まで (COVID-19 以降) のそれと比較したプロット図を示す。対象とした問題は図 5 と同じである。図から、全体的には対数オッズは COVID-19 以前よりも COVID-19 以降の方が大きいことがわかる。これらの 2 つの図は、CAR_{問題} についてのおおまかな分布も示してくれる。

CAR の値が大きくなっているということは、解けている問題の数が増えているということで、単純には学習効果が上がっているというように思える。

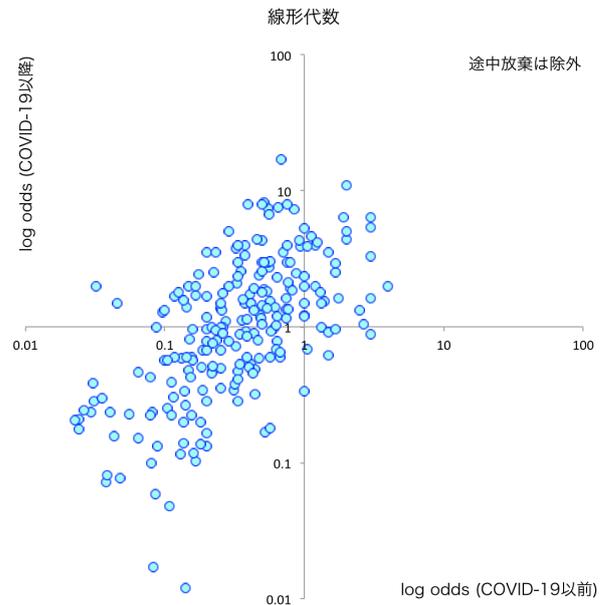


図 6: 線形代数における COVID-19 以前と COVID-19 以降の対数オッズの比較

用いたものと同様である。ここで、線形回帰は、問題 j に対して、すべての対象となるアクセス毎に、説明変数 (x_s, s : 年度) を年度に、目的変数を年度毎の応答値 ($y_{ij}^s = 1$ (正答), $y_{ij}^s = 0$ (誤答), i はアクセス番号) としている。

表 4 に、1) CAR 値の年トレンド変化の線形回帰の傾きが正であるかどうか、2) COVID-19 以前 (2016 年度から 2019 年度まで) と COVID-19 以降 (2020 年度から 2022 年度まで) の CAR 比が改善されているかどうかについて検定を行い、有意水準 5% で (変化がないという) 帰無仮説が棄却された問題数を示した。表の下部に、トレンドと CAR 比と記載されているのは、1) と 2) の両方を同時に満たすときの問題数である。CAR が COVID-19 以前よりも COVID-19 以降で改善されたと明確に言えるのが 105 個の問題についてで、改善されなかったというのがわずかに 2 問になる。CAR に注目する限り、学習改善はなされていると見るべきであろう。なお、改善されなかったうちの 1 問については、後の考察の節で詳しく述べる。

表 4: CAR 値の年トレンド変化と COVID-19 前後での変化

	上昇した問題数	下降した問題数
トレンド	120	2
CAR 比	117	2
トレンドと CAR 比	105	2

2016 年 4 月 1 日から 2023 年 2 月 2 日までのキャリブレーションを行った 327 問の問題の場合。(有意水準 5%)

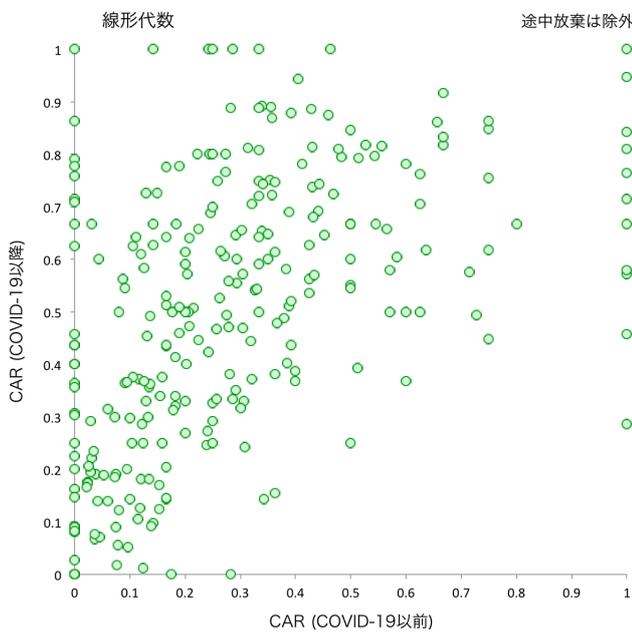


図 5: 線形代数における COVID-19 以前と COVID-19 以降の CAR 値の比較

3.5 CAR 値の年トレンド変化と COVID-19 前後での変化

上で述べたことは、CAR 値やオッズ値が COVID-19 以前と COVID-19 以降で違っているかどうかを目視して学習効果の改善の可能性を指摘しているにすぎない。そこで、ここでは問題毎に、1) CAR 値の年トレンド変化を線形回帰して、直線の傾きが正であるかどうか、また、2) COVID-19 以前 (2016 年度から 2019 年度まで) と COVID-19 以降 (2020 年度から 2022 年度まで) の CAR 比が改善されているかどうかについて検定を行い、対象となるすべての問題数の中のうちの程度の問題数に対して受験者の学習効果が表れているかを確認する。対象とする問題は上の 3.4 で

3.6 CAR 値, ability 値, 問題困難度の関係

図5で, CAR 値の平均が COVID の前後でどのように変わったかを示した. 一見すれば, このことから学習効果の改善が見られるとも思われる. 確かに, すべての受験者に全く同じ問題と同じ制限時間内に解かせるような従来の評価テストでは, 正答率が上がったことと受験者の能力値が向上したことは同じになる. しかし, ここで取り扱っているデータはアダプティブテストによる結果である. アダプティブテストでは, 出題される問題は受験者の潜在能力値に合わせた困難度を持つ問題が自動的に選ばれている. そこで, ここでは, そのような背景を考慮した上での学習効果の改善について調べてみる.

図7は, 線形代数において, 各問題に対する能力値の平均を COVID-19 以前と COVID-19 以降で比較したものである. 全体的に見ると, どの困難度の問題に対しても, 図5に見られたような COVID の前後での大きな違いは見られない. しかし, 注意深く見ると, COVID-19 以降の能力値の方が少しばかり COVID-19 以前の能力値よりも大きいようにも見える. キャリブレーションした問題の困難度はこの図では隠れているので, 次に, 問題の困難度が見えるように, 問題の困難度に対する能力値をプロットしてみた.

図8に, キャリブレーション後の各問題の困難度に対する能力値の平均の関係を示す. 図を見ると, まず, 問題困難度と受験者の能力値には非常に高い相関 (相関係数は 0.925) がみられるので, 出題される問題の困難度が受験者の能力値に合わせてアダプティブに設定されていることが確認できる. つまり, アダプティブシステムは正常に機能していることがわかる.

次に, 図8では, COVID-19 以前と COVID-19 以降に分けて, 両者に違いがあるかどうかを確認できるようにしているが, 両者を比較すると, COVID-19 以前の能力値は COVID-19 以降の能力値よりも少し上側にシフトしているように見える. つまり, 同じ問題を与えたときの, COVID-19 以前の受験者の習熟度よりも COVID-19 以降の受験者の習熟度の方がわずかに高いということがわかる. ただし, 目視では確認が難しいので平均の差の検定を問題毎に行ってみた. 表5にその結果を示す. 表の下部に, ability 値平均とトレンドと CAR 比と記載されているのは, ability 値平均, トレンド, CAR 比すべてが同時に改善されたときの問題数である. このことから, 同じ問題に取り組んだ受験者の能力値が高くなっていった傾向が確認できる. また, 表4の結果と合わせても, COVID-19 以降には COVID-19 以前と比較して学習効果の改善が見られると考えられる.

習熟度が高い集団になれば問題困難度も同時に高い方にシフトしていくと考えられる. そこで, 図9では, COVID-19 以前と COVID-19 以降のアクセス問題の比較を行ってみた. 図の縦軸は, すべての問題へのアクセス総数に対する当該問題へのアクセス数の比率である. 図からは, COVID-19 以降のアクセス率が高いのは問題困難度が大きいときに多

表 5: ability 値平均の COVID-19 前後での改善

	改善された問題数	改善されなかった問題数
ability 値平均	130	21
ability 値平均とトレンドと CAR 比	83	0

2016年4月1日から2023年2月2日までのキャリブレーションを行った327問の問題の場合. (有意水準5%)

いように見える. あるいは, 同じアクセス率を示す問題困難度は COVID-19 以前に比べて COVID-19 以降の方が右にシフトしているように見える. このことは, アダプティブシステムが提供する問題の困難度が受験者の習熟度に合わせて自動的にシフトされていることを表していると考えられる.

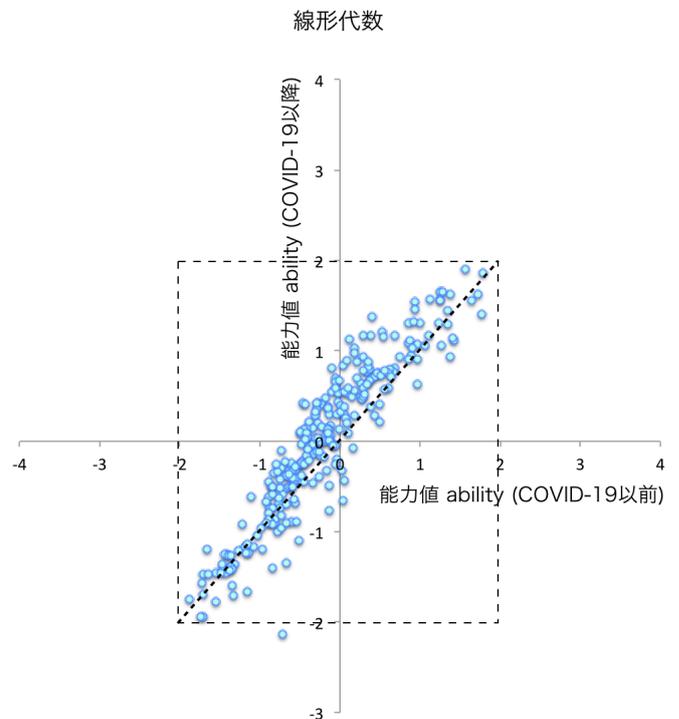


図 7: 線形代数における COVID-19 以前と COVID-19 以降の能力値の比較

4 考察

本稿で取り扱っているアダプティブテストでは, 受験者の学習効果が上がるように, 受験後には解答例だけでなく解答に至る説明も開示されている. もし同じ受験者に同じ問題が提示されれば, 以前閲覧した解答例とその説明の記憶をもとに, 二回目の受験時には正答率が上がるように働く可能性を否定できない. ただ, この行為は, 連続する受験

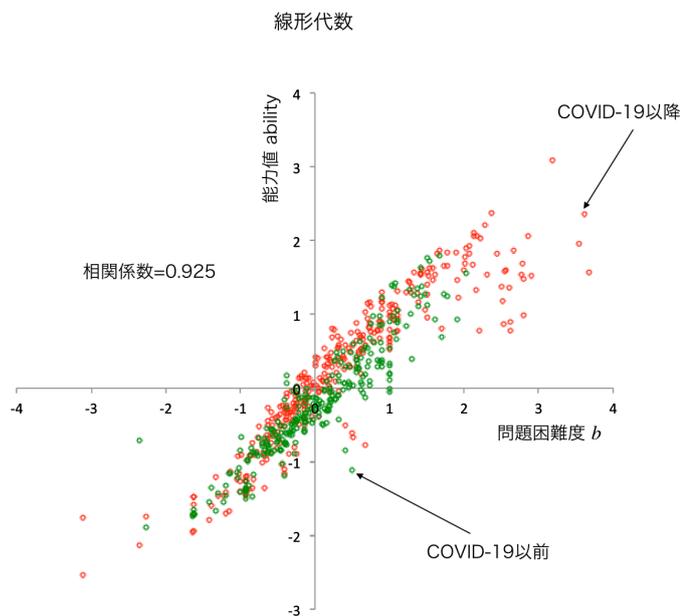


図 8: 線形代数における問題困難度を介した COVID-19 以前と COVID-19 以降の能力値の比較

機会の間は長くはないだろうということから、COVID-19 の前後で正答率が変化するように機能するとは思えない。長期間にわたっての問題と解答の開示による正答率の上昇のもう一つの可能性はチートである。一人の受験者で閉じた受験を行うのではなく、複数の受験者との協働によって問題と解答が保持され引き継がれていくという可能性である。そのようなことが行われているかどうかについて確認をとるのは難しいように思われる。しかし、偶然ではあるが、チートの可能性を否定できるような事象をデータベースから発見できた。ここでは、3.5 で言及したように、CAR に関して、COVID-19 以前よりも COVID-19 以降で改善されなかったうちの 1 問について詳しく述べる。

表 6 に、問題 id782 と問題 id783 の CAR の年変化をそれぞれ示している。表には、参考として、問題の識別力 a と困難度 b の設定値も付けている。これらの設定値は、システムが稼働し始めた 2014 年度から 2015 年度末まで使った初期値と、IRT によってキャリブレーションを行った後で 2016 年度から 2022 年度まで使った値である。これら 2 つの問題は、三角関数を要素に持つ 2 つの行列の積を求める問題で、内容はほとんど同じである。従って、CAR のトレンドも同じようになり、また、COVID-19 前後での学習効果の改善も同様な傾向が見られることが想定される。しかしながら、問題 id782 では明確な学習効果の改善が見られる（表 4 の 105 のうちの 1 つ）のに対して、問題 id783 では逆に改善が見られない。この問題 id783 が表 4 で改善が見られなかった 2 問のうちの 1 つである。

問題 id782 では、CAR の値が 0.467 であるため、 b のキャリブレーション値は 0 に近くなっており、また、CAR の値も少し上昇しながら 0.5 付近を保っている。一方、問題

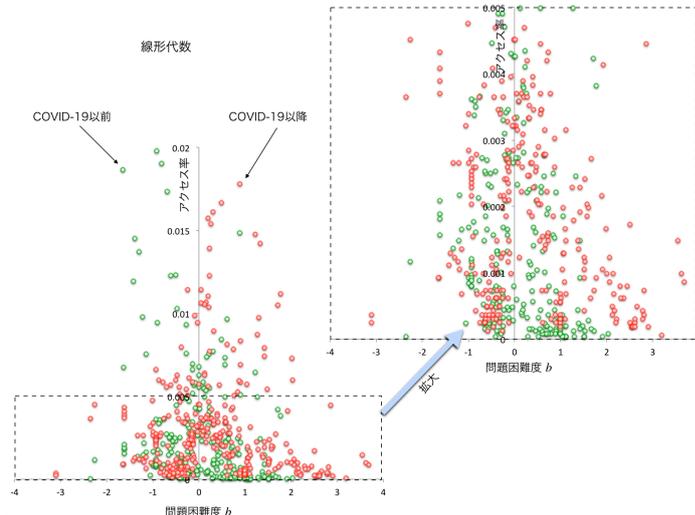


図 9: 線形代数における COVID-19 以前と COVID-19 以降のアクセス問題の比較

id783 では、問題の困難度は問題 id782 と同じであるにもかかわらず、CAR の値は 0.045 と非常に低い値で、また、そのため、 b のキャリブレーション値は 1.71 と高い値になっている。しかも、年々 CAR 値は下降気味である。

この矛盾するような現象の原因を探している最中に、問題 id783 の中の小問 8 問のうち 2 問の解答に間違いがあることが見つかった（正確には 2 問の正答の順序が逆になっていた）。この間違いは偶然 9 年間誰も気がつかずに放置されていた。その結果、このような矛盾した事象が起こったと考えるとこの不合理的が理解できる。そして、さらに重要なことは、もし、チートが頻繁に行われていたと仮定すれば、システムの中の解答を見つけた受験者がそれをそのまま入力して正答にできるはずで、正答率 CAR はこんなにも悪くはないし、ましてや年々下降することも考えられない。つまり、仮定したようなことはなかったと考える方が自然で、チートは行われていない、あるいは行われていたとしても無視できる程度の量ではなかったかと考えられる。なお、データベース内での問題 id783 の解答訂正は 2022 年 12 月 27 に終了している。

5 まとめ

項目反応理論は、TOEFL を含む様々な場面で、受験者の能力を正確かつ公正に評価するために使用されてきている。項目反応理論を用いた厳密な評価テストでは、通常、問題項目は複数回使用できるように隠されており、能力評価の推定結果の信頼性を高めているように思われる。しかし、項目反応理論では問題の困難度も同時に推定できるため、アダプティブテストにも利用可能である。アダプティブテストでは、受験者の学習意欲を高めるために問題の解説や解答を開示することが求められることも多い。開示された場

表 6: 問題 id782 と問題 id783 の CAR の年変化と a と b

年度	id782 CAR	id782 a	id782 b	id783 CAR	id783 a	id783 b
2014	-	1.00	0.00	-	1.00	0.00
2015	0.467	1.00	0.00	0.045	1.00	0.00
2016	0.318	1.08	0.119	0.038	0.801	1.71
2017	-	1.08	0.119	-	0.801	1.71
2018	0.310	1.08	0.119	0.175	0.801	1.71
2019	0.214	1.08	0.119	0.056	0.801	1.71
2020	0.533	1.08	0.119	0.016	0.801	1.71
2021	0.567	1.08	0.119	0.010	0.801	1.71
2022	0.500	1.08	0.119	-	0.801	1.71

合、推定された能力値の信頼性が気になる。

ところが、完全な回答マトリクスが利用できる場合でも、大規模テストにおける項目開示の有用性に関する研究は限られている。また、アダプティブオンラインテストによる学習が効果的かどうか重要であるが、そのような研究成果も乏しいようである。

そこで、本稿では、問題と解答例およびその解説を開示した条件下でのアダプティブオンラインテストにおける学習効果について検討した。用いたデータベースには、教科書の付録ツールとして、線形代数、微積分、基礎解析、確率・統計、常微分方程式、基礎物理などの大学の学部向け Web 支援型アダプティブオンラインテストシステムによって蓄積されたものを使った。本稿では特に、2020 年の初め頃から流行が始まった COVID-19 が学習効果に与えた影響を見るため、2019 年度以前のデータと 2020 年度以降のデータを比較した。分析にあたっては、項目反応理論におけるパラメータの推定値を解析するだけでは学習効果の分析は困難と思われるため、本稿では正答率やオッズの解析も評価法に加えた。学習改善効果の有無などの統計的な検証には、2 次元コンティンジェンシー表解析や線形回帰分析も用いた。

分析の結果、Web 支援型適アダプティブオンラインテストシステムは正常に機能していることが確認でき、また、それを利用するユーザーには、COVID-19 以降特に学習効果の改善が見られることがわかった。

謝辞

Web アシスト演習での問題と解答の解説の作成には、佐藤好久、池田敏春、岡崎悦明、岡山友昭、齋藤夏雄、田上真、廣門正行、若狭徹、大澤智興、桑田精一、田中公一、藤原真、藤野友和、高藤政典、山内雄介、小山哲也（敬称略）のみなさまにご協力いただきました。ここに謝意を表します。

参考文献

- [1] de Ayala R. The theory and practice of item response theory. Guilford Press; 2009.
- [2] Baker FB, Kim SH. Item response theory: Parameter estimation technique, 2nd edn. Marcel Dekker; 2004.
- [3] Hambleton R, Swaminathan H, Rogers HJ. Fundamentals of item response theory. Sage Publications; 1991.
- [4] van der Linden WJ. Handbook of item response theory. Chapman and Hall/CRC; 2016.
- [5] Hirose H. Prediction of success or failure for examination using nearest neighbor method to the trend of weekly online testing. International Journal of Learning Technologies and Learning Environments. 2019; 2:19–34.
- [6] Hirose H. Current failure prediction for final examination using past trends of weekly online testing. In: 9th International Conference on Learning Technologies and Learning Environments; 2020. p. 142–148.
- [7] Hirose H. Meticulous learning follow-up systems for undergraduate students using the online item response theory. In: 5th International Conference on Learning Technologies and Learning Environments; 2016. p. 427–432.
- [8] Hirose H, Takatou M, Yamauchi Y, et al. Questions and answers database construction for adaptive online irt testing systems: Analysis course and linear algebra course. In: 5th International Conference on Learning Technologies and Learning Environments; 2016. p. 433–438.
- [9] Hirose H. Learning analytics to adaptive online irt testing systems “ai arutte” harmonized with university textbooks. In: 5th International Conference on Learning Technologies and Learning Environments; 2016. p. 439–444.
- [10] Park YS, Yang EB. Three controversies over item disclosure in medical licensure examinations. Medical Education Online. 2015;23:1–5.
- [11] Bock RD, Muraki E, Pfeifferberger W. Item pool maintenance in the presence of item parameter drift. Journal of Educational Measurement. 1988;25:275–285.
- [12] Feinberg RA, Raymond MR, Haist SA. Repeat testing effects on credentialing exams: Are repeaters misinformed or uninformed? Educational Measurement: Issues and Practice. 2015;34:34–39.
- [13] Wagner-Menghin M, Preusche I, Schmidts M. The effects of reusing written test items: A study using the rasch model. ISRN Education. 2013;:1–7.
- [14] Tang X, Schultz M. The effect of repeat exposure to simulation based items. Practical Assessment, Research, and Evaluation. 2020;25:1–10.

- [15] Raymond MR, Swygert KA, Kahraman N. Psychometric equivalence of ratings for repeat examinees on a performance assessment for physician licensure. *Journal of Educational Measurement*. 2012;49:339–361.
- [16] Selvi H. Should items and answer keys of small-scale exams be published? *Higher Education Studies*. 2020; 10:107–113.
- [17] Ferreira AI, Almeida LS, Prieto G. The role of processes and contents in human memory: An item response theory approach. *ISRN Education*. 2011; 23:873–885.
- [18] Wood TJ, Swygert KA, Kahraman N. The effect of reused questions on repeat examinees. *Advances in Health Sciences Education*. 2009;14:465–473.
- [19] Gilmer JS. The effects of test disclosure on equated scores and pass rates. *Applied Psychological Measurement*. 1989;13:245–255.
- [20] 桂利行, 池田敏春, 佐藤好久, 廣瀬英雄. 理工系学生のための線形代数 : Web アシスト演習付. 培風館; 2015.
- [21] 廣瀬英雄, 藤野友和. 確率と統計 : Web アシスト演習付. 培風館; 2015.
- [22] 桂利行, 岡崎悦明, 岡山友昭, 齋藤夏雄, 佐藤好久, 田上真, 廣門正行, 廣瀬英雄. 理工系学生のための微分積分 : Web アシスト演習付. 培風館; 2017.
- [23] 大澤智興, 桑田精一, 田中公一, 藤原真, 廣瀬英雄, 小田部荘司. 理工系学生のための基礎物理学 : Web アシスト演習付. 培風館; 2017.
- [24] 廣瀬英雄, 高藤政典, 山内雄介, 小山哲也. 1変数の微積分 : Web アシスト演習付. 培風館; 2020.
- [25] 桂利行, 若狭徹, 岡山友昭, 佐藤好久, 田上真, 廣瀬英雄. 理工系学生のための微分方程式 : Web アシスト演習付. 培風館; 2021.

特異値分解から見た項目反応理論の新評価： 大学数学の CBT によるテスト

廣瀬 英雄（久留米大学バイオ統計センター，中央大学研究開発機構）

hirose_hideo@kurume-u.ac.jp, hhideo001@g.chuo-u.ac.jp

1 はじめに

最近の統計検定にも見られるように，CBT（Computer Based Testing）は，PBT（Paper Based Testing）と比較して受験者が増加する傾向にある [1,2]．それは，CBT には，公平・公正性の特徴に加え，即時性はもとより DX への適用性も高いと考えられるからである．また，今後は，大学入学共通テスト「情報 I」の出題に関する調査研究 [3] にも見られるように，CBT ならではの対話性も加味できる可能性を持つと考えられ，CBT は今後も更に発展していくと思われる．

CBT を用いれば，受験者の能力と問題の困難度を同時に精度よく評価できる項目反応理論（IRT, Item Response Theory）（[4-7]）を容易に利用することもできる．この IRT は，その汎用性と信頼性から，今日では受験者のパフォーマンスを評価するための標準的な手法の一つと見なされている．そのため，TOEFL をはじめとする様々な公的試験で利用されている．

IRT では，ある統計的パラメトリックモデルを仮定したとき，ある受験者がある問題項目に回答したときの正否から構成されるマトリクス（これをここでは「観測応答マトリクス」と呼ぶ）を入力データとすれば，最尤推定法を用いることによって，IRT パラメータである各受験者の能力と，各問題の困難度の推定値とその信頼区間を出力として求めることができる．IRT は，事前に各問題に配点を与えて，各問題の得点の合計点を評価値とする古典的評価法と異なり，問題の困難度を同時推定した上で受験者の能力値を推定できる点が特徴である．また，このことによって，問題の困難度を受験者の能力値に合わせたアダプティブテストも可能になり，少ない問題数で精度の良い評価値を得ることができる．

最尤推定法によって求められた推定量は，ある条件下で漸近正規性が成り立つことが知られている [8]．また，最尤推定量以外に漸近平均 2 乗誤差が小さくなる推定量は存在しない．これは，想定した数学モデルとそのパラメータ空間の下では，その推定量が最良であることを意味する．従って，パラメトリックな IRT モデルの範囲内では，得られるパラメータ推定量が最良である．また，それを用いた関数の推定量も最良になる．つまり誤差は最小になる．

この方法を，IRT のパラメータの推定値を用いて再構築

した項目応答マトリクス（これをここでは「推定応答マトリクス」と呼ぶ）に適用することを考える [9]．この意味は，仮定されたモデルの範囲内での最尤推定量の振る舞いを評価するのではなく，推定応答マトリクスが観測応答マトリクスにどの程度近いかという近似度を，直接マトリクスノルムを用いて評価し，IRT が持っている推定能力を再評価しようということである [10]．ここではこの近似度を，特異値分解（Singular Value Decomposition, SVD）によって観測応答マトリクスの低ランクマトリクスを求めた上で，推定応答マトリクスがどのランクの低ランクマトリクスに匹敵するか比較することによって表現する．

ただし，推定された IRT パラメータは漸近的に正規分布に従う変動を伴っているため，ここで求める近似度についてもその変動について考慮する必要がある．つまり，推定ではなく予測の意味での近似度を評価する必要がある．ここでは，低ランクマトリクスの変動を評価する際にはブートストラップ法を用いる．

この方法を，実際に大学で実施される多数の数学テストに CBT を使い，得られた観測応答マトリクスに適用した結果，パラメトリックな IRT モデルから得られた推定応答マトリクスはかなり良い結果をもたらすことがわかった．つまり，IRT は，パラメトリックモデルとして制限された範囲であるにもかかわらず，予測の意味で，観測応答マトリクスを表現できている低ランクマトリクスに十分近い推定応答マトリクスを出力していることがわかった．

2 IRT と推定応答マトリクス

2.1 IRT モデル

標準的な IRT では，観測応答マトリクスを入力データとしたとき，受験者の能力値 θ_i ($i = 1, \dots, n$) と問題の困難度 a_j, b_j, c_j ($j = 1, \dots, m$) を同時に推定することができる．通常，この $n \times m$ マトリクスの要素は $1/0$ の値 δ_{ij} から成り立っている．ここで， δ_{ij} には，受験者 i が問題 j に正解したとき 1 を，正解しなかったとき 0 を与える．

IRT モデルでは，受験者 i が問題 j に正答できた確率を

次のロジスティック確率 p_{ij} で与える.

$$\begin{aligned} p_{ij}(\theta_i; a_j, b_j, c_j) &= c_j + \frac{1 - c_j}{1 + \exp\{-1.7a_j(\theta_i - b_j)\}}, \\ &= 1 - q_{ij}(\theta_i; a_j, b_j, c_j), \end{aligned} \quad (1)$$

ここで, θ_i は受験者 i の能力値, a_j, b_j, c_j は, それぞれ, 問題 j に対する, 識別度パラメータ, 困難度パラメータ, 当て推量パラメータと呼ばれている. q_{ij} は受験者 i が問題 j に正答できなかった確率になる. ここで取り扱う問題では多肢選択の選択数が多く当て推量が困難と思われる状況を考えるため, 当て推量パラメータは用いない.

最尤推定法を用いれば, パラメータ θ_i, a_j, b_j に対する最尤推定量 $\hat{\theta}_i, \hat{a}_j, \hat{b}_j$ は, 次の尤度関数

$$L = \prod_{i=1}^n \prod_{j=1}^m (p_{ij}^{\delta_{ij}} \times q_{ij}^{1-\delta_{ij}}). \quad (2)$$

を最大化することによって求められる.

パラメータ θ_i および a_j, b_j, c_j をまとめて Θ で表し, 観測応答マトリクスを $\Delta = (\delta_{ij})$ で表すとき, 最尤推定法を用いてパラメータ推定を行うことを, 簡単に,

$$\Delta \rightsquigarrow \hat{\Theta}. \quad (3)$$

と表現する.

2.2 推定応答マトリクス

推定された $\hat{\Theta}$ を (1) 式の p_{ij} に代入すれば, $[0, 1]$ で連続な $\hat{\Theta}$ の関数 $\hat{\delta}_{ij}$ が得られる. この操作を,

$$\hat{\Theta} \rightsquigarrow \hat{\Delta}, \quad (4)$$

で表す. ここで, $\hat{\Delta} = (\hat{\delta}_{ij})$ である. この操作で得られたマトリクス $\hat{\Delta}$ を「推定応答マトリクス」と呼ぶ. (1) 式から, $\hat{\delta}_{ij}$ の値は, 受験者 i が問題 j に正答する確率を推定することになる. $\hat{\Theta}$ は漸近正規性を持つため, $\hat{\delta}_{ij}$ も漸近正規性を持ち, (誤差を最小にするという意味で) 最良の推定値となる.

3 特異値分解と低ランク応答マトリクス

3.1 特異値分解, SVD

$A = (a_{ij})$ を $m \times n$ マトリクスとする. このとき, $A^T A$ は $n \times n$ の対称マトリクスに, AA^T は $m \times m$ の対称マトリクスになる. ここで, A^T は A の転置を表す. $A^T A$ および AA^T に対する固有値と固有ベクトルは, (存在すれば) 同じになる. $A^T A$ の固有値と固有ベクトルを $\{\xi_1, \xi_2, \dots, \xi_n\}$ および $\{v_1, v_2, \dots, v_n\}$ とする. つまり,

$$A^T A v_i = \xi_i v_i. \quad (5)$$

である.

固有値を, $\xi_1 \geq \xi_2 \geq \dots \geq \xi_r > 0, \xi_{r+1} = \dots = \xi_n = 0$ のように, 降順に並べ替える. ここで, r は $A^T A$ のランクとする. $A^T A$ は対称であるから, 固有ベクトルは直交系をなす. つまり, $v_i \cdot v_j = I_{ij}$ である. ここに, I_{ij} は, $I_{ii} = 1, I_{ij} = 0 (i \neq j)$ となるインジケータ関数である. u_i を, $\sigma_i = \sqrt{\xi_i}$ として, $u_i = A v_i / \sigma_i, (i \geq r)$ によって定義する. すると, これらのベクトルから作られるマトリクス $U = (u_i)$ と $V = (v_j)$ によって, A は, $A = U \Sigma V^T$, あるいは同じことであるが, $A = \sum_{l=1}^r \sigma_l u_l v_l^T$ と表すことができる. ここに, Σ は対角要素に σ_i を用いた対角マトリクスである. これが典型的な特異値分解法である ([11-13]).

3.2 低ランクマトリクスの生成

A_k を, 特異値分解されたマトリクス U と V の最初の k 列を用いて,

$$A_k = \sum_{l=1}^k \sigma_l u_l v_l^T, \quad (6)$$

のように作る. このときの A_k は A の低ランクマトリクスになっている.

Eckart-Young の定理 [14] によれば,

Theorem 1 (Eckart-Young).

- 1) $\text{rank}(A_k) = k$
- 2) For any $m \times n$ matrix $B, (\text{rank}(B) \leq k),$

$$\|A - A_k\|_F = \min_{B, \text{rank}(B) \leq k} \|A - B\|_F = \left(\sum_{l=k+1}^n \sigma_l^2 \right)^{1/2},$$

ここで, $\|\cdot\|_F$ は Frobenius のマトリクスノルム, つまり, $\|(a_{ij})\|_F = (\sum_{i,j} |a_{ij}|^2)^{1/2}$ である.

この定理は, A_k は, マトリクスのランクが $k+1$ よりも小さいマトリクスの中では, マトリクスノルムの意味で A の最良近似になっていることを示している.

3.3 低ランク応答マトリクス

A を観測応答マトリクス Δ , A_k を Δ_k とみなすことによって, Δ の低ランクマトリクス A_k を低ランク応答マトリクス Δ_k とみなすことができるので, SVD によって, 観測応答マトリクス Δ から低ランク応答マトリクス Δ_k を構成することができる.

4 大学数学 CBT の例: 線形代数の場合

図 1 左に, 大学数学 CBT (線形代数) の観測応答マトリクス $\Delta = (\delta_{ij})$ を示す. 実際の実験者数 n は 216 人, 問題数 m は 31 であるが, 図では受験者数は 28 人分しか見えていない. これをケース A としよう.

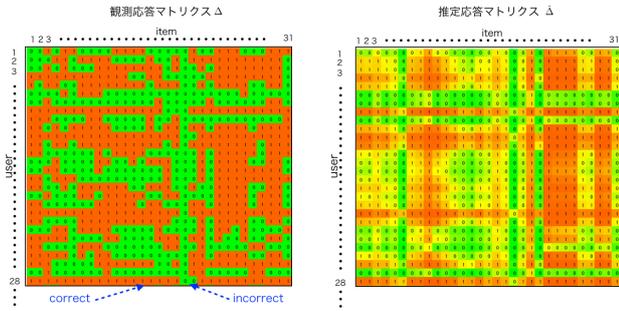


図 1: 観測応答マトリクス Δ と推定応答マトリクス $\hat{\Delta}$ [9]

4.1 IRT による推定応答マトリクス

観測応答マトリクス $\Delta = (\delta_{ij})$ に最尤推定法を適用し、パラメータ θ の推定値 $\hat{\theta}$ を求め、これを使って推定応答マトリクス $\hat{\Delta}$ を求めたものを、図 1 右に示す。先に述べたように、尤度原理からはこのマトリクスは IRT モデルのもとでは最良になっている。しかし、図 1 の $\hat{\Delta}$ と Δ を比較すると、推定応答マトリクス $\hat{\Delta}$ から観測応答マトリクス Δ をおおまかに見ることはできが、あまりよく近似されている印象はない。

これを数値的に確認するために、Frobenius のマトリクスノルムを求めてみる。同じ大きさのマトリクス $A = (a_{ij})$ と $B = (b_{ij})$ の違いを Frobenius のマトリクスノルムで表すと、

$$\begin{aligned} \text{RMSE}(A, B) &= \sqrt{\frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m (a_{ij} - b_{ij})^2} \\ &= \sqrt{\frac{1}{nm} (\|A - B\|_F)^2}. \end{aligned} \quad (7)$$

となるので、これを推定応答マトリクス $\hat{\Delta}$ と観測応答マトリクス Δ との違いを見るために計算してみると、 $\text{RMSE}(\hat{\Delta}, \Delta)$ は 0.3915 となった。このことは、観測値 δ_{ij} が推定値 $\hat{\delta}_{ij}$ と平均しておおよそ 0.3915 離れていることを示す。直感的に、この数値は小さい値ではない。

4.2 特異値分解による低ランク応答マトリクス

上に述べたように、IRT を用いた推定値 $\hat{\delta}_{ij}$ は最良な推定値であっても、推定応答マトリクス $\hat{\Delta}$ は観測応答マトリクス Δ を正確に再現しているようには見えない。そこで、ここでは、 $\hat{\Delta}$ が SVD によって生成される低ランクマトリクス Δ_k のどのランクのマトリクスに最も近いかをマトリクスノルムの意味で調べて、それを推定応答マトリクス $\hat{\Delta}$ の観測応答マトリクス Δ への近似度として定義する。

表 1 に、 $k = 1, \dots, 5, 10, 20, 31$ のときの $\text{RMSE}(\Delta_k, \Delta)$ の値を示す。表には、 $\text{RMSE}(\hat{\Delta}, \Delta)$ の値も併記している。表から、 $\text{RMSE}(\hat{\Delta}, \Delta)$ は、 $\text{RMSE}(\Delta_1, \Delta)$ と $\text{RMSE}(\Delta_2, \Delta)$ の間に位置していることがわかる。 Δ のランク数は 31 であ

るにもかかわらず、推定応答マトリクス $\hat{\Delta}$ はわずかランク数 2 程度の低ランクマトリクス Δ_2 よりも観測応答マトリクス Δ への近似度が悪い。つまり、IRT 推定応答マトリクスの観測応答マトリクスへの再現度は高くはなさそうということが確認できる。このことは、IRT モデル下での最尤推定量から作られる応答マトリクスは最良であるという観点からは衝撃的な結果である。

図 2 に、 $k = 2$ と $k = 10$ のときの SVD から作られる低ランクマトリクスを示す。 $k = 2$ のときの低ランクマトリクス Δ_2 は図 1 の推定応答マトリクス $\hat{\Delta}$ とよく似ており、 $k = 5$ のときの低ランクマトリクス Δ_{10} は図 1 の観測応答マトリクス Δ をより近似していることがわかる。

表 1: 観測応答マトリクス Δ と低ランクマトリクス Δ_k の差の RMSE[9]

k	$\text{RMSE}(\Delta_k, \Delta)$	$\text{RMSE}(\hat{\Delta}, \Delta)$
		0.3915
1	0.4066	
2	0.3851	
3	0.3652	
4	0.3479	
5	0.3306	
10	0.2562	
20	0.1325	
31	0	

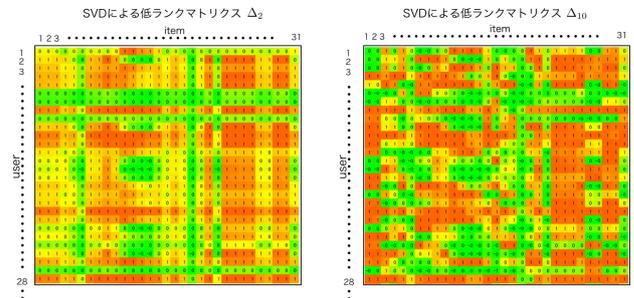


図 2: 低ランク応答マトリクス Δ_2 と Δ_{10}

しかしながら、これは、ある大学の線形代数のテストに CBT を用いた結果の一例であり、上に述べたような性質が他の場合にも成り立つかどうかはわからない。そこで、他の CBT の結果についても同様のことを調べてみることにした。

4.3 42 ケースの CBT の結果

ここでは、表 2 に示す 42 ケースの科目の CBT 結果での観測応答マトリクスについて、推定応答マトリクスと低ランク応答マトリクスの違いを調べてみた。42 ケースにはケース A も含まれている。受験者の対象は大学 1-2 年生で、問題の科目は、確率論、統計学、常微分方程式、微積分、線

形代数などである。すべての問題は多肢選択式で回答されるようになっており、正当なら1を誤答なら0を観測応答マトリクスデータベースに返すようになっている。

表 2: 42 ケースの科目とマトリクスサイズ [15]

id	科目	n	m	id	科目	n	m
1	PS	44	14	22	LA	132	84
2	PS	41	19	23	LA	177	49
3	PS	34	17	24	LA	142	45
4	P	97	32	25	LA	46	39
5	S	57	15	26	LA	39	45
6	S	75	14	27	LA	181	45
7	C	40	19	28	LA	229	84
8	PS	44	15	29	C	1131	77
9	PS	72	21	30	LA	1101	84
10	ODE	41	13	31	C	215	36
11	ODE	49	25	32	LA	47	39
12	PS	54	21	33	C	209	36
13	C	70	26	34	C	868	6
14	C	9	16	35	C	215	36
15	C	45	42	36	LA	585	84
16	LA	36	39	37	LA	39	39
17	LA	566	6	38	C	209	31
18	C	66	33	39	C	209	67
19	S	97	12	40	C	216	31
20	C	76	30	41	LA	585	49
21	LA	132	49	42	C	145	34

PS: 確率・統計, P: 確率論, S: 統計学,
ODE: 常微分方程式, C: 微積分, LA: 線形代数

図3に、42 ケースの $RMSE(\hat{\Delta}, \Delta)$ を示す。図からわかるように、横軸の id は縦軸の $RMSE(\hat{\Delta}, \Delta)$ が昇順になるように順序付けされている。図には、 $RMSE(\Delta_2, \Delta)$ も併記しているが、これを見ると、すべての id について、 $RMSE(\Delta_2, \Delta) < RMSE(\hat{\Delta}, \Delta)$ となっていることが確認できる。つまり、IRT の推定応答マトリクスは SVD から作られる低ランクマトリクス ($k=2$) よりも観測応答マトリクスを近似していないということは、上の線形代数の例が特殊な例ということではなく、他の科目の CBT でも普通に見られる現象であるということを表している。

しかしながら、この結果は、 Δ の要素すべてを用いたときの推定結果である。機械学習の言い方をすれば、すべてのデータをトレーニングデータとして用いたときの結果であり、過学習が含まれているかもしれない。予測の観点からは、 Δ の要素をトレーニングデータとテストデータに分けた処理法のもとで、テストデータを用いた結果がどのようなになるかが重要になる。

5 テストデータを用いた結果

マトリクス Δ と同じ大きさのマトリクスをトレーニングデータ S とテストデータ T に分ける。テストデータには Δ の要素からランダムに 10% の要素を抽出し、残りの 90% をトレーニングデータとする。トレーニングデー

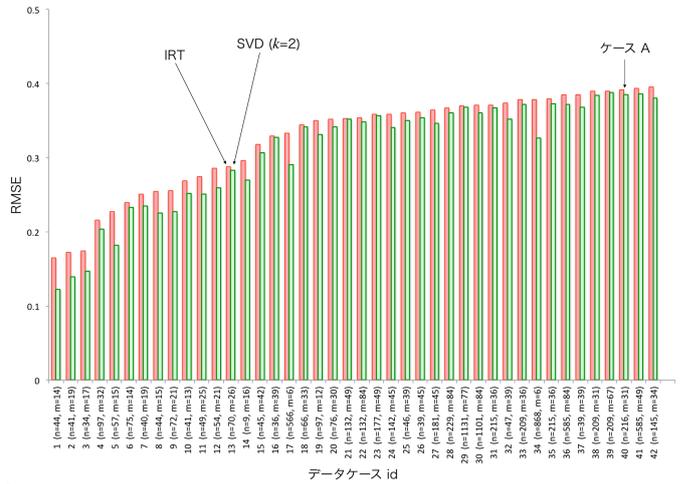


図 3: 42 ケースの $RMSE(\hat{\Delta}, \Delta)$ と $RMSE(\Delta_2, \Delta)$

タ、テストデータによって再構築（再構築法は [15] に示す）された応答マトリクスを、IRT での推定応答マトリクスの場合、 \tilde{S} , \tilde{T} で、SVD でのランク k の低ランクマトリクスの場合、 \tilde{S}_k , \tilde{T}_k で表す。このトレーニングデータ、テストデータをランダム抽出する操作を 10 回繰り返し、それぞれの $RMSE(\tilde{S}, S)$, $RMSE(\tilde{T}, T)$, $RMSE(\tilde{S}_k, S)$, $RMSE(\tilde{T}_k, T)$, およびそれらの平均、 $\mu(RMSE(\tilde{S}, S))$, $\mu(RMSE(\tilde{T}, T))$, $\mu(RMSE(\tilde{S}_k, S))$, $\mu(RMSE(\tilde{T}_k, T))$ を求める。

5.1 ケース A : 線形代数

図4に、先のケース A（線形代数）について、IRT から求められた $RMSE(\tilde{S}, S)$ と $RMSE(\tilde{T}, T)$, および、 $k=1$ から $k=10$ のときの SVD から求められた $RMSE(\tilde{S}_k, S)$ と $RMSE(\tilde{T}_k, T)$ を示す。

図を見ると、トレーニングデータでの RMSE は想定されるとおり k が増加するにつれて単調に減少していつている。また、RMSE のゆらぎは小さく、すべて $RMSE(\tilde{S}_1, S)$ と $RMSE(\tilde{S}_2, S)$ の間にある。このことは Δ のすべての要素をトレーニングデータとして用いたときの性質と変わらない。

一方、テストデータでの $RMSE(\tilde{T}_k, T)$ は、これも想定されるとおり、 k が小さいときは k の減少関数に、大きくなると増加関数になる V 曲線 [16] を描いている。ブートストラップによるそれぞれの $RMSE(\tilde{T}_k, T)$ 曲線が最小になるのは、 $3 \leq k \leq 5$ のときであるように見えるが明瞭ではない。

そこで、表3に、10 回のブートストラップから得られた $RMSE(\tilde{T}_k, T)$ の平均 $\mu(RMSE(\tilde{T}_k, T))$ を $k=1, \dots, 10$ の場合について示してみた。表には、IRT から作られる推定応答マトリクスのテストデータにおける $\mu(RMSE(\tilde{T}, T))$ の値も併記している。表から、 $\mu(RMSE(\tilde{T}_k, T))$ が最小になる k の値は $k=5$ であることがわかる。表ではこのときの値を太字で示している。

このことから、 $\text{RMSE}(\tilde{T}_k, T)$ は SVD から作られる低ランクマトリクスのランク数を上げてモデルの複雑度を増したとしても、ある一定の RMSE 値以下にはならないことが示されている。しかも、そのランク数は小さい。また、V 曲線では最小になる $\text{RMSE}(\tilde{T}_k, T)$ の値 0.3985 は、 $\mu(\text{RMSE}(\tilde{T}, T))$ の値 0.4056 の値に近いこともわかる。つまり、SVD から作られる低ランク応答マトリクスの複雑度を観測応答マトリクスにいくら近づけても、IRT による推定応答マトリクスから得られる予測精度を大きく上回ることはなく、ほぼ同等な精度を示している。

言い方を変えると、IRT による推定は、予測の意味でも最良な状態に近いことがわかる。しかし、これも、ケース A だけに見られる現象かもしれない。そこで、次に、42 ケースの中から適当に 8 個のケースを抽出して、ケース A と同様な特徴がみられるかどうかを検証してみる。

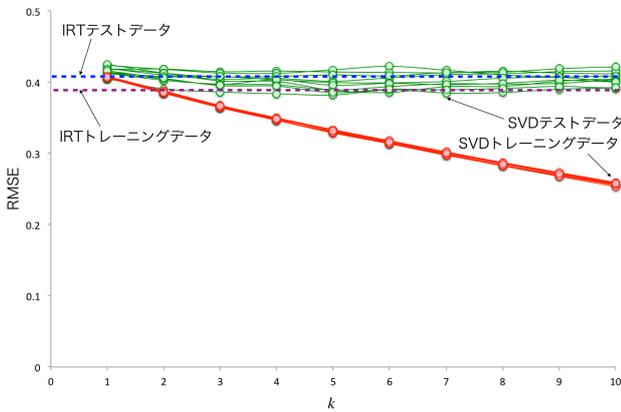


図 4: トレーニングデータおよびテストデータでの RMSE (ケース A)[15]

表 3: 10 回のブートストラップから得られた $\text{RMSE}(\tilde{T}_k, T)$ の平均 $\mu(\text{RMSE}(\tilde{T}_k, T))$ [15]

k	$\mu(\text{RMSE}(\tilde{T}_k, T))$	$\text{RMSE}(\hat{\Delta}, \Delta)$
		0.4056
1	0.4168	
2	0.4085	
3	0.4026	
4	0.4025	
5	0.3985	
6	0.4009	
7	0.4019	
8	0.4025	
9	0.4048	
10	0.4052	

5.2 42 ケースの中の 8 ケースでの結果

42 ケースの中から、ケース A を含む 8 個のケースについて、上と同様なことを調べてみる。表 4 に、

$\mu(\text{RMSE}(\tilde{T}_{k_{opt}}, T))$ と $\mu(\text{RMSE}(\tilde{T}, T))$ を示す。ここに、 k_{opt} は、 $k = 1, \dots, k_{max}$ の中で $\mu(\text{RMSE}(\tilde{T}_{k_{opt}}, T))$ が最小になる k を表す。

表を見ると、ほとんどのケースで、 $\mu(\text{RMSE}(\tilde{T}_{k_{opt}}, T))$ は $\mu(\text{RMSE}(\tilde{T}, T))$ に近く、しかもそのときの k_{opt} の値は小さい。例外はケース B (id は 30) で、マトリクスの大きさは $n = 1101, m = 84$ のときで、受験者数がかなり多いときで、 $\mu(\text{RMSE}(\tilde{T}_{k_{opt}}, T))$ は $\mu(\text{RMSE}(\tilde{T}, T))$ よりも明らかに小さくなっている。そこで、次に、ケース B について、 $\text{RMSE}(\tilde{T}_{k_{opt}}, T)$ と $\text{RMSE}(\tilde{T}, T)$ の平均値ではなく、個々のブートストラップでの両者の関係について更に詳細に調べてみる。

表 4: 8 ケースでの 10 回のブートストラップから得られた RMSE の平均 [15]

id	$\mu(\text{RMSE}(\tilde{T}_{k_{opt}}, T))$	k_{opt}	$\mu(\text{RMSE}(\tilde{T}, T))$	ケース名
5	0.2935	1	0.2908	case B
10	0.3547	1	0.3591	
15	0.3457	1	0.3344	
20	0.3801	1	0.3726	
25	0.3701	5	0.3789	
30	0.3442	16	0.3771	
35	0.3982	2	0.3964	
40	0.4053	3	0.4068	case A

図 5 に、SVD を用いたときのトレーニングデータでの結果とテストデータでの結果を示す。図で、丸印は各ブートストラップでの $\text{RMSE}(\tilde{S}_k, S)$ と $\text{RMSE}(\tilde{T}_k, T)$ を示す。参考のため、図には、 $\mu(\text{RMSE}(\tilde{S}, S))$ と $\mu(\text{RMSE}(\tilde{T}, T))$ の値を水平な点線を用いて示している。ここでは、 $k = 1, \dots, 30$ のときについて示されている。

図から、 $\mu(\text{RMSE}(\tilde{T}, T))$ は $\text{RMSE}(\tilde{T}_1, T)$ と $\text{RMSE}(\tilde{T}_2, T)$ の間に位置していることがわかる。また、 $\mu(\text{RMSE}(\tilde{S}, S))$ は $\text{RMSE}(\tilde{S}_1, S)$ と $\text{RMSE}(\tilde{S}_2, S)$ の間に位置することもわかる。この性質は、1 ケースのトレーニングデータを用いたときの $\text{RMSE}(\tilde{S}, S)$ が $\text{RMSE}(\tilde{S}_1, S)$ と $\text{RMSE}(\tilde{S}_2, S)$ の間にあった性質と同じである。このような結果は偶然に起こったのかもしれない。しかし、 $\mu(\text{RMSE}(\tilde{T}_k, T))$ が $\mu(\text{RMSE}(\tilde{T}, T))$ とはそれほど異なっていないという特徴は、他の 7 ケースと同様、保たれている。

別の言い方をすれば、IRT が観測応答マトリクスを近似する推定応答マトリクスが、観測応答マトリクスの SVD での、ランク数が 2 とかその程度の低い、低ランク応答マトリクスに対応しているのは、トレーニングデータだけの帰結に限らず、トレーニングデータ・テストデータを用いたときにも同じような結果になるということである。推定応答マトリクス $\hat{\Delta}$ は、観測応答マトリクス Δ のランクと比較してかなり低いランクの低ランクマトリクスと同等になってはいるが、 $\text{RMSE}(\tilde{T}_{k_{opt}}, T)$ が $\mu(\text{RMSE}(\tilde{T}, T))$ と同

程度になるため、IRT の予測能力はかなり高いと言ってもよいと考えられる。IRT の受験者数が 1000 人以下で、問題数も 100 以下のような比較的中規模の項目応答マトリクスであれば、観測応答マトリクスだけを使って、IRT から得られる予測情報を上回るような情報を得るのは困難であると考えられる。

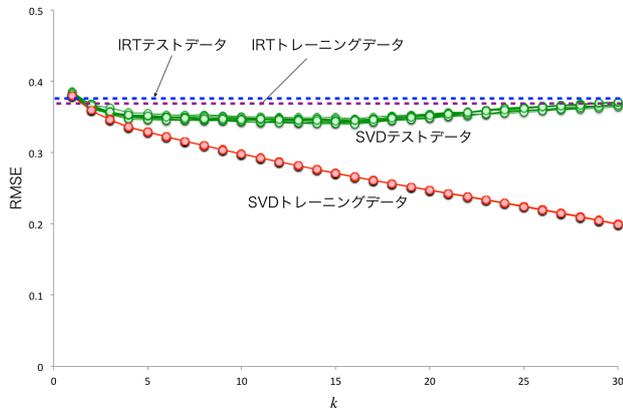


図 5: トレーニングデータおよびテストデータでの RMSE (ケース B)[15]

6 まとめ

本稿では、項目反応理論における予測の有効性を特異値分解の観点から再確認した。

観測された項目反応マトリクス（観測応答マトリクス）が与えられれば、項目反応理論におけるパラメータの最尤推定値が推定できて、その推定値を用いて項目反応マトリクスの推定値（推定応答マトリクス）を再構築することができる。これから、観測応答マトリクスと推定応答マトリクスの差のノルムを求めることができる。つまり、観測マトリクスと推定マトリクスがどれだけ近いかという、2つのマトリクス間の近似度を求めることができる。

一方、特異値分解では、観測応答マトリクスから低ランクマトリクスを生成することができ、観測応答マトリクスと生成された低ランクマトリクスの差のノルムを求めることができ、2つのマトリクス間の近似度を求めることができる。

そこで、推定応答マトリクスから得られる近似度と低ランクマトリクスから得られる近似度を比較すれば、推定応答マトリクスに近い低ランクマトリクスのランクを求めることができる。更に、推定応答マトリクスと低ランクマトリクスをトレーニングデータとテストデータに分けることで、マトリクスノルムの変動も求めることができるため、予測の意味での項目反応理論の有効性を確認できる。

実際の大学の基礎数学の CBT テストに用いられた試験データにこのことを適用してみると、推定応答マトリクスと等価な低ランク近似マトリクスのランクは非常に低くな

ることがわかった。しかしながら、特異値分解から得られた低ランク近似マトリクスと観測応答マトリクスの差のノルムの最小値は、項目反応理論から得られた推定応答マトリクスノルムと観測応答マトリクスの差のノルムとほぼ等しいことから、観測応答マトリクスを入力とする限り、項目反応理論の予測能力は十分に高いところにあると考えられる。

このことは、2値情報から成る観測応答マトリクスしか得られない状況では、2パラメータロジスティックモデルなどの項目反応理論が十分に有効であることを示している。より信頼性の高い受験者能力の推定値を得るためには、新たな数学的モデルの開発と同時に、項目応答マトリクスとは異なる新たなデータの収集が必要である。

参考文献

- [1] 統計検定受験者数の推移 (CBT 方式) [<https://www.toukei-kentei.jp/info/past/page-14294/>]; 2022.
- [2] 統計検定受験者数の推移 (PBT 方式) [https://www.toukei-kentei.jp/info/past/nun_tran/]; 2022.
- [3] 大学入学共通テスト「情報 I」の出題に関する調査研究 [https://www.dnc.ac.jp/research/cbt/cbt_houkoku.html]; 2022.
- [4] de Ayala R. The theory and practice of item response theory. Guilford Press; 2009.
- [5] Baker FB, Kim SH. Item response theory: Parameter estimation technique, 2nd edn. Marcel Dekker; 2004.
- [6] Hambleton R, Swaminathan H, Rogers HJ. Fundamentals of item response theory. Sage Publications; 1991.
- [7] van der Linden WJ. Handbook of item response theory. Chapman and Hall/CRC; 2016.
- [8] Kendall MG, Stuart A. Advanced theory of statistics. Macmillan Pub Co; 1983.
- [9] 廣瀬英雄. 推薦システム：マトリクス分解の多彩な姿 (統計学 One Point 22). 共立出版; 2022.
- [10] Hirose H. Matrix decomposition perspective for accuracy assessment of item response theory. arXiv. 2022; stat.CO:2203.03112.
- [11] Golub GH, Van Loan CF. Matrix computations. Johns Hopkins Univ. Press; 2012.
- [12] Strang G. Multiplying and factoring matrices. The American Mathematical Monthly. 2018;125:223–230.
- [13] Strang G. Introduction to linear algebra. Wellesley-Cambridge Press; 2021.
- [14] Eckart C, Young G. The approximation of one matrix by another of lower rank. Psychometrika. 1936;1:211–218.

- [15] Hirose H. Prediction effectiveness of item response theory from matrix decomposition perspective. submitting. 2022.
- [16] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Springer; 2009.

プログラミングによる統計計算の実装を演習とする初等統計学教育の試み

京都大学教育学研究科 西山慧

nishiyama.satoru.2u@kyoto-u.ac.jp

1. はじめに

本報告は、発表者が担当した心理統計学入門についての実践報告である。

心理・教育学系に所属するいわゆる「文系」学生は、数学に対する理解度が低い、あるいは苦手意識を持つ学生が「理系」学部よりも多いことが想定される。おそらくそのことを反映して、心理・教育学系の統計学の講義では、説明の際に数式を示すことはあっても、演習では既存のソフトウェアを利用し数式・計算手順を意識せずに統計計算を実施できるようにしていることが多いという印象がある（調査などの結果に基づくものではなく、発表者が書籍やオンライン上の資料を確認した範囲での個人的な印象である）。

ソフトウェアは数ステップで目的の分析に関する結果を出力する。その利用方法に習熟することは実用的である。一方で、その操作自体に統計計算の理解は必要ではないため、ソフトウェアを用いた演習は統計学の理解にあまり貢献しないと発表者は考える。なお、プログラミング言語を用いれば良いというわけではない。すでにパッケージ化された関数を利用する場合は、ソフトウェアのボタン操作が関数名の記述に代わっただけであり、同様に理解には寄与しないだろう。

以上の理由から、発表者は自身が担当した統計学入門の講義において、統計量の計算式や推定・検定の手順をすべて 1 からプログラミング（具体的には Python）で実装するという演習を実施した。数値例に対して正しい計算結果を求めるのではなく、その計算結果を出力するプログラムを書く

ことに主眼を置いた。受講生が自身の手で計算手順を実装することで統計計算をブラックボックスとしないこと、そして、数式をコードに変換することを通して、数学記号の理解向上・苦手意識の低減を期待した。

2. 実践内容

講義の概要 心理・教育学系の大学 1 年生を対象とする心理統計学入門の講義で、受講生は 28 名であった。講義内容（全 15 回、各回 90 分）は表 1 に示した。心理学系の大学生を対象とした統計学入門の講義としては一般的な構成である（例えば、清水, 2021）。初回を除き、各回の講義内容に関する Python を用いた演習を行った。演習は全て Google Colaboratory 上で実施した。講義は情報端末室で実施され、受講生は持参した PC か大学が提供する共有 PC を用いて授業内の演習を行なった。

表 1. 講義の内容

回	内容
1	統計学の導入
2	Python の演習（四則演算・変数）
3	Python の演習（リスト・for 文）
4	Python の演習（関数）・尺度水準
5	データの可視化
6	1 変数の統計量（平均・標準偏差）
7	2 変数の統計量（共分散・相関係数）
8	相関係数の注意点（偏相関係数など）
9	推定 1：確率・標本抽出
10	推定 2：点推定
11	推定 3：区間推定
12	推定 4：差の推定
13	検定 1：差の検定（t 検定）
14	検定 2：サンプルサイズ設計
15	振り返りと今後に向けて

講義の進め方 第 2~4 回では統計学に関する講義は行わず、以後の演習に最低限必要な Python の基礎事項に関する説明・演習を行なった。これは、どの受講生も Python を用いたプログラミングを経験したことがなかったことを考慮してのことである。

第 5 回以降の講義は各回三部構成で実施され（前回の演習の解説、新たな内容の講義、講義内容に関する演習）、それぞれに用いた時間は概ね 30 分ずつであった。演習の解説では、受講生全員から見える位置にあるスクリーンに講師の画面を表示し、口頭で説明しながらコードを書いていくという形式をとった。講義のパートではスライド資料を使用し、内容の説明を行なった。最後の演習時間では、受講生は演習問題に取り組み、講師は講義室を巡回し質問対応を行った。教員は受講生同士の会話を許可し、お互いに相談し教え合いながら取り組むことを勧めた。受講生は講義から 1 週間以内に演習を提出するように求められた。また、講義内容に関するアンケートも提出した。演習課題とアンケートについては以下に詳しく説明する。

演習課題 演習課題は標準問題と発展問題の 2 つで構成された。受講生は少なくとも標準問題を全て解くように求められた。どの問題でも基本的には、計算結果が提示されており、それを出力するコードを書くことが求められた。また、指定された場合を除き、既存のパッケージを利用せず、Python の基本的な文法のみを用いて、統計計算（平均や標準偏差を計算する関数を含む）を実装させた（具体的な課題は以下に示すリンク先から確認してほしい）。

受講生にいきなり一から統計計算を実装させるのではなく、計算手順を分解し目標の計算を少しずつ実装できるように配慮した。また、第 7 回以降の演習問題では毎回、平均や標準偏差などの基本的な統計量を計算する関数を必ずはじめに実装させ、知識の定着を期待した。

提出された演習課題は次の講義までに講師が確認し、フィードバックを行なった。具体的には、全体の総評、ミスや特筆すべき箇所への言及、アンケートに記載されたコメントへの返答を行なった。

演習に使用したファイルは全て [GitHub `https://github.com/snishiyama/intro-psych-stats-exercises`](https://github.com/snishiyama/intro-psych-stats-exercises) で公開されている。

アンケート 演習課題の他に、受講生は毎回講義内容に関する簡単なアンケートを提出した。アンケートは 5 つの項目で構成されていた。具体的には「講義に関するコメント」「楽しく授業を受けることができたか」「講義内容は理解できたか」「講義内容を他人に教えることはできそうか」「講義内容は難しかったか」であった。最初の項目は自由記述での回答を求め、それ以外では、1（全くそう思わない）~5（とてもそう思う）の 5 段階で回答させた。

期末課題 最終回である第 15 回では、講義内容全体を復習する最終課題を課した。その問題の一つとして最後に、受講生が講義全体を通じた自身の学びの振り返りを求めた。具体的には、毎回提出したアンケートのうち、数値として収集したものについては図示させ、それらとコメントを参考に、理解度や授業に取り組む姿勢の変化、今後への課題について記述させた。

3. 結果

本講義では試験や小テストは実施しておらず、学生の理解を数量化することはできてない。また、提出された演習課題も、受講生同士でコードを写すことがあった（これを講師は認めている）ため、受講生の理解を評価するには適さない。

そこで本稿では、毎回のアンケートと期末課題において得られた自由記述を参考に、受講生の講義・統計学に対する態度を推察し、プログラミングによる統計計算の実装

を演習させた本講義における学生の学びを概観することとした。

講義内で収集した学生のコメントからは、当初プログラミング自体に困難を抱えていた受講生らが、毎回の課題に取り組む中で、プログラミングのコツを掴み、独力で演習問題を解けるようになったという様子が多数見受けられた。それらの学生は、演習が解けるようになってるとともに統計学に対して前向きな姿勢を持つようになった。また、実装の際、数式が示す計算手順を言語化することもあった（これはコメントではなく授業中に観察された様子である）。

期末課題の振り返りでも、今後の統計学の学習に対する高い意欲を示す記述が多かった。その中には元々数学に対する苦手意識を持っていたと述べる者もいた。数名に限られるが、数式をヒントにコードを組み立てるようになったと報告する者もいた。

一方で、数名の受講生は、数式とプログラミングのどちらの理解にも困難を感じてしまい、期末課題のコメントでは統計学に対する強い拒否反応を示す記述が見られた。また、今後の意欲を示した受講生からも、推定や検定についてはあまり理解できていないという報告が散見された。

4. 考察

以上の結果から、本講義は、統計学への苦手意識を生じさせづらく、意欲的な姿勢を保ちやすい設計になっていたと考えられる。隣接する回の内容は関連が強く、ある回の演習が一つ前の内容の復習を含むことが多かったため、自身の理解を確認したり実感したりする機会が確保されていたことがその理由かもしれない。もちろん、講義でも復習は行うが、問題を解くためには知識を運用する必要があるため、聞いているだけよりも演習に取り組むほうが自身の理解を確認するきっかけとなったのだろう。

理解の確認を促すという点では、手計算による演習でも同様であるだろう。しかしながら、プログラミングの場合は計算自体を計算機が実行するため、計算に必要な時間的・認知的なコストが取り除かれる。したがって、受講生は計算手順を考え・試行錯誤することにより時間をかけることができると考えられる。

数式とコードを対応づける受講生がいたこともプログラミングで演習を行なったことの重要な成果であるだろう。数式を理解するための手立てが増えることは、今後の統計学の理解に資するに違いない。

これら良い面の一方で、数名は講義についていけなかった。改善点として、講義内で数式とコードを対応づけた説明を増やすことが挙げられる。数式などの計算手順を講義内で積極的に言語化することでコードを書きやすくするかもしれない。

また、推定や検定を十分理解できなかった学生がいたことも本実践の改善の必要性を示している。実際、プログラミングによる実装は統計計算については扱いやすいが、背後にある考え方までは取り上げづらい。説明により時間をかける必要があると思われるが、実際に実験・調査を行うことも理解に重要だと思われるため、カリキュラムを通して扱っていく必要があるだろう。

5. 今後の課題

本実践に関して、学生の報告から明らかとなった課題点については上述した通りである。ここでは、それら以外の課題について2点取り上げる。

まず、本形式の実践成果を他形式の講義のものと比べる必要がある。発表者自身は、「はじめに」で取り上げた講義よりも効果的であると信じているが、実際のところは不明である。

次に、実践の効果の評価方法も重要な問題である。これは1点目とも関連する。筆

記試験で知識を問うのか，与えられたデータセットに対して妥当な分析手法を行えるかのようなパフォーマンス評価を行うのか。個人的には，運用できることも重要であるが，必要に応じて統計学を継続的に学べるかどうかを重視したい。そのような評価方法を模索していきたい。

6. 引用文献

清水裕士（2021）心理学統計法 放送大学教育振興会

授業で使うデータに関する試みと提案

島谷健一郎（統計数理研究所）

立川市緑町 10-3, 050-5533-8590, shimatan@ism.ac.jp

統計学の授業において、どのようなデータを用いるかは、小学校から大学、社会人に至るまで、一貫して教員を悩ます問題である。

1. 書籍やインターネットで公開されているデータ。
2. 人工的に作った数値。
3. 授業時に学生自身で調査をしてデータを作るところから始める。

2は統計手法の演習には便利だが、結果を見てデータの背後にある現象について考察することができない。

1では、なるべく多くの学生が関心を抱いてくれそうなデータを探すことに教員は腐心する。しかし、サッカーや野球などの人気スポーツでも、ペンギンやクマなどよく知られた動物でも、関心を示さない学生は少なからず出てしまう。

3は一見、最も好ましいように見える。しかし、いざ調査から始めるとなると、その調査デザインにおいて、学生も教員もおおいに悩むところから始めるのが本筋で、統計手法の学習までには長い道のりがある。また、教員（または学生）に調べる対象に関する調査方法についてある程度の素養がないと、調査を始められないか、実践してもいい加減なデータを得る事態となる。

本発表者は、当初は自らが野外で収集した生き物のデータ（上記1）を使っていた。しかし、この数年は2の人工データを使っている。それは、やり方によって、1や3と同等かそれ以上に、統計解析結果を踏まえた考察の練習もできると考えるからである。

大学生が最初に学習する統計手法として、説明変数が連続的、従属変数が0-1のデータに対するロジスティック回帰モデルを選んでいる。授業では、0-1データという概念から始めている。その好例が大学の合否を0-1とするものである。そして、模擬試験の点数を説明変数とすると合格可能性はまさしく0-1データの予測モデルである。

模擬試験の点数	合否	合格=1, 不合格=0
65	合格	1
41	不合格	0
23	不合格	0
37	不合格	0
55	合格	1
24	合格	1

他にも身近に0-1データはたくさん見られる。そこで、以下の作文とロジスティック回帰の計算練習を並行して進めている。

作文課題

1. 0か1で表されるデータの例を考えてください。

- 連続的な数値をとりうる x が変化すると、0 と 1 しかとらない y は 0 が多いから 1 が多いに变化するような例を考えてください。
- そんなデータについてロジスティック回帰モデルを適用するとどんな知見（予測）が得られるでしょう。

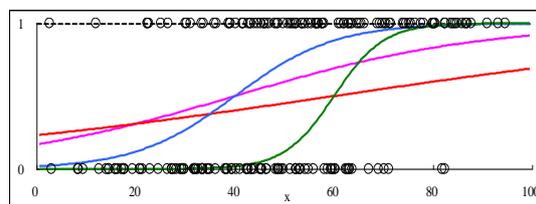
インターネットや書籍でみつけたデータでもいいし、想像して作文してもかまわない。計算練習は最尤法によるパラメータ最適化だが、結果として出てくるロジスティック曲線からわかってくることを、作文例を題材に考察する。作文は（氏名以外）教室で公開し、講評も行う。

作文は、想像でかまわないとしているのもあって、わりと自然に学生の関心ごとを反映しているのではないか。

- x 軸に新型コロナウイルスに感染した患者の年齢をとる。そして y 軸に患者が重症化(人工呼吸器を必要とする状態)した場合を 1、重症化しなかった場合に 0 とする。

- サッカーの試合における枠外シュートの数（変数 x ）と勝敗の関係（勝利 = 1、敗北 = 0）一般にサッカーの試合において、枠内シュートを打たないとゴールは決まらない。積極的にシュートを打つことを求められるが、枠外シュートは試合の勝敗に影響があるかを調査するために利用できる。この調査を一年間、様々な国で行われた試合で実施する。この調査において、正の相関が見られれば枠内、枠外を問わずシュートを打つことは試合の勝利につながる。だが、この 2

変数の相関が見られなければ、枠外シュートを打つことで試合の勝敗に影響せず、むやみにシュートを打つチームが減ると考察できる。



例えば、以下のようなスマホに関する作文は毎年、数名の学生から出ている。

- x をスマートフォンやゲームなど画面を見る時間とおき、メガネやコンタクトをしていたら 1、していなかったら 0 として調べる。画面を長く見ていればメガネ率は高いと考えられる。

好んでスマホ漬けの生活を好んでしているように見える学生たちだが、実は自身の目の健康管理にも関心と危機感を抱いている学生も少なくないようである。

0-1 データを考えているが、ロジスティック回帰より、説明変数と従属変数を入れ替えて t -検定などが望ましい作文もある。

よく小中学生の頃には、「朝ごはんを食べないと頭が回らなくなり、集中できなくなる」との話を聞きましたが、毎朝朝食を食べる生徒と、食べない生徒と、実際に取ったテストの点数には本当に相関はあるのだろうか。 x 軸にテストの点数、 y に朝食の有無で関数を取りたい。

作文の講評は、データの科学の視点に限ることを意識している。社会

問題を取り上げる作文も数多く見られるが、社会批評は控えている。

・女性の肌の露出度(x)と痴漢の被害(被害に遭った=1 被害に遭わなかった=0)

女性が性被害に遭うと女性の危機管理能力の欠如を責められる二次被害が多い。

「露出の多い服装だから被害に遭う」という主張と「地味でおとなしそうな服装のほうが狙われやすい」という主張が論争を繰り広げている。女性の肌の露出度をパーセンテージで表し、被害の有無を調べることができれば、この論争は終わり、二次被害もなくなるのではないか。

調査デザインの問題として格好の題材である。こうした問題では、被害女性について調べ、共通点を探す中から対策を論じる場合が多い。しかし、データとして分析するなら、被害に遭っていない女性との比較

(対照区の設定)が望ましい。ところが、被害に遭っていない女性は数も多いし状況も多様で、はて、どのように抽出したらよいか。

実は同じ問題が生態学の野外調査で頻繁に見られる。希少種を保護するべく、その営巣地を調べ、共通点を整理し、それに基づいて保護区を定める。しかし、営巣していない場所との比較こそが肝要であり、それを伴わない調査は科学調査に値しない。ところで、地球の大半の地域で特定希少種は営巣していない。では、営巣していない場所はどうか抽出するか(ちなみに生態学では一つの

解決策が2010年台になって確立されている)。

作文講評と称しているが、作文の良し悪し評価ではなく、実際のところ、データ・統計・調査デザインに関する解説をするネタ元として作文を利用している。この世界にどんなデータがあるか(ありそうか)。そして統計分析法が決まっている場合などめったになく、統計解析経験を豊富に持っていて、新しい課題、新しいデータに対峙したら、たいていの場合、まずじっくり考える所から始めざるを得ない。こうした現実を、大学生が聞いても聞くだけで通じないことは承知の上で、とにもかくにも1度聞かせている。

受講者が10名なら10、40名なら40の事例について、調査デザインや統計解析法の厄介さに関する解説を聞かせられる。これは、人工データのひとつの大きな利点と考えられる。

ひとつの課題とデータに絞って統計解析結果を考察したり調査デザインから始める冒頭1や3のデータと比べ、人工データは劣るのか有益なのか。本発表は、この問題を提起して終わる。

高校生の探究活動の方法としての統計

橋本三嗣・広島大学附属中・高等学校
〒734-0005 広島市南区翠一丁目1番1号
TEL.082-251-0192 FAX.082-252-0725
E-mail: mhashimo@hiroshima-u.ac.jp

1. はじめに

現在、小中高の各段階において新たな「学習指導要領」への移行が段階的に進められており、そこで特に注目されるキーワードが「探究」である。教科学習では、各教科の固有の知識や個別のスキルを学び、各教科の本質に根ざした問題解決能力や学び方、考え方などを育むのに対して、探究学習では、教科にとらわれない、横断的、総合的な問題解決の能力を育む。2022年度高等学校入学生より、従来の「総合的な学習の時間」は「総合的な探究の時間」と名称を変えて実施する。そこでは生徒が自ら課題意識を持って、その意識を連続発展させていくことを重視し、「課題の設定」、「情報の収集」、「整理・分析」、「まとめ・表現」を繰り返しながら進める。また新たに設定された「理数探究基礎」、「理数探究」ではこれまでのスーパーサイエンスハイスクール（以降「SSH」と記す）校の取り組みを参考として課題研究を行うことになっている。いずれの授業においても統計分析を行う機会が従来よりも増えるといえる。

広島大学附属高等学校では2003年にSSH事業に採択され、それ以降、学校設定科目として継続して統計分析の手法を学ぶカリキュラムを開発し、実施している。それは高校生が課題研究で統計的な手法を用いることができるようになることを目的にしている。

本稿は、高校生の課題研究で統計的な手法を用いたグループ（通称：メイクチーム）の事例を報告する。

2. 課題研究計画（高1の3学期、生徒が作成）

(1) 動機・目標

オシャレを研究したい。流行りのメイクについて、数値化して傾向を明らかにし、これから流行るメイクを提案する。

(2) 研究計画

メイクに関する先行研究を概観し、対象と状況を限定してメイクの選好傾向を調べる。日本の歴史的な変

遷を調べるとともに、海外の高校生と研究交流を行い、類似点と相違点を調べる。

【高2：4月～6月】

分析する女優の写真を選び、部位に関する位置データの収集

【高2：7月～9月】

収集したデータをもとに、傾向の分析

【高2：11月～12月】

分析結果をもとに、メイクのテンプレートを作成

【高2：1月～2月】

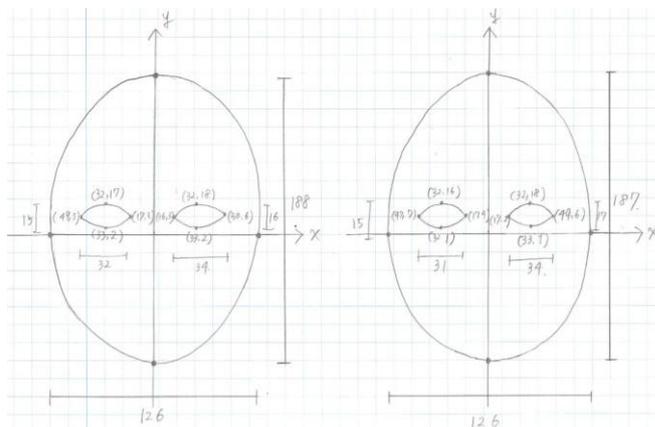
雑誌の表紙写真をもとに、女優のメイクの肌と唇の色から色見本を作成

3. 数値化の方法（高2の1学期）

①y軸を設定

②y軸をもとにx軸を設定

③目頭、目尻等の座標を計測



測定誤差を考慮して1枚の写真を3回測定して平均値を算出した。

4. 課題の焦点化（高2の2学期）

人によって顔の形や縦横の長さが異なるため、メイクのテンプレートを作成する際に、女優の写真の顔の縦横の比率（比の値）が変わっても使用できるように工夫した。また肌と唇の色から色見本を作成する際には、アプリを使用した。

5. 生徒の作成したポスター（高2の3学期）

広島大学附属高等学校
女優のメイクから美の傾向を探る

発表者 指導教員

研究目的

生来の顔を変えずにメイクでなりたいたい顔を作るための技法を見つけた

先行研究から研究内容と方法を焦点化

日本の有名な女優である8人の女優に注目し、正面からの写真画像から眉と目の位置や形に注目して数値化しその特徴を調べる

結果と考察

(1)眉の形について
図2の結果から太めの眉の形をしている女優は少ない

女優	眉の形	数
A: 宇野実里子	太めの眉	2
B: 宇野実里子	細めの眉	1

図2

(2)目の形について
表1 分析結果(名前が敬称略)

女優	縦幅(比率)	横幅(比率)
宇野実里子	0.1295	0.2428
宇野実里子	0.1040	0.2318
宇野実里子	0.1146	0.2329
宇野実里子	0.0822	0.2259
宇野実里子	0.1038	0.2474
宇野実里子	0.1109	0.2407
宇野実里子	0.1131	0.2398
宇野実里子	0.1023	0.1953
宇野実里子	0.1148	0.2287

研究方法

(1)眉の形について
図1をもとに研究班の二人が該当する眉の形を分類

(2)目の形について
女優の写真を見ながら探す

- ①y軸を設定
- ②y軸をもとにx軸を設定
- ③座標を算出

図1 From: <https://aquadonimg.jp/mch011>

結果の活用

結果にもとづいて
Beauty Makeup Templateを作成

今後の展望

分析対象を増やし肌の凹凸や色について調べる。またメイクをしていない状態からの変化についても調査、分析したい。

広島大学附属高等学校
女優のメイクから美の傾向を探る

発表者 指導教員

研究1

日本の有名な女優である8人の女優に注目し、正面からの写真画像から眉と目の位置や形に注目して数値化しその特徴を調べる

方法
(1)眉の形について
図1をもとに研究班の二人が該当する眉の形を分類

(2)目の形について
女優の写真を見ながら探す

- ①y軸を設定
- ②y軸をもとにx軸を設定
- ③座標を算出

結果
(1)眉の形について

女優	眉の形	数
A: 宇野実里子	太めの眉	2
B: 宇野実里子	細めの眉	1

(2)目の形について
左目の縦幅の比率 0.1091
右目の縦幅の比率 0.1061
左目の横幅の比率 0.2363
右目の横幅の比率 0.2373

テンプレートの作成
Beauty Make-up Template

研究2

雑誌の表紙を飾る女優のメイクに着目し、月ごとの肌の色と唇の色を調べる

方法
①3種類の雑誌の2022年の表紙の画像を集める(無料アプリGIMPを使用)
②RGB(光の三原色)におけるそれぞれの要素を抽出(数値化)
③それぞれの数値の平均値を算出
④色見本を作成

結果(制作した色見本)

肌

1月	2月	3月	4月	5月	6月
7月	8月	9月	10月	11月	12月

唇

1月	2月	3月	4月	5月	6月
7月	8月	9月	10月	11月	12月

考察
・月ごとで色味に違いがある
・冬は比較的薄い色のリップを使用している
・6月、7月のリップはBlueの要素が少なく、黄色ベース(Red、Greenの数値が高い)である
・肌(ファンデーション)の色と唇(リップ)の色に関連がみられる

今後の展望

研究1の眉の形、BMTに加えて、研究2の結果をもとに色味も考慮した月ごとのメイクを提案する

校内「課題研究中間発表会（11月）」にて発表

校内「SSHの日：課題研究発表会（2月）」にて発表

Hiroshima University High School
Mathematics Creates Beauty:
Calculating the Best Makeup

Problem

Our research purpose is to find a makeup technique to become ideal face without our face by not changing.

We found a previous work which made by University Faculty and Company that makes cosmetics. It created several makeup proposals based on images of female subjects' natural face and an application developed by the person in question to choose what she likes. It can find out women's makeup selection trends and irregularities by using it. We made contact its researcher and asked what they know and what they do not know. And we focused the research purpose and method.

we focus on eight famous actresses in Japanese and quantify their eyebrow, eye position, eye shape using their photo taken from the front. We decided to look into this trend. And compare with the actresses.

Our definition of beauty is as follows:
The shape and size of the eyebrow and eyes of the actress popular in Japan

Fig.1 Japan Fig.2 Thailand
Similarities and differences

Findings

1) Classification of eyebrow shapes
Two people made the decision based on classification, if the two judgments did not agree, each was set to 1[Table1].

Table1. Results of classification of eyebrow shape

女優	眉の形	数
A: parallel straight eyebrow	太めの眉	2
B: natural straight eyebrow	細めの眉	1

2) Quantifying eye position

Table2. Eye position results

女優	Left eye	Right eye	Width of left eye	Width of right eye
宇野実里子	0.1295	0.1040	0.2428	0.2407
宇野実里子	0.1040	0.1038	0.2318	0.2398
宇野実里子	0.1146	0.1131	0.2329	0.2347
宇野実里子	0.0822	0.1040	0.2259	0.2474
宇野実里子	0.1038	0.1040	0.2474	0.2474
宇野実里子	0.1109	0.1040	0.2407	0.2407
宇野実里子	0.1131	0.0945	0.2398	0.2347
宇野実里子	0.1023	0.1040	0.1953	0.2398
宇野実里子	0.1148	0.1131	0.2287	0.2398

Table3. Lengthwise-width(ratio)

女優	縦幅(比率)	横幅(比率)
宇野実里子	0.1295	0.2428
宇野実里子	0.1040	0.2318
宇野実里子	0.1146	0.2329
宇野実里子	0.0822	0.2259
宇野実里子	0.1038	0.2474
宇野実里子	0.1109	0.2407
宇野実里子	0.1131	0.2398
宇野実里子	0.1023	0.1953
宇野実里子	0.1148	0.2287

Table4. Width(ratio)

女優	左目の縦幅の比率	右目の縦幅の比率	左目の横幅の比率	右目の横幅の比率
宇野実里子	0.1091	0.1061	0.2363	0.2373

Framework

1) Classification of eyebrow shapes
Judged visually from the photo according to the classification(Fig.1)

- parallel straight eyebrow
- natural straight eyebrow
- feminine arch eyebrow
- natural arch eyebrow
- natural thick eyebrow
- simple arch eyebrow
- elegance arch eyebrow
- wide straight eyebrow
- wide natural eyebrow

Fig.3. Eyebrow shapes

2) Quantifying eye position
The x-axis and y-axis were set in the photo, and the position of the eye was quantified(Fig.4).

Step1: The y-axis is a straight line connecting the midpoint of the eye socket and the midpoint of the nasal wing.

Step2: Take the intersection of the y-axis, the hairline, and the chin. The x-axis is the perpendicular bisector of the two points.

Step3: Measure the length and width of the face. Coordinates for the position of the inner and outer corner of the eye. Each coordinate, etc., expressed as a ratio of height and width.

Interpretation and Conclusion

Based on the results of the analysis, we made Beauty Makeup Template(BMT). Use a photo on top of the template.

Urassya Sperbund(Fig.2) is different from the template.
Finchbrook Lucio-dopajou(alias Balforn) is similar to the template.

Fig.5 BMT

References

Fig.1. From: <https://www.instagram.com/@yuuuuukko>
Fig.2. From: <https://www.instagram.com/@urassyasj>
Fig.3. From: <https://aquadonimg.jp/mch011>

TJ-SIF2022（12月）にて発表

6. まとめ・今後の課題

高校生の関心からスタートした課題研究は、女優の写真の数値化に始まり、メイクのテンプレートの作成、色見本の作成と少しずつ進んできた。興味・関心の低下を感じたときに、どのような支援をすればよいか、手探りであったが、様々な発表の機会を通して、高校生のモチベーションを回復することができた。特にタイのICTフェア（TJ-SIF2022）の発表では、英文要約、発表動画等、限られた時間での作成になったが、タイの女優との比較を取り入れる等、楽しみながら取り組むことができた。分析の方法として使った統計は簡単なものであり、十分な考察ができていたとは言い難い。今後、論文指導を行う際に（高3）、適切なアドバイスをを行い、探究活動を深める支援を行いたい。

参考文献

- (1) 熊原啓作, 渡辺美智子 (2012) 『身近な統計』, NHK 出版.
- (2) 文部科学省 (2018), 高等学校学習指導要領 (平成30年告示) 解説 数学編 理数編.

国公立大学の2025年度入学者選抜における教科「情報」の対応調査に関するレポート

Survey of Public Universities Regarding Information Studies and Admissions for the 2025 Academic Year

下山恭平、大戸祥也（株式会社ラーズ マーケティング・営業部）



株式会社ラーズ

国公立大学の2025年度入学者選抜における教科「情報」の対応状況

1 「教科『情報』の大学入試情報まとめ」

ラーズの教育情報

新課程情報

教科「情報」の
大学入試情報まとめ



入試最新情報をラーズがレポート

新課程情報

教科「情報」の大学入試情報まとめ

2022年度から実施された高校新課程では「情報Ⅰ」が必修（全員履修）になりました。その背景には、現代社会における高度な情報技術の進展が挙げられ、「情報Ⅰ」を通じて身につける情報の科学的な理解、および情報活用能力がこれからの社会における「国民的素養」として必要なものになったといえます。

一方、大学では近年「情報」に関する学部・学科等の新設が目立ち、2023年度では一橋大がソーシャル・データサイエンス学部、名古屋市立大がデータサイエンス学部を新設予定で、私立大でも多くのデータサイエンス関連学部の新設が予定されています。また、「社会科学」や「健康・医療・スポーツ」といった従来の領域に「データサイエンス」を加えた学部の増設が目立ち、これは文系学部にもデータサイエンス系を広げるための文科省の施策が背景にあります。

最新情報

- 2023/02/21 徳島大学が教科「情報」に関する入試情報を更新しました
- 2023/02/17 敦賀市立看護大学が教科「情報」に関する入試情報を公表しました
- 2023/02/16 長野県看護大学が教科「情報」に関する入試情報を公表しました
- 2023/02/16 群馬大学が教科「情報」に関する入試情報を公表しました

最新の一覧表を取得する

125

※ダウンロードしたファイルはZIP形式で圧縮しています。解凍するとデスクトップにフォルダが作成されます。

情報

ラーズWEBサイトで
国公立大学の情報を調査し
て発信しています。

<https://www.learn-s.co.jp/edu/t/t1034/>



株式会社ラーズ

2 「教科『情報』の大学入試情報まとめ」

更新一覧

A	B	C
1	更新履歴	各大学の公表情報は下記シートにて国立・公立大学別で一覧にまとめているので是非ご覧ください！
2	更新日	更新内容
3	2023年3月7日	筑波技術大学の教科「情報」に関する入試情報を更新しました
4	2023年2月27日	公立千歳科学技術大学の教科「情報」に関する入試情報を公表しました
5	2023年2月27日	国立大学／公立大学
6	2023年2月27日	リンク先（情報掲載元）
7	2023年2月21日	https://www.chitose.ac.jp/exam/admissions_info
8	2023年2月21日	https://www.ac-nagano.ac.jp/admission/service/2023
9	2023年2月21日	https://admission.aui.ac.jp/info/677/

公表情報のリンク先
履歴と予定、特記事項など

C		D	E	F	G	H	I	J	K	L	M	N
SP1549		「情報 I」の2025大学入試関連情報（国立大学一覧）										
エリア		都道府県名	大学	公表日更新日	確度	概要、配点等	共通テスト「情報 I」の実施			個別試験での「情報 I」の実施	最新情報リンク先サイト	備考（公表の履歴・今後の予定等）
							一般選抜	総合型選抜（共通テストあり）	学校推薦型選抜（共通テストあり）			
北海道	北海道	北海道	北海道大学	2023年9月16日	予備	「令和7（2025）年度入学者選抜（一般選抜）」における実施教科・科目の予告について【入学者選抜内容】 大学入学共通テストの成績、個別学力検査等の成績及び推薦書等を総合して合格者を決定し、大学入学共通テストの成績は配点に活用する。 成績同等者の順位決定にあたっては、個別学力検査等の成績を重視し、個別学力検査等の成績と同等の場合は、大学入学共通テストの成績を活用する。	実施する			実施しない		「内容については、今後変更される場合がありますので、本学ウェブサイトや7年度入学者選抜要項及び学生募集要項必ずご確認ください。また、本内容、令和7年度入学者選抜に係るものであり、令和8年度以降も変更される可能性があります。おいて、フロンティア入試については、後日改めて公表いたします。」
北海道	北海道	北海道	北海道教育大学	2023年12月12日	決定	「令和7年度（令和6年度実施）以降の大学入学共通テストの受験を要する教科・科目について」配点は未公表。	実施する 【教員養成特選入試】 実施する 【教員養成課程、国際地域学、芸術・スポーツ文化学科芸術・スポーツビジネス専攻】 実施する 【教員養成課程、国際地域学、芸術・スポーツ文化学科芸術・スポーツビジネス専攻】 実施する 【教員養成課程、国際地域学、芸術・スポーツ文化学科芸術・スポーツビジネス専攻】	実施する 【教員養成特選入試】 実施する 【教員養成課程、国際地域学、芸術・スポーツ文化学科芸術・スポーツビジネス専攻】 実施する 【教員養成課程、国際地域学、芸術・スポーツ文化学科芸術・スポーツビジネス専攻】	実施する 【教員養成課程、国際地域学、芸術・スポーツ文化学科芸術・スポーツビジネス専攻】 実施する 【教員養成課程、国際地域学、芸術・スポーツ文化学科芸術・スポーツビジネス専攻】		https://www.hokkaido-u.ac.jp/	「その他、以下の事項等については、「令和7年度大学入学共通選抜に係る大学入学共通選抜要項及び学生募集要項」及び「令和7年度大学入学共通選抜に係る大学入学共通選抜要項及び学生募集要項」の公表内容等に基づき、検討し、決定する。大学入学共通テストの教科・科目等の配点・大学入学共通選抜及び一般選抜専修要項に記載の「大学入学共通テストの扱い」
北海道	北海道	北海道	室蘭工業大学	2023年4月9日	予備	「令和7年度大学入学共通選抜に係る大学入学共通テストの受験を要する教科・科目について」配点は未公表。	実施する （芸術・スポーツ文化学科 音楽文化専攻、美術文化専攻、スポーツ文化専攻） 選択可 （芸術・スポーツ文化学科 音楽文化専攻、美術文化専攻、スポーツ文化専攻）	実施する （芸術・スポーツ文化学科 音楽文化専攻、美術文化専攻、スポーツ文化専攻） 選択可 （芸術・スポーツ文化学科 音楽文化専攻、美術文化専攻、スポーツ文化専攻）	実施しない		https://muran-i.ac.jp/uploads/sites/6/2022/03/2025gakub	「予告内容については、今後変更される場合がありますので、令和7年度入学者選抜要項及び学生募集要項で必ず確認してください。」
北海道	北海道	北海道	小樽商科大学	2023年11月6日	予備	「令和7（2025）年度小樽商科大学入学選抜」	実施する （選抜コース） 選択可	実施する （選抜コース） 選択可	実施しない			「配点等の詳細については、大学入学共通テストの配点が決定後、令和7年度入学者選抜要項及び学生募集要項で必ず確認してください。また、本内容については、今後変更される場合がありますので、本学ウェブサイトや7年度入学者選抜要項及び学生募集要項必ずご確認ください。また、本内容、令和7年度入学者選抜に係るものであり、令和8年度以降も変更される可能性があります。おいて、フロンティア入試については、後日改めて公表いたします。」

公表日
予告・確定
文書のタイトル、概要、配点
共通テストを課す選抜方式での扱い
一般選抜での扱い



3 調査の目的

教科「情報」の2025共通テストに向けた「指導環境の提供」

普段忙しい学校の先生に代わって、徹底的に「情報」をリサーチし、より役立つ指導環境を提供する。

教科「情報」に関するまとめ（ラーズWEBサイト）

- 「情報 I」のご指導に関するアンケート結果報告
- 2025共通テスト「実施大綱の予告・補遺」まとめ
- 2025年共通テスト「試作問題」の分析と指導提案

教科「情報」の大学入試情報まとめ（毎月約4000アクセス）

<https://www.learn-s.co.jp/edu/pg/info22090/>

「進研WINSTEP 情報 I」
（2023年11月発刊）

「2025共通テスト対策【実力完成】直前演習 情報 I」
（2024年1学期発刊）



情報

教材



4 調査の手順

- ① 「確認用ファイル」で進捗管理
 - ② 各大学のWEBサイトを確認
 - ③ 公表・更新情報を転記、刷り合わせ
 - ④ 「公表ファイル」を更新
 - ⑤ WEBで発信
- メルマガとFacebookで更新のお知らせ
- (①～⑤を社員2名で行い、2週間で全国公立大学を1巡するペースです)

① 「確認用ファイル」

エリア	都道府県名	大学	公表日 更新日	最新情報リンク先サイト	チェック日	2022/02/16 https://www.obihiro.ac.jp/wp-content/uploads/2022/02/44b2955ef7b9dca7025e2c070e950f1e	2月27日	3月13日	3月27日
北海道	北海道	帯広畜産大学	2023年2月16日	https://www.obihiro.ac.jp/	1月30日				
北海道	北海道	旭川医科大学	2022年6月14日	https://www.asahikawa-med.ac.jp/bureau/nyus/contents/pdf/notice_2022_01.pdf	1月30日				
北海道	北海道	北見工業大学	2022年9月9日	https://www.kitamit-it.ac.jp/	1月30日				

↓色付きのセルが、現在公表されている大学です。

各大学のリンクをコピー

② 各大学のWEBサイトを確認

金沢大学 受験生特設サイト

拡張機能
アクセス不要
これらの拡張機能は、このサイトの情報の表示、変更を必要としません。
 Pasty
拡張機能を管理

5 調査の手順

- ① 「確認用ファイル」で進捗管理
 - ② 各大学のWEBサイトを確認
 - ③ 公表・更新情報を転記、刷り合わせ
 - ④ 「公表ファイル」を更新、WEBで発信
 - ⑤ メルマガとFacebookで更新のお知らせ
- (毎週更新)

④ 「公表ファイル」

SP1549	「情報 I」の2025年度大学入試関連情報 (国立大学一覧)										株式会社 ラーンズ		
	エリア	都道府県名	大学	公表日 更新日	確度	概要、配点等	共通テスト「情報 I」の実施			個別試験での「情報 I」の実施		最新情報リンク先サイト	備考 (公表の履歴・今後の予定等)
							一般選抜	総合型選抜 (共通テストあり)	学校推薦型選抜 (共通テストあり)				
北海道	北海道	帯広畜産大学	2023年2月16日	予告	「令和7年度帯広畜産大学学部入学選抜 (令和6年度実施) について (予告) (第2報)」- 畜産科学課程-一般選抜 (後期日程) の募集人員を25人から20人へ変更し、総合型選抜の募集について新規公表。募集人員は25人、配点は150点。また、総合型選抜の基本方針が示される。	実施する	実施する (畜産科学課程)	実施する	実施しない	https://www.obihiro.ac.jp/news/45653	総合型選抜の基本方針-「総合型選抜では、畜産科学を修得するために必要な理系教科の基礎知識及びその応用力の他、畜産分野においてニーズが高まっているデータサイエンスについて、情報、統計、計算及びビジュアル表現(能力)、そして学習意欲を評価の対象とする。」 2022/5/26 「令和7年度帯広畜産大学入学選抜 (令和6年度実施) について (予告)」-一般選抜、学校推薦型選抜で共通テスト情報 I が課されます。配点は未公表。 「より詳細な内容については、2024年 (令和6年) に公表する入学選抜要項及び学生募集要項で確認ください。」		
北海道	北海道	旭川医科大学	2022年6月14日	予告	「令和7年度旭川医科大学入学選抜における出題科目等について (予告)」-配点は未公表。	実施する	実施する	実施する	実施しない	https://www.asahikawa-med.ac.jp/bureau/nyus/contents/pdf/notice_2022_01.pdf	「配点等の詳細については、令和7年度旭川医科大学入学選抜要項の発表までお知らせする予定です。」		
北海道	北海道	北見工業大学	2022年9月9日	予告	「令和7 (2025) 年度工業部一般選抜における実施教科・科目等について (予告) (第1報)」-「地理歴史・公民」及び「情報」も受験している者については、「地理歴史・公民」及び「情報」のうち得点の高い科目の成績を用いる。(ただし、「地理歴史・公民」において複数科目を受験している者については、第1報資料目録「情報」の「得点の高い科目の成績を用いる」) 配点は未公表。	選択可			実施しない	https://www.kitamit-it.ac.jp/	「内容については今後修正がある場合がありますので、必ず令和6 (2024) 年度に公表する入学選抜要項及び学生募集要項並びに本学ホームページ (https://www.kitamit-it.ac.jp/) にて確認してください。」		
東北	青森県	弘前大学	2022年12月26日	予告	「令和7 (2025) 年度入学選抜方法について (第2報)」-配点は未公表。	実施する	実施する (医学部医学科及び医学部保健学科の総合型選抜 II)		実施しない	https://nyushi.hirosaki-u.ac.jp/news/faculty/9467/	2022/12/26 「令和4年7月29日付」令和7 (2025) 年度入学選抜からの新教科「情報」について「予備知識が追加されたが、追加が事項については、決定次第速やかに公表いたします。」 2022/7/29 「令和7 (2025) 年度入学選抜からの新教科「情報」について」 https://nyushi.hirosaki-u.ac.jp/news/faculty/9467/ 「上記内容は現時点でのものであり、今後、変更となる場合があります。なお、令和7 (2025) 年度入学選抜方法の詳細や変更等については、令和4年12月までお知らせする予定です。」		

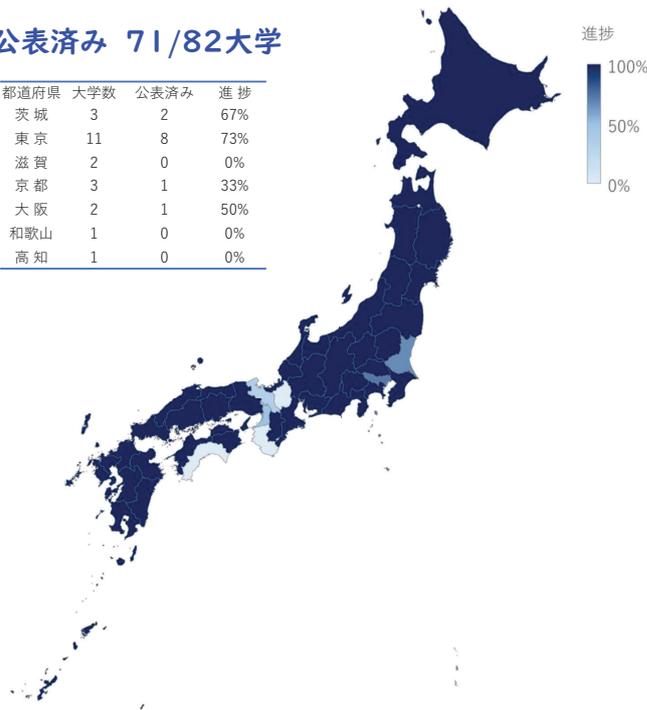
6 公表状況（2023年3月7日時点）

国立大学

公立大学

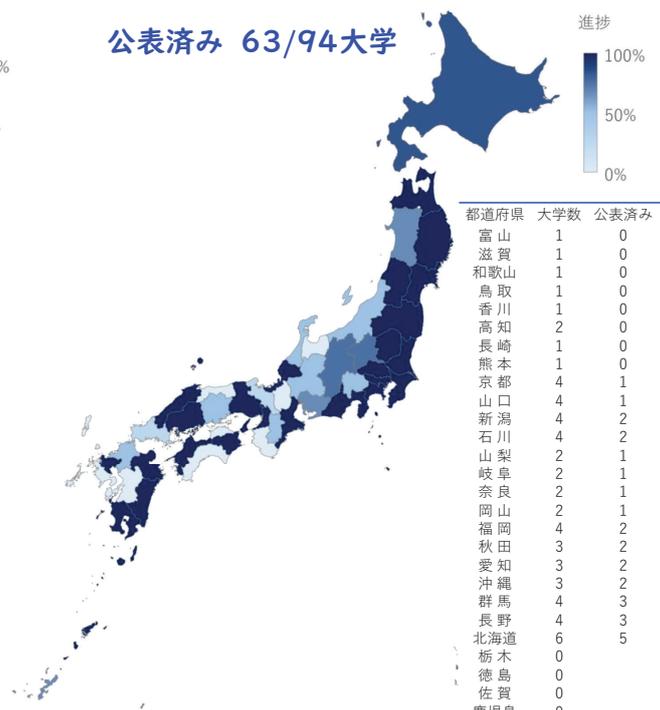
公表済み 71/82大学

都道府県	大学数	公表済み	進捗
茨城	3	2	67%
東京	11	8	73%
滋賀	2	0	0%
京都	3	1	33%
大阪	2	1	50%
和歌山	1	0	0%
高知	1	0	0%



公表済み 63/94大学

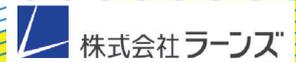
都道府県	大学数	公表済み	進捗
富山	1	0	0%
滋賀	1	0	0%
和歌山	1	0	0%
鳥取	1	0	0%
香川	1	0	0%
高知	2	0	0%
長崎	1	0	0%
熊本	1	0	0%
京都	4	1	25%
山口	4	1	25%
新潟	4	2	50%
石川	4	2	50%
山梨	2	1	50%
岐阜	2	1	50%
奈良	2	1	50%
岡山	2	1	50%
福岡	4	2	50%
秋田	3	2	67%
愛知	3	2	67%
沖縄	3	2	67%
群馬	4	3	75%
長野	4	3	75%
北海道	6	5	83%
栃木	0	0	0%
徳島	0	0	0%
佐賀	0	0	0%
鹿児島	0	0	0%



	公表	未公表
国立大学	86.6%	13.4%
公立大学	67.0%	33.0%

Powered By Bing
© GeoNames, HERE, MSFT

Powered By Bing
© GeoNames, HERE, MSFT



7 公表状況（2023年3月7日時点）

表 国立大学、公立大学それぞれの「予告」「決定」の割合

	「予告」	「決定」
国立大学	90.1% (64大学)	9.9% (7大学)
公立大学	95.2% (60大学)	4.8% (3大学)

※ 公表済みの大学数に対する割合(%)です。

「予告」…文書のタイトルや内容に「予告」とあるもの。

「決定」…文書のタイトルや内容に「予告」等がなく、次のように言い切っているもの。

「入試区分、実施方法等が決定しましたので公表いたします。」

「以下のとおり変更する。」

「以下のとおりとします。」



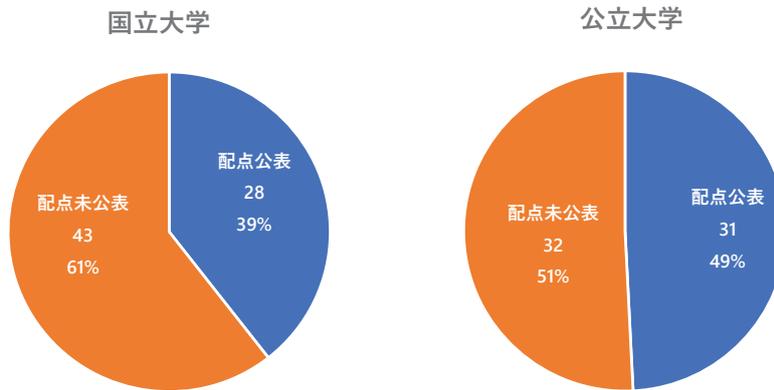
8 「配点」の状況（2023年3月7日時点）

表 一般選抜における共通テスト情報Ⅰの「配点」

	配点公表	100点	100点未満	100点超
国立大学	28/71	10	16	3
公立大学	31/63	20	9	7

※ 「試作問題」では60分、100点満点の解答時間・配点でした。

※ 同じ大学でも募集単位によって配点異なる場合がありますため、大学数の重複があります。



9 「配点」の状況

配点ウェイトが高い大学

愛知県立大学(情報科学部)

「一般選抜前期日程と学校推薦型選抜(全国枠)では150点満点、一般選抜後期日程と学校推薦型選抜(愛知県内枠)では200点満点に換算して利用する。」

学 科	試 験 区 分	教 科						面接	配点 合計	
		国語	地歴	公民	数学	理科	外国語			情報
情 報 科 学 科	大学入学共通テスト	200	100		200	150	◎200	○150	—	1,000
	個別学力検査等	—	—		500	—	—	—	—	500
	計	200	100		700	150	200	150	—	1,500

大学入学共通テストで受験を要する教科・科目		配点
国語	「国語」	100
地理歴史※1	「地理総合、地理探究」「歴史総合、日本史探究」「歴史総合、世界史探究」「地理総合、歴史総合、公共」	1科目選択 100
公民※1	「公共、倫理」「公共、政治・経済」「地理総合、歴史総合、公共」(再掲)	
数学	「数学Ⅰ、数学A」 「数学Ⅱ、数学B、数学C」	500
理科※2	「物理」「化学」「生物」「地学」	1科目選択 150
外国語	「英語」	◎200
情報	「情報Ⅰ」	○150
(6教科7科目) 合計		129
		1,200

10 「配点」の状況

共通テストで情報Ⅰを課すが点数化しない／配点0点

北海道大学

「大学入学共通テストの成績、個別学力検査等の成績及び調査書等を総合して合格者を決定します。ただし、**大学入学共通テストの情報Ⅰの成績は配点しません**。なお、本学が指定した個別学力検査等の教科・科目等の全てを受験していなければ、合格者としません。また、成績同点者の順位決定にあたっては、個別学力検査等の成績を重視します。個別学力検査等の成績も同点の場合は、大学入学共通テストの情報Ⅰの成績を活用します。」

徳島大学(2026年度入試から点数化)

「令和8年度入学者選抜(令和7年度実施)までは**「総合判定の参考」とし点数化を行いません**が、令和9年度入学者選抜(令和8年度実施)より点数化を行う予定です。」

香川大学(2026年度入試から点数化)

「共通テスト「情報Ⅰ」を受験していることを出願要件とし、出願書類(調査書)とともに、**総合的に判定するための資料として利用します**。(経済学部(学校推薦型選抜ⅡB)を除く)」

東京工業大学(総合型選抜の第2段階選抜は10点または0点)



11 「配点」の状況

公平・公正な入学者選抜への配慮

九州工業大学(2022/3/25にいち早く「情報」の実施を公表。2022/12/27に次の公表)

「大学入学共通テストの新教科「情報」については、他教科の必修単位数等とのバランス、ならびに、現状では学校現場における「情報科」教員の配置等の実施体制に地域間や学校間で大きな差があることをふまえ、令和7年度入学者選抜においては、**配点比率を低く設定すること**としました。本学では、社会における情報教育の重要性を強く認識しておりますので、引き続き、学校現場における教育環境の整備や学びの状況、大学入学共通テストの内容等を注視しながら、教科「情報」の配点変更ならびに個別学力検査・適性検査への導入について検討していきます。」

筑波大学

「導入1年目は特に高等学校卒業生への影響の大きさを考慮して、受験者全員に対する配慮として、**例えば配点を低くするといった負担軽減策**を講じます。配点や配慮の方法については、決まり次第公表します。」



12 個別試験で教科「情報」を実施する大学

電気通信大学

「情報理工学域 I 類(情報系)の総合型選抜(昼)および学校推薦型選抜においてCBTを活用した「情報 I」を含む入学者選抜を実施します。CBTによる選抜には本学の教員が参画して独立行政法人大学入試センターで開発されたシステム I) を活用する予定です。このシステムでは、①実際のプログラミング環境でプログラムを編集・実行しながら解答を求める問題、②データ解析ツールを用いて実際のデータを分析しながら解答を求める問題など、従来の紙による試験や口頭試験だけでは測定ができなかった力を評価する問題を出題することができます。」

帯広畜産大学

総合型選抜の基本方針…「総合型選抜では、農畜産科学を修得するために必要な理系教科の基礎知識及びその応用力の他、農畜産分野においてもニーズが高まっているデータサイエンスについて、情報、統計、計算及びビックデータを読み解く能力、そして学習意欲を評価の対象とする。」

13 個別試験で教科「情報」を実施する大学

宇都宮大学

「宇都宮大学データサイエンス経営学部(仮称・設置構想中)において実施する令和7年度(2025年度)入学者選抜予定について」

一般選抜(前期日程、後期日程)および総合型選抜A(一般)を行い、総合型選抜の論述試験では、経営に関する数理的な分析とデータサイエンスに関する基礎的な素養及び論理的思考力を評価し、面接ではコミュニケーション能力と社会への貢献、経営に関する数理的な分析とデータサイエンスに対する関心や意欲を評価します。

14 個別試験で教科「情報」を実施する大学

広島市立大学(情報科学部後期日程(配点300点))

〈実施内容及び評価の視点〉

1次選考の配点は活動報告書50点、志願理由書50点です。

情報工学・情報科学に関する学修や活動※についてアドミッション・ポリシーに基づき、「関心・意欲」「主体性・協働性」を重点的に評価し、審査・選考します。

2次選考対象者は募集人員の2倍程度の範囲内とします。

※《例》

- ・本学で実施する「高校生による情報科学自由研究」への参加、取組
- ・本学で実施するオープンキャンパスの模擬授業への参加
- ・本学または他大学が実施する講座や高大接続事業への参加、取組
- ・高等学校での「総合的な探究の時間」における取組
- ・情報科学に関する活動(数学に関する研究、アルゴリズム・プログラムの作成など)

群馬県立女子大学(個別試験で情報Ⅰは課さないが、情報に関連することを出題)

文学部文化情報学科の一般選抜…「文化情報学科の「総合問題」は、社会や情報に関する文献や資料等について理解力や論理的思考力を問います。」

文学部文化情報学科の総合型選抜…「文化情報学科の教員による社会や情報をテーマとした講義を受講した後、講義内容についての理解を問う試験を行います。」

このような要件を満たす生徒の育成を、学校現場の限られたリソースで、どのように行っていくかが課題です。



15 個別試験における教科「情報」に関する言及

新潟大学

「注）本学が実施する試験において、「課さない」としている場合であっても、試験(小論文等)によっては、「情報」に関連する内容が盛り込まれる場合があります。」



16 学部、学科等の新設

秋田大学

「令和6年度（2024年度）の開設に向けて、ICT・データサイエンス系の新学部設置および教育文化学部と理工学部の学部改組を計画しています。」

宇都宮大学

「宇都宮大学データサイエンス経営学部（仮称・設置構想中）において実施する令和7年度（2025年度）入学者選抜予定について」

注）大学のWEBサイトから見つけれられたもので、網羅しているものではありません。

17 大学から受験生へのメッセージ（情報を学ぶ意義）

琉球大学 2022/11/30

琉球大学を目指すみなさんへ

令和7年度琉球大学入学者選抜における「情報Ⅰ」の活用について

重要なお知らせ

琉球大学では、令和6年度（2024年度）に実施する令和7年度（2025年度）入学者選抜において、大学入学共通テストの新教科「情報Ⅰ」を課すことになりました。対象は大学入学共通テストを課すすべての選抜区分です（別紙1参照）。大学入学共通テストでは「情報Ⅰ」を受験しておいてください。

なぜ新教科「情報Ⅰ」が必要なのか

みなさんの中には「自分が受験のときに教科が増えるなんて・・・」と思っている人もいるでしょう。大学入学共通テストに新しく教科が追加されるのは、前身の大学入試センター試験、共通一次試験を通して初めてのことです。これには次のような理由があります。

Society5.0 超スマート社会がやってくる

Society5.0|とという言葉聞いたことがあるでしょうか？ これは「超スマート社会」とも呼ばれ、AI、ビッグデータ、IoT等の先端技術が高度化して、社会に浸透し、社会の在り方そのものが大きく変わった新しい世界のことです。この社会はもう目の前に迫っており、みなさんは将来的にこの新しい世界で生きることになります。このような世界では、ただ情報を集めた知識だけ持っていて、その情報に振り回されるだけでは、「適応できていない」とは言えないでしょう。その膨大なデータを有効活用し、数理的に扱い、その他の知識と結び付けて新たな知識や価値を生み出すことで、「適応する」ことができるようになると思われれます。

18 大学から受験生へのメッセージ（情報を学ぶ意義）

データを扱う能力・スキルはみんなに必要

新しい社会に適応して生きるために必要な能力・スキルですから、分野を問わず、すべての大学生が適切にデータを扱えるようになる必要があります。そのために今、教育が変わっています。その一部が、高校で「情報Ⅰ」が必修化されたことです。そして、それがどの程度身につけているのかを評価するために、大学入学共通テストの「情報Ⅰ」が課されることになりました。大学入学共通テストで「情報Ⅰ」を課すことで、大学入学後の教育のための準備が、どのくらいできているのかを評価するのです。

琉球大学の取り組み

みなさんが学んできた、データを扱う能力・スキルは入試で評価されて終わりではありません。むしろ社会に出てからより必要になる能力・スキルなので、大学ではさらに磨きをかけます。琉球大学では、すべての学部の学生が、「データを分析し、その中にある情報や知識を読み取り、それを基に論理的に思考して解釈し、さらに議論をすることができるようになるための科目」を履修します。そして、卒業までに「数理的な思考力」と「データを分析する能力・スキル、活用する能力」を身に付けることを目標としています。

琉球大学は、文部科学省の事業である「島嶼地域社会の自律的・持続的発展に向けた地域共創型数理・データサイエンス・AI 教育モデル普及展開事業」（社会科学分野、ダイバーシティ）に認定されており、データを扱う能力・スキルについての高度な教育を、広く深く受けることができます。

みなさんが高校で身に付けた能力・スキルは、大学入試において評価され、それが大学でさらに磨かれ、Society5.0の中でよりよく生きるために、みなさん自身が活用していくこととなります。

「情報Ⅰ」が大学の入学者選抜に取り入れられた理由は理解いただけでしょうか？ 琉球大学だけでなく、多くの大学が「情報Ⅰ」を入試で課しています。それだけ今後のために必要な能力であると、国も大学も考えているのです。このことを理解し、日々の学びを大切にしてください。なお、「情報Ⅰ」をはじめ、大学入学共通テストでの配点等は今年度内に、他の教科の活用方法と合わせて公表する予定です。

19 大学から受験生へのメッセージ（情報を学ぶ意義）

電気通信大学 2023/1/17

2025（令和7）年度入学者選抜における科目「情報Ⅰ」の取り扱いについて

要旨

高等学校学習指導要領（平成30年告示）に対応した入学者選抜が、2025（令和7）年度から実施されるにあたり、今後の情報化社会を先導する人材を育成するために、本学では令和7年度入学者選抜（令和7年2月実施予定）から、情報理工学域一般選抜前期日程試験において科目「情報Ⅰ」を課して実施することとしました。

現在、我が国のデジタル人材不足は社会的問題とされ、高度な人材育成は急務となっています。

「情報」を研究領域とする本学においても、様々な教育プログラムを実施してまいりましたが、今後の情報化社会を先導する人材を育成するためには、より「情報」に対する高い素養を持った学生を受入れ、高度な知識および技術を学んでもらうことが必須と考えています。

そのため、高等学校学習指導要領において必修とされた「情報Ⅰ」を学んだ高等学校卒業者が大学に進学する令和7年度入学者選抜から、科目「情報Ⅰ」を試験科目として選択可能とした入学試験を実施することとしました。

「情報」を試験科目として課すことで、多様な学生が入学し本学の教育プログラムの中で切磋琢磨しながら日本社会を牽引するデジタル人材が育成されるものと期待しています。

20 大学から受験生へのメッセージ（情報を学ぶ意義）

広島市立大学 2023/3/6

2025年度(令和7年度)入学者選抜における「情報I」の導入について

現在、情報通信技術は飛躍的な進展を遂げつつあり、それに伴って私たちの生活は大きく変化してきています。私たちは今後あらゆる活動において情報を適切に選択・活用していくことが不可欠といえます。

それを踏まえて改訂された学習指導要領(平成30年3月告示)では、問題の発見・解決に向けて、いろいろな物事や現象を、情報とその結び付きの視点から捉え、情報技術を適切かつ効果的に活用する力を全ての生徒に育む「情報I」が共通必修科目となりました。

本学では、その有用性・重要性を考慮し、2025年度(令和7年度)入学者選抜から、大学入学共通テストの利用科目として「情報I」を導入することとします。国際学部と芸術学部では選択科目として利用します。特に、開学時から情報科学の分野で教育研究を進める情報科学部では大学入学共通テストの「情報I」を必須とし、一般選抜後期日程の個別学力検査で「情報I」を課すこととします。

情報科学部からのメッセージ

情報科学部では、2025年度(令和7年度)の入学者選抜から、一般選抜前期日程及び後期日程の大学入学共通テストの利用科目として「情報I」を必須にするとともに、一般選抜後期日程の個別学力検査に「情報I」を導入します。

広島市立大学は1994年に情報科学部を設立しました。恒久の平和を見つめ未来を照らすようにと設立した学部です。遠くを照らす灯台が荒波にさらわれぬしっかりした基礎を持つように、開学時から約30年かけてしっかりと皆さんに情報科学を教育する基礎を築いてきました。私たち情報科学部の教員は、皆さんを情報科学で未来を切り拓ける人に育てたいと思っています。

皆さんの高等学校での学びは大きな未来につながります。その学びの成果を入試で表現していただくと共に、なによりも「『情報』が好き」な皆さんにぜひ門を叩いて欲しいと思い、入試に「情報」の科目を設定することにしました。皆さんも考えているように、社会では「情報」人材が求められています。AIも電卓のように道具として活かす時代になります。皆さんの未来を照らす灯台のように、しっかりと基礎を築いてきた広島市立大学情報科学部で、皆さんの好きな情報科学を学びましょう。

21 大学から受験生へのメッセージ（情報を学ぶ意義）

アドミッションポリシーに「情報」

岡山県立大

「看護学科、栄養学科、現代福祉学科、子ども学科、情報工学部の入学者受入れの方針…情報Iにおける内容の理解と応用力」

「デザイン学部の入学者受入れの方針…問題の発見・解決に向けて情報と情報技術を適切かつ効果的に活用するための能力(知識及び技能)」

22 2025入学者選抜までのスケジュール

大学入試には、いわゆる「2年前予告」というルールがあり、大きな変更がある場合は、原則2年程度前に予告を公表することになっています。（「大学入学者選抜実施要項」、文科省）

時期	おもなイベント、公表予定の情報	発信元、実施主体
2022年度中	「2年前予告」…国公立大の個別試験の教科・科目の設定等	各国公立大学
2022年11月9日	共通テスト各教科・科目の「問題作成の方向性」（地歴、公民、数学、情報（経過措置問題を含む）の試作問題が示された）	大学入試センター
2023年6月初旬までに	「2025新課程共通テスト実施大綱」（令和7年度大学入学者選抜に係る大学入学共通テスト実施大綱） 現行課程における共通テストの実施方針は廃止され、予告した出題教科・科目等を含む試験の実施方針、実施期日等が示されます。	文部科学省
2023年6月までに	「令和7年度大学入学者選抜に係る大学入学共通テスト出題教科・科目の出題方法及び問題作成方針」（経過措置を含めた出題方法、科目選択の方法など）*	大学入試センター
2024年6月までに	「2025新課程 大学入試実施要項」（令和7年度大学入学者選抜に係る実施要項）	文部科学省
（上記に続いて）	「2025新課程共通テスト実施要項」（令和7年度大学入学者選抜に係る大学入学共通テスト実施要項）	大学入試センター
2024年9月	2025新課程共通テスト「受験案内」の配布	大学入試センター
2024年9月～	総合型選抜の実施	各国公立大学
2024年11月～	学校推薦型選抜の実施	各国公立大学
2025年1月	2025新課程共通テストの実施	大学入試センター
2025年2月～	一般選抜の実施	各国公立大学

※ 「令和7年度大学入学者選抜に係る大学入学共通テストに関する検討状況」（2021.10.1、大学入試センターより）

23 「2年前予告」で確認しておきたい事項

- 1) 文書のタイトル
- 2) 一般選抜、学校推薦型選抜、総合型選抜それぞれのおもな変更点
- 3) 「予告」か「決定」かの確認
- 4) 共通テストの「情報」の取り扱い（実施する／選択可／実施しない）
- 5) 共通テストの実施教科・科目
- 6) 共通テストの配点
- 7) 個別試験の教科・科目
- 8) 経過措置（既卒生への対応）
- 9) 入試日程、入試方式、募集人員の変更
- 10) 変更、予告等

例) 「公表内容は現時点のものであり、今後変更する可能性があります。」

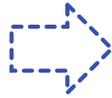
例) 「配点につきましては、2023年3月末までに、確定次第公表いたします。」

例) 「令和6年度に発行される令和7年度入学者選抜要項及び学生募集要項で必ず確認してください。」

私立大学を含めた対応状況を調べるには

検索のキーワード

- 「令和7年度入学者選抜」
- 「令和7(2025)年度入学者選抜」
- 「令和6(2024)年度実施入学者選抜」
- 「平成30年告示高等学校学習指導要領に対応した」



Google 令和7(2025)年度入学者選抜

約 203,000 件 (0.40 秒)

もしかして: **令和7(2025)年度入学者選抜に係る予告**

https://www.ris.ac.jp/news/news20230228_03 :
令和7(2025)年度入学者選抜に関わる予告について - 立正大学
 5 日前 — 令和4(2022)年度高等学校入学者から新学習指導要領が適用されることに伴い、本学における令和7(2025)年度入学者選抜に関する予告を下記の通り ...

https://www.ris.ac.jp/pdf/r7_exam-system PDF :
【立正大学】令和7(2025)年度入学者選抜に関わる予告 ...
 5 日前 — 本学の令和7(2025)年度入学者選抜について、各試験制度の変更点等を下記の通りお知らせいたします。なお、今回の公表内容は現時点での予定で ...

高校現場の意識

Q19 「情報Ⅰ」のご指導に関する次の①～⑱の事項について、当てはまるお考えや気持ちを、1～3のうちからお選びください。[1課題・不安はない/2どちらともいえない/3課題・不安がある]

教科指導～受験指導



まだ高1指導の段階ですが、受験指導に関することに課題や不安が集中しています。

出典:弊社「情報Ⅰ」のご指導に関するアンケートより(2022年5月18日～7月1日 WEBにて実施 n=347)

教科「情報」の特性と現場の課題

- 共通入試として初の出題教科
- 国公立大学の実施状況・配点等
- 必履修科目（他の必履修科目との単位数等のバランス）
- 教員配置、地域間・学校間の指導体制の差
- 高校現場の温度感（情報の先生の立場、管理職の意識）
- 情報教育そのものが過渡期であること（教科書の内容）



2025共通テストの他教科のポイント

- 共通テスト国語は「国語」1科目
- 「論理国語」と「文学国語」（とくに理系の文学への対応）
- 地理歴史・公民は、「地理総合、地理探究」「歴史総合、日本史探究」「歴史総合、世界史探究」「公共、倫理」「公共、政治・経済」「地理総合、歴史総合、公共」の6科目から最大2科目選択
- 数学は「数学Ⅰ、数学A」「数学Ⅰ」「数学Ⅱ、数学B、数学C」の3科目へ
- 理科は、「物理」「化学」「生物」「地学」「物理基礎、化学基礎、生物基礎、地学基礎」の5科目へ
- 共通必履修科目「歴史総合」「地理総合」「公共」「情報Ⅰ」や、「数学C」（特にベクトル）の共通テストでの扱い
- 難関国立大が採用する入試科目（がはっきりしないとカリキュラムが組めない）
- 地歴公民の負担増大（日本史と世界史の両方を学習、教科書を終わらせられるか）
- 理系の生徒にとっての地理総合、歴史総合の位置づけ



トピックス 「年内入試」

「年内入試」の拡大

「教育改革」の流れ

- ・学力だけでは測れない適性を見極める。
- ・多面的・総合的な評価
- ・多様な生徒、学生を集め、各校の特色ある教育を促進させる。

京都大学法学部

「試験実施方式として後期日程を廃止し、学校推薦型選抜を実施します。」

徳島大学 生物資源産業学部生物資源産業学科

募集人員の変更（前期48→44名、学校推薦型選抜8→12名）

帯広畜産大 畜産科学課程

募集人員の変更（後期25→20名、総合型選抜の新設→5名）

東京工業大学

総合型・学校推薦型選抜で143人の「女子枠」を導入

横浜国立大学

学校推薦型選抜の理工学部化学・生命系学科化学教育プログラム・化学応用教育プログラム・バイオ教育プログラム、数物・電子情報系学科電子情報システム教育プログラムは新設（定員15名）。

学生の「早期確保」

- ・「少子化」が根底にあり、膨らんだ学校数、大学数、公立私立間の学費差の緩和などによる各校の生き残りが背景。
- ・コロナ禍での2020年度入試以降、「早く進路を決めたい」という生徒・保護者の意識により年内入試が加速した。



株式会社 ラーンズ

トピックス 「東京23区内の大学入学定員規制の緩和検討」

（日経ニュースより）

政府は23区内の大学定員規制を2024年度にも緩和する検討に入る。

原則禁止している23区内の学部学科新設や既存学部の定員増を、情報系に限り、期限付きでの規制緩和。

内閣官房の有識者会議が近く定員規制の緩和案を示し、

政府はこれに基づいて全国知事会などと相談して判断し、政省令で決定。

既存の学部学科の定員を増やす案もある。

ITやデジタル産業が集積する23区内で情報・デジタル分野で産学連携し、成長産業への人材育成や、国際競争力を向上させることがねらい。

経営余力がある大規模私立大を想定。

増員は期間限定で、一定期間を経て元に戻す。

内閣官房によると、

2022年時点で23区内に本部を置く大学数は101（全国の1割）

入学定員は2021年時点で約122,000人（全国の2割）



株式会社 ラーンズ

2025年度に入学する学生の統計スキルについて

大橋 真也・順天堂大学 数理・データ科学教育研究センター
s.ohashi.au@juntendo.ac.jp

1 はじめに

新学習指導要領で学んだ高等学校1年生がこの3月で修了し、進級する。高等学校の先生方は、新しい学習内容を指導し、さらに観点別評価の導入もあり、混迷の1年間であったかもしれない。新課程の年次進行が完了する2年後、大学に入学する学生の資質・能力が変化するとされている。

本稿では、情報科の高等学校学習指導要領等の作成にも関わり、2021年度まで高等学校において、高等学校現場で数学科と情報科の指導をしてきた自分の経験をもとにし、新課程の高校生の資質・能力がどのように変化し、2025年の大学入学生がどのように変化していくのかについては、考察していく。

2 高等学校新学習指導要領における統計分野の扱い

2.1 数学 I

前課程から「(4) データの分析」として、数学 I に導入されたが、前課程にあった箱ひげ図(外れ値なし)や四分位範囲などは中学校2年生で学習する内容になり、新たに数学 I には「仮説検定の考え方」が導入された。また質的データ、量的データの扱いが以前よりも明確になった。前課程からの学習内容の主な変化は次のようなものである。

- 箱ひげ図については、外れ値を表記したものを扱う。外れ値についても学ぶことになった。
- 質的データに関しては、分割表を用いてその関係を理解する。
- 量的データに関しては、散布図を用いるが、大量のデータであると点の重なりが出ることから、箱ひげ図も併用することもある。
- 仮説検定の考え方では、コイントスなどのシミュレーションをもとにして、確率的に考える。

2.2 数学 A

「(2) 場合の数と確率」として、前課程と同様の扱いで導入されている。新学習指導要領解説には、「主観確率」「ベイズ推定」等の用語が登場するが、それらの用語が出てこない教科書も多い。また情報 I との関連で「モデル化とシミュレーション」との関連が解説されている。

2.3 数学 B

「(2) 統計的な推測」として、統計分野が導入されている。前課程との違いは、従来の「数学 B」にあった「ベクトル」が新課程では「情報 C」に移動し、「統計的な推測」をより多くの学校で展開しやすいような配列になっている。ただし、「情報 B」と「情報 C」の履修条件の順序性はなく、令和7年度からの共通テストの科目も「数学 IIBC」に変更され、必答問題に加えて「数列(数学 B)、統計的な推測(数学 B)、ベクトル(数学 C)、平面上の曲線と複素数平面(数学 C)の4項目に対応した出題とし、4項目のうち3項目の内容の問題を選択解答する。」としたために、各学校では新教育課程の編成途中で大きな変更を強いられた。またこの選択により、「統計的な推測」の選択者は増えるとみられ、高等学校での実施状況も変化が出てくると考える。

旧課程との内容の変更に関しては、従来の「確率変数」「確率分布」「二項分布」「正規分布」「母集団と標本」「推定」に加えて新たに、「仮説検定」が導入されたことである。「仮説検定」を高等学校で指導するのは、採択数は少なかったが1978年に告示された高等学校学習指導要領(1982～1993年)における科目「確率・統計」以来である。比較的多く履修されていた科目では、1970年告示の高等学校学習指導要領下の科目「数学 III」以来であろう。すなわち、現職の教員でこの内容を指導したことのある教員はごくわずかと考えられる。

現状の見本の内容を、公開することが難しいので、「新課程チャート式 数学 II+B」(数研出版)によれば、「数学 I」の新出項目である「仮説検定の考え方」を受け、「母比率の検定」「母平均の検定」などが扱われている。

2.4 情報Ⅰ

共通必修科目「情報Ⅰ」での統計分野の扱いは、「(4) 情報通信ネットワークとデータの活用」の単元である。(4)自体は大きく「ネットワーク技術」「データベース」「データの分析」の3つの内容に分けることができ、「情報Ⅰ」が2単位科目であることも考えると、扱いはかなり少なく感じる。しかし、大学入試センターの令和7年度の共通テストに「情報Ⅰ」が導入され、これが国公立大学を含めた多くの大学で採用されることになっており、またそのサンプル問題や試作問題が公開されるにしたがって、「データの分析」の扱いは大きくなってきている。

「情報Ⅰ」における統計分野の扱いは高等学校学習指導要領解説(情報編)によると、次のような内容である。

- 質的データ、量的データ、(名義、順序、間隔、比例尺度、連続、離散)の扱い方
- 外れ値、欠損値の扱い方
- 構造化データ、非構造化データの扱い方(可視化や分析方法、自然言語解釈など)
- キー・バリュー形式のデータや時系列データを扱う。
- 単回帰分析による予測
- 仮説検定の考え方

後述するが、現在出版されているほとんどの教科書に関してはこれらの内容を満たしていないものが多くみられ、実際に学校現場での指導内容に関しては問題が生じていると考える。

2.5 情報Ⅱ

共通選択科目「情報Ⅱ」に関しては、「情報Ⅰ」を修了したあとに履修する科目として設定されている。「情報Ⅱ」における統計分野の扱いは、「(3) 情報とデータサイエンス」の単元で扱われている。「情報Ⅱ」は選択科目であり、多くの高等学校出に採択はないと見込まれていたが、朝日新聞DIGITALの2022年9月3日の記事「情報科の教員採用、20年で1人だけの県も 公立高、地域格差の実態」の中では、東京都で84校と5割以上の学校での採用があり、神奈川県(63校)、埼玉県(48校)、千葉県(46校)と4割以上の学校が採択されているという。全国的な正確な調査の結果は公表されていないが、約2割弱の普通科高等学校での採択があると言われている。ただし採択が全くない県も存在し、地域間の格差が開いていくことが考えられる。今後、採択数は着実に増加傾向にあると考えることから、「情報Ⅱ」の内容を履修してくる学生も増加していくことと考えられる。

「情報Ⅱ」のデータサイエンスに関する内容は、文部科学省作成の高等学校情報科「情報Ⅱ」教員研修用教材によると次のような項目がある。

- データと関係データベース
- 大量のデータの収集と整理・整形
- 重回帰分析とモデルの決定
- 主成分分析による次元削減
- 分類による予測
- クラスタリングによる分類
- ニューラルネットワークとその仕組み
- テキストマイニングと画像認識

教科書に関しては、実教出版、東京書籍、日本文教出版の3社が出版しており、その内容に関しては様々である。教科書によっては、一般の実用書のように手法のみを解説して、どのような事例で扱い、どのように問題解決に役立てるのかなどの説明がほとんどないものもあり、今後の改善が望まれる。

2.6 カリキュラム・マネジメント

新学習指導要領においては、「カリキュラム・マネジメント」が謳われており、その中の1つの視点である「教師が連携し、複数の教科等の連携を図りながら授業をつくること」という「教科横断連携の視点」が統計分野の教育では重要になる。少なくとも「数学」「情報」の2教科の連携は必要であると考えられるし、「地理」「家庭」「保健・体育」などとのデータ等の連携等も考えられる。新学習指導要領解説では、「数学」と「情報」の連携を強く行うために解説の付録として、相互の解説内容の抜粋が含まれている。学校現場における円滑な教科間の連携が期待されている。

3 共通テストと統計教育

大学入試センターでは、2025年度の共通テストから「情報Ⅰ」を受験科目とすることを発表し、国立大学協会でも基本方針で「情報Ⅰ」を採用することを発表した。これに伴いほとんどの大学で「情報Ⅰ」を共通テストで受験科目とするようになってきている。

大学入試センターでは、これまで3回にわたって「情報Ⅰ」の試作イメージを公表している。2020年には最初の試作問題を公開しており、2021年3月にサンプル問題、2022年11月には本番での配点や構成なども意識した試作問題が公開されている。

2020年公開の試作問題には、データの分析分野の出題はなかった。2021年に公開されたサンプル問題では、ワールドカップのサッカーを題材にした問題が出題された。2022

年の試作問題では、第4問として配点25点でコンピュータやスマートフォンの利用時間と学習時間や睡眠時間との関係をグラフ等から傾向をつかみ、回帰直線により予測との乖離を残差分析のグラフで評価する問題が出題されている。また2022年には第4問参考問題としてエアコンとアイスクリームの売り上げ個数に関する時系列分析の問題も同時に公開されている。

特徴的な部分は、これらの3つの問題中2つの問題に提示されている散布図相関行列である。散布図行列や散布図相関行列に関しては、掲載されている教科書は全くない。教科書外でありながら、それを出题した根拠は、学習指導要領解説(情報編)の次のような記載がもととなっていると考える。

多くの項目のあるデータに対して、項目間の相関を見るためにデータを漏れのないように組み合わせて複数の散布図などを作成し、相関関係の見られる変数の組合せを見出し、その変数の組合せに関して回帰直線を考え、データの変化を予測する力を養うことが考えられる。

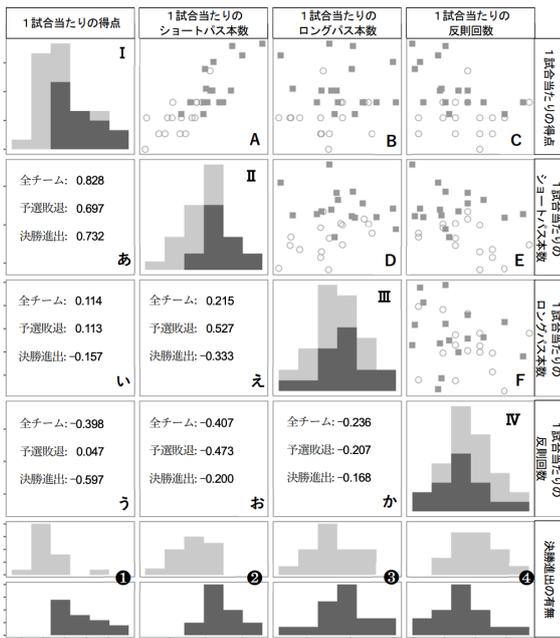


図1 各項目間の関係

また、2022年の参考問題にある時系列分析についても高等学校学習指導要領解説(情報編)や高等学校情報科「情報I」教員研修用教材でも扱われていたにもかかわらず、ほとんどの教科書で扱われていなかった内容である。この参考問題では、散布図行列以外にも自己相関関数のグラフを読み取らせたり、移動平均のグラフを選択させたりする問題も掲載されていた。

2022年の11月9日に公開された試作問題に関しては、大学生の受験モニター調査も行われたが、その結果に関しては残念ながら現時点では公表されていない。

4 統計教育と教科書

新しい高等学校学習指導要領解説(数学編 理数編)や(情報編)における新しい統計分野の学習が増えてきている。しかしながらここまででもいくつかの例を挙げたが、学習指導要領解説と教科書の間にはギャップがある。数学に関しては「仮説検定」に関する事が主なことであることから大きな乖離はみられないが、情報に関しては、次のような項目に関して「情報I」に関しての教科書の乖離がみられる。

- 単回帰分析を予測のために使うことを目的としているが、多くの教科書では直線を描いただけで終わっている。
- 時系列分析に関してほとんど教科書で扱っていない。
- 散布図相関行列について、教科書に掲載はなく、その使い方についても説明はない。
- 共通テストでは、問題解決の手順をストーリーとして展開しているが、教科書は手法のみを解説しているものが多く、実データを用いた例も少ない。
- データクリーニングやデータクレンジングについて記載のある教科書は少ない。
- 疑似相関や交絡因子についての記載のある教科書は少ない。

なぜこのようなことが起こるのだろうか。学習指導要領解説の内容を具体化したものが教科書ではないのだろうか。これには以下のような原因が考えられる。

- 教科書の編集コストの問題
教科書の内容を大幅に刷新することを控え、旧課程の教科書を内容の配列を変更して、最低限の新出項目に関して執筆者に依頼し、教科書を作成していることがある。
- 教科書執筆者の認識不足
執筆者が高等学校学習指導資料寮開設をきちんと読み解いていない。また、統計分野やデータサイエンスに詳しい執筆者が見つからなかった。

すべての教科書ではないが、一部の教科書に他の一般書籍や過去の教科書と記載が同一であるものが見つけられることから、上記のような懸念は捨てきれない。またここでは挙げなかったが、情報IIの重回帰分析や主成分分析の扱いに関してもいくつか問題点があると感じた。

大学入試センターの共通テストは、それらの教科書での記載内容の不足に関してのバラツキが受験生の有利や不利

にならないように、高等学校学習指導要領に沿った内容をきちんと解説しながら出題していると考えられる。

高等学校の教科担当(特に情報科)の先生方は、このような教科書の不備な部分を理解し、正しい内容を生徒に指導する必要が出てくると考える。また教科書に関しても、改訂のタイミングでこのような不備を少しでもなくしていく努力をすることを期待する。

5 2025年度の入学生

さて、今回のテーマは「2025年度の入学生の統計スキル」についてであった。数学などでの仮説検定を理解し、情報でも単回帰分析および残差分析などを学んできた学生が入ってくる。また首都圏で採択数の多い選択科目の情報IIを学んだ学生に関しては重回帰分析や主成分分析、統計的機械学習などは既知の内容として入学してくることと考えられる。

しかしながら、教科書の不備な部分もあり、学生の統計スキルの変化はある程度大きいですが、急激な変化ではなく、共通テストや教科書の改訂などともなっていて徐々に増えていくことであると考え。そのため、入学生の統計分野における知識や能力の格差は初年度には大きいですが、徐々にその格差も小さくなっていくことであろう。

6 おわりに

現在、大学において数理・データサイエンス・AI教育や統計学の授業を担当している自分にとっても2年後の変化に関しては、今後とも注視していかなければならないと考えている。AI等の進歩が急激に起こる中、常に変化のある内容を扱う情報科の科目を学んできた生徒がどんな学生として入学してくるのか楽しみでもある。ChatGPTなどで簡単にレポートの記述問題やプログラミングの課題がほとんど自分の知識がなくても解決してしまう時代になり、これからの統計教育も含め、何を教えていくことが大切なのかを、再検討する必要があると感じている。

参考文献

- [1] 文部科学省：高等学校学習指導要領，https://www.mext.go.jp/content/20230120-mxt_kyoiku02-100002604_03.pdf，2018.
- [2] 文部科学省：高等学校学習指導要領解説(数学編 数編)，https://www.mext.go.jp/content/20230217-mxt_kyoiku02-100002620_05.pdf，2018.
- [3] 文部科学省：高等学校学習指導要領解説(情報編)，

https://www.mext.go.jp/content/1407073_11_1_2.pdf，2018.

- [4] 文部科学省：高等学校情報科「情報I」教員研修用教材，https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1416756.htm，2019.
- [5] 文部科学省：高等学校情報科「情報II」教員研修用教材，https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/mext_00742.html，2020.
- [6] サンプル問題(『地理総合』、『歴史総合』、『公共』、『情報』)(令和3年3月24日)，https://www.dnc.ac.jp/kyotsu/shiken_jouhou/r7ikou/,2022.
- [7] 令和7年度試験の問題作成の方向性，試作問題等，https://www.dnc.ac.jp/kyotsu/shiken_jouhou/r7ikou/r7mondai.html,2022.

中等教育におけるデータサイエンス教育の実践と高大連携

神戸大学附属中等教育学校数学科教諭
神戸大学数理・データサイエンスセンター客員研究員
林 兵馬(h-hayashi@people.kobe-u.ac.jp)

1. はじめに

近年の AI・データサイエンスの普及とともに、令和3年度より実施されている中学校・高等学校の学習指導要領においても、AI・データサイエンスの取り扱いが益々増加している。

神戸大学数理・データサイエンスセンターの協力を得て、神戸大学附属中等教育学校では学校設定科目「データサイエンス I・II」の授業開発、総合的な学習・探究の時間における課題研究などを通じて統計教育の推進を進めてきた。

本発表では、下記2～4の内容について現状と今後の展望について報告する。

2. 学校設定科目「データサイエンス I・II」

令和2年から学習指導要領の改訂に先立ち、学校設定科目「データサイエンス I・II (以下 DS I・II)」を数学科として実施してきた。

・DS I：高校1年生、週1時間

内容：数学 I・B の統計内容を中心

・DS II：高校1年生、週1時間

内容：情報IIの DS 内容の一部

学習指導要領の改訂に先立ち、数学科として DS I・II を実施してきたため、「新しい教科書がまだ無い」「他校での実践がほとんどない」のような手探りの状態であったため、カリキュラムや授業内容の検討など様々な紆余曲折を経てきた。その紆余曲折の過程についても報告する。

3. 課題研究とデータサイエンス

DS I・II は、総合的な学習・探究の時間において K P (Kobe Port Intelligence Project) と称して実施している課題研究を支援するという側面を有している。

本校では課題研究を1人1テーマで実施をしている。令和3年の高校3年生の論文では「回帰または検

定」が登場する論文は10編/110編程度であったが、令和4年の高校3年生では30編/125編であり、生徒の中で統計分析が市民権を得ていると考えられる。また、生徒が課題研究で得た結果などを使い、様々なコンテストやコンクールで受賞してきた。

一方で、「仮説検定の誤用」「重回帰分析の誤用」などの各種統計手法を正しく使用できているか疑問がある論文は少なからず存在する。

また、本校の研究において「アンケート分析」を利用した個人研究が多いが、アンケートの質や分析手法の問題があると考えている。

4. 神戸大学との高大連携

神戸大学附属中等教育学校では、神戸大学数理・データサイエンスセンター（以下 CMDS）と連携を行ってきた。

神戸大学附属中等教育学校の「データサイエンス I・II」の授業開発指導、生徒への課題研究への助言、神戸大学1年生向け「データサイエンス基礎」の受講機会の提供などを通して、SSH 事業の運営にも協力いただいている。また、CMDS 主催で「中学生・高校生データサイエンスコンテスト」を2回開催した。

1. 記述統計の変化

統計教育が変化している。その要因として、データサイエンス (DS) の勃興、プログラムの発展 (特に統計グラフ描画)、オープンデータの充実、データ取得機会の拡大 (センサーなど)、初等中等教育における統計教育の充実等を挙げることができよう。筆者は、これらは、情報学や AI の影響はあるものの基本的に統計学の中での変化・発展と捉えている。DS の目的は、データを活用することによる課題解決や価値創造と云われる、宝物探しという形容もある。PPDAC サイクルも定着した。これらはデータ解析を生業とする伝統的な分野においてはさほど真新しいものではない。日本行動計量学会の Web には昔から存在する学会紹介ページがあり、たとえば、PPDAC という用語は使っていないものの、ほぼ同意の手续が掲載されている¹。

2. 代表値について何を理解させるのか

統計教育の中で、大きく影響を受け変わるべきなのが記述統計ではなかろうか。しかし、多くの統計学や DS 入門のテキストでは記述統計は補助的扱いで旧態依然という気がする。例えば、データの代表値の一つである中央値は、定義を記した後、外れ値に対して頑健、収入や貯蓄のような歪んだ分布の代表値として用いられるといった説明が多い。中央値や最頻値の一般的な目的はデータ分布・母集団分布のバルク²を捉えることであろう。平均値、中央値、最頻値の使い分けについては、単峰の歪んだ分布に対していわゆるピアソンの経験則のような大きさの順序、例えば、左に偏った (右に歪んだ) 単峰の分布の場合、

$$\text{最頻値} < \text{中央値} < \text{平均値}$$

となることはよく紹介される (e.g., 図 1)。しかし、これだけの情報で、「代表値の特徴を把握したうえで、今行おうとしている分析に適しているかどうかを判断しなければいけません」と結ばれても、学習者や実務家は迷うだけである。

代表値に関して少し高等なレベルになると、それぞれの推定量の理論的な性質や特徴づけ、

例えば、

$$\begin{aligned} \bar{x} &= \operatorname{argmin}_{a \in \mathbb{R}} \sum_{i=1}^n (x_i - a)^2 \\ \text{Med} &= \operatorname{argmin}_{b \in \mathbb{R}} \sum_{i=1}^n |x_i - b| \\ \mu_p &= \left(\frac{1}{n} \sum_{i=1}^n x_i^p \right)^{1/p} \quad (p = \pm 1, p \rightarrow 0; x_i > 0) \end{aligned}$$

などが語られる。裾が重く平均 (積分) が存在しない分布に対して平均値は一致性を持たないことやその数値実験、また特徴づけの一つとして母集団分布と MLE との関係もある。さらに上級の場合は、漸近分布や最適性 (e.g, 最小分散不偏推定量)、頑健性の理論であろうか。しかし、このよう内容はリテラシー教育には不適で不要である。

リテラシー教育・記述統計で何が抜け落ちているのかというと、代表値の利用目的である。

3. 代表値の利用目的と平均値と中央値

何らかの寿命のデータ X_1, \dots, X_n を考え、その母集団分布は左に歪んでいるとしよう (右に尾を引いた単峰の非対称分布, e.g. 図 1)。非対称な分布であるが故、教科書的には中央値を代表値として考えがちである。しかし、たとえば、 r 個のランプ在庫があり切れたら交換するという方法で何時間連続点灯させることができるかという問題に対して、寿命のデータ X_1, \dots, X_n を採取し代表値 x を求め $r \times x$ とすることを考える。このとき、 x として中央値 M を用いて rM で推定すると過小になり、実際の点灯時間は rM より長くなることが多い。適切な代表値は (算術) 平均値 \bar{X} である (MTTF, Mean Time to Failure)。その理由は $r\bar{X}$ は連続点灯時間の不偏推定量であるが、 rM は不偏ではないからである。リテラシー教育では不偏という専門用語は使わないので、 r 個を順にランプが切れたら置き換えていくと、どこかで相当に長い寿命のランプが出てくる (それが右に尾っぽを引いていることの意味)、だから結構長くなるのだ、と説明する。

では中央値の出番はないのか。そんなことはない。 n 人が同じランプを買って使用しそれぞれの寿命を比べるとする。もしあなたが購入したランプの寿命が中央値より短ければ運が悪かった

¹<https://bms.gr.jp/about/intro/>

²bulk, ひと塊り。

(or パチもんをつかまされた³)と感じるだろう。寿命が中央値より長ければ(少しは)ラッキーだったと感じるのではないか。あなたが購入したものがラッキーかアンラッキーかの比較対象として中位は機能するだろうということである。

このように、同一のデータであっても、目的によって適切な代表値は異なる。合計に注目する場合は平均値、順位ならば中央値というわけである。

信頼性工学では製品の保証すべき性質として寿命分布の90%点などが用いられる。考え方は順位という意味で中央値(50%点)と共通するものがある⁴。

世帯の収入や貯蓄の分布も同じように考えることができる(発表当日示す。)

4. 代表値の利用目的と平均値, 中央値, 最頻値

同一のデータで目的によって上述の3つの代表値を使い分ける例を見つけるのはなかなか難しい。図1は、大学周りに位置する学生向け賃貸集合住宅(アパート)の月額家賃のヒストグラムである(豊中市の石橋界限)。右に尾を引く非対称分布である。親の立場で子が住むアパートを考える場合、最頻値を参照するのではないか。多くの学生が住んでいるという理由である。大学が一定数の学生の家賃を補助する場合(そんなことはありえないが)、予算との関係で負担総額が問題になり、平均値が適切である。学生どうして家賃を話題にしたとき、上半分に入っているとリッチ、下半分だとあきらめる。友人の家賃の分布を参照し、自身の住環境(家賃)の順位が気になるのではないだろうか。

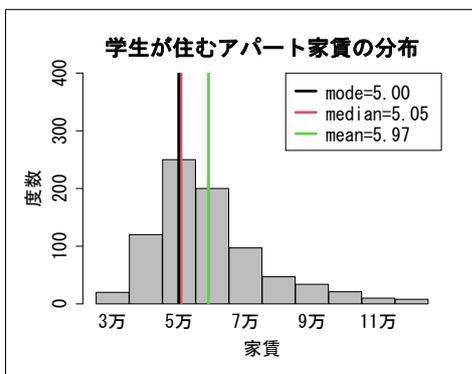


図1: 代表値の比較

5. まとめ

データや母集団の代表値は統計学であれDSであれ最初に登場する一見易しく見える履修項目である。定義は初等中等教育で必修になっており計算の習熟度も高い。では、代表値について高等教育や社会人教育(リカレント, リスキリング)では何を身に付けるべきなのか。計算機で置き換えられるような技術は今更磨いても仕方がない。ここで強調したいことは、代表値の選択においては、その分布と利用目的が重要であること、とくに利用目的の方が軽視されているのではないかということである。先に例を挙げたように、同一データであっても採用すべき代表値はその利用目的によって変わるのである。そして、そういった選択はAIや情報技術で置き換えられにくく、人間が判断する技術ではないか、ということである。

そのための重要な性質は、代表値の利用目的が、平均値はデータの合計に、幾何平均はデータの積に、調和平均はデータの逆数の合計に、中央値はデータの50%点に意味・関心がある場合に選択されるということである。たとえば、データの合計に意味があるときはその分布は考慮しなくてよいということがある。

$$\begin{aligned} \text{算術平均: } x_1 + \dots + x_n &= \overbrace{a + \dots + a}^{n \text{ 個}} \\ \therefore a &= \frac{1}{n} \sum_{i=1}^n x_i \end{aligned}$$

$$\begin{aligned} \text{幾何平均: } x_1 \times \dots \times x_n &= \overbrace{b \times \dots \times b}^{n \text{ 個}} \quad (x_i > 0) \\ \therefore b &= \left(\prod_{i=1}^n x_i \right)^{1/n} \end{aligned}$$

$$\begin{aligned} \text{調和平均: } \frac{1}{x_1} + \dots + \frac{1}{x_n} &= \overbrace{\frac{1}{c} + \dots + \frac{1}{c}}^{n \text{ 個}} \quad (x_i > 0) \\ \therefore c &= \frac{1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}} \end{aligned}$$

統計学・データサイエンスは、データやデータの分析結果を背景情報・ドメイン知識を照らし合わせながら、人間が逐次判断し分析を進めていくことで、現象を理解し最終判断に結びつける方法論である。常にデータと向き合い「考える」という姿勢を強調しておきたい。

参考文献

狩野裕(2016/01)「平均値 versus 中央値: 代表値を巡って」月間統計, pp.43-47, 日本統計協会.

³関西弁かもしれない。

⁴昨今、寿命分布は対称になることも多い。

文系学生のためのデータサイエンス教育

経済学部における実践と課題

田中昌宏・福岡大学経済学部
福岡市城南区七隈八丁目 19 番 1 号
文系センター棟 1109
gspddlnt45@toki.waseda.jp

1. 問題意識

データ活用人材に対するニーズの高まりを背景に、学習指導要領の改訂、データサイエンス(DS)に関連した学部等の新設など、近年、わが国における統計学・DS教育は急速に変化している。その一方で、昨今の統計学・DS教育を巡る議論は原則として学力上位層を念頭に置いており、“普通の”学生は議論から排除されているように思われる。しかし、社会におけるデータ活用を推進するには、“普通の”学生の教育を充実させ、データ活用人材の裾野を広げることも重要ではないだろうか。

本報告では、“普通の”私立大学文系学部におけるDS入門科目の実践例として、福岡大学経済学部で報告者が担当した「計量経済分析」を紹介する。また、“普通の”大学でDS教育を展開するためには、反転教室の成否が鍵であることを指摘する。

2. 実践例紹介

本報告で紹介する「計量経済分析」は、一切の前提知識のない学生を対象に90分×30コマでPythonプログラミングと機械学習の基礎を講じる科目である。

DS入門科目の先行事例の中には、一科目でプログラミング、確率・推測統計、機械学習のすべてをカバーするケースもみられるが、“普通の”大学で実施するにはカバレッジにメリハリをつける必要がある。この科目では確率・推測統計は一切扱わず、およそ2/3をプログラミング、残りを機械学習に充てた。プログラムについては、正規表現やオブジェクト指向といったやや進んだ事項は割愛した一方で、興味をもった学生が着実に二歩目に進めるよう基本事項はしっかりと網羅した。機械学習については、主要タスク(分類、予測、クラスタリング、次元削減)の考え方と応用例をカバーした一方で、各タスクで取り

上げる手法は一つずつに限定した。

3. 課題：反転教室

“普通の”大学でDS教育を展開するための鍵は反転教室である。プログラミング実習を含む科目は以下の理由から反転教室で実施することが望ましい。

(1)プログラミング経験、数学力、英語力などの影響により難易度の個人差が大きい。(2)習得するまでに大量の練習を必要とする。(3)初学者は自力でエラーを発見することが難しい。(4)“普通の”大学ではティーチングアシスタントを用意できないため学生同士で教え合ってもらう必要がある。

“普通の”大学で反転教室を実施するにあたって取り組むべき課題を二つ指摘する。

第一の課題は事前学習を促す工夫である。“普通の”大学には授業時間外に学習する習慣のない学生が多いため、事前学習を徹底させることが難しい。授業時間外の学習習慣の形成も含めた総合的なアプローチが求められる。

第二の課題は障害、特に発達障害・精神障害をもつ学生への合理的配慮である。発達障害・精神障害をもつ学生は年々増えており、“普通の”大学では配慮要請は決して珍しいものではなくなっている。こうした学生の修学上の困りごとは、反転教室を構成する重要な要素の一つであるアクティブラーニングの狙いと重なることが多い。例えば、自閉スペクトラム症の場合、口頭でのやりとり、対人関係の構築、緊張への対処などが困難となる可能性があるが、アクティブラーニングの本質を変えずにこうした困難に配慮することは非常に難しい。合理的配慮を実現するための体制づくりやアクティブラーニングのユニバーサルデザインについて業界全体で議論を深めると共に、各教育現場での経験を共有していく必要がある。

どのように共通教育と専門教育の「統計学」の橋渡しをするかII?

-2013年度と2022年度の学生に対する質問紙調査に基づく考察-

北海学園大学経営学部

関 哲人

北海学園大学工学部

速水 孝夫

金沢大学人間社会研究域学校教育系・学校教育学類 米田 力生

代表連絡先(関 哲人): nseki@ba.hokkai-s-u.ac.jp

1.背景

統計学を学ぶにあたって、問題解決のサイクルを意識することになる[1]。また、近年はデータサイエンスの社会的要請が高まっている。現状を的確に理解するために正しくデータを分析・活用し、適切な意思決定を遂行し、課題を解決し新しい価値を創ることができる人材(データサイエンティスト)が求められており [2]、文部科学省もデータサイエンスの普及を推進している[3]。このことから大学教育に統計学の必要性がより高まっていると言える。

2.本学における統計学講義の現状

本報告の問題意識は前項で述べたような一連の行為を授業に反映させることになる。他方、大学における大人数教育は知識中心の授業デザインであり、「憶える、そしてできるだけ正確に再生する」[4]ことを主眼に置くものである。

例えば現状からデータ問題への帰結、意思決定、価値創造といった行為は知識の応用・実践である。よって大人数での講義形式による統計学の講義では、これらの一連の行為を念頭に置きつつ、これらに少しでも関わりを持たせることが求められる。

2.1 共通教育(一般教育)における統計学

大学における統計学の講義は共通教育(一般教育)と専門教育で大別される。本学は経済学部、経営学部、法学部、人文学部、工学部による総合大学である。本学での共通教育はこれら5学部に対し、一般教養の学びの機会を提供するのが責務である。共通教育科目は基盤科目(言語、身体、情報)と教養科目(人文科学、社会科学、自然科学、北海道学)にて構成されている。共通教育科目の自然科学分野では、さらに小分野(環境、普遍)があり、小分野の普遍における講義科目として、「数

学概論I,II」、「統計学I,II」、「物理学概論」が開講されている。

文系学部向けの数学概論I,IIでは、近年はそれぞれ線形計画法と魔法陣の数理を扱っており、微積分や線形代数といった大学の専門教育で用いる数学の基礎を積極的に扱ってはいない。「いろいろな数学を知る」ことを主眼としており、教養としての側面が強い。

統計学の科目としては統計学I,IIがそれぞれ半期ずつ、工学部を除く全学部で開講されている。統計学I,IIは問題解決の一連の行為における現状からデータ問題への帰結に重点を置く形になるが、統計学を通じて教養としての数学を学ぶ目的も有している。統計学I,IIはそれぞれ1部3クラス、2部1クラスの開講である。経済学部・経営学部の受講生が例年半数以上を占めている。教科書は石村園子(2006)『やさしく学べる統計学』、共立出版を使用している。なお、工学部においては、文系学部における統計学I,IIに該当する科目として、文系学部での授業内容を若干理系向けにアレンジした科目を、1・2年生対象の専門科目として開講している。

2.2 専門教育における統計学

本学経営学部では、統計学の専門講義として経営統計学概論I,IIが位置づけられている。本講義では吉田耕作(2006)『経営のための直感的統計学』、日本経済新聞社を教科書に指定している。すでに課題が設定されていることを前提に、データの記述と分析、結果の統計的解釈を重視したものとしている。

本講義は1年次配当科目で、経営学部で必要な統計学の修得を目的としており、情報関連科目、心理学関連、マーケティング関連の科目との結びつきを踏まえた内容にしている。上級年次では、本講義を学んでいることを前提に、専門科目で統

計学を用いた講義が存在する。さらに、本学部では情報教育に力を入れているため、情報実習科目における操作・算法の背景を説明する役割も持つ。この点で実用性も重視した内容である。

また、本学部は文系学部であるため、できる限り数学を用いなくて統計学を修得できるように配慮している。そのため、本学部における統計学は多くの統計学の講義で見られるような数式主体、板書中心ではなく、図形描画、表計算、Excel操作を中心とした講義としている。ただし、数式の証明は最小限にとどめている。統計学の理論の数式の証明については一般教育に期待する面を持たせている。

表1 本学における統計学講義の単元対応表

	共通教育		専門教育	
	統計学I	統計学II	経営統計学概論I	経営統計学概論II
確率の基礎	○		○	
条件付き確率	○			
確率変数 確率分布	○			
正規分布	○		○	
記述統計	○		○	
質的データの 処理	○			○
点推定		○		
区間推定		○		○
検定		○		○
回帰分析		○		
t分布		○		○
χ^2 乗分布		○		

2.3 共通教育と専門教育の連携の必要性

こうした背景から本学でも共通教育と専門教育の「統計学」講義の連携を模索する必要がある。本稿は、連携の模索に該当する議論である。

共通教育と専門教育の単元の対応は表1になる。専門教育の統計学は数式をあまり用いない、板書主体の講義ではないためか、共通教育と比べ扱う単元が少ない。共通教育では専門教育で扱っていない確率や回帰分析を取り扱っている点にも注目したい。

3. 統計学教育の改善例

「統計学」講義において共通教育で主眼とする数学的素養と専門教育で主眼としている実用性・具体性を結びつけるべく、2014年1月にそれ

ぞれの受講生に対して質問紙調査を実施し、調査結果から模索を試みた [5]。専門教育の統計学については、板書を主体としていない講義スタイルであるため、高校の数学の授業で学ぶ統計学と違いを感じるようであった。他方、共通教育の統計学とは扱う内容が重複するものもあり、共通教育の統計学の受講生が専門教育の統計学とあまりギャップを感じているわけではないようであった。また、共通教育と専門教育で重なる単元について興味を持っていることも調査結果から分かった。これを踏まえて、それぞれの科目で次の対策を講じた。

3.1 共通教育における教育の改善の実践

共通教育の統計学では、学問としての統計学、そして数学やデータサイエンスとの関わりについても触れながら講義を進めている。

授業形態については、これまで主として板書による説明を中心としていたが、やはり数式の扱いに習熟していない文系の学生にとっては少しストレスになるとは思われた。また、2015年度以降入学生の多くにとっては、学習指導要領の改訂により、記述統計学の基礎的な内容について高校のときにすでに学習済みの内容となってしまった。もちろん、数学が得意ではない学生への配慮も必要であることから、記述統計学の内容をカットすることもできず、この部分を板書で長々と説明するのも学生にとってはストレスになっていたと思われる。このこともあってか、一時期、共通教育の受講生数が低下してしまった時期があった。このような状況を改善すべく、まず学部によるクラス指定を撤廃し、履修しやすくした。そして、共通教育の統計学の授業でもスライド資料の活用も試みることにした。説明事項などについてはスライド資料を十分に活用し、問題演習については板書を十分に活用するというスタイルをとることとした。受講生にとっては、特に高校の復習の内容を扱う場面では概ね好評だったようである。

さらに、専門の統計学への接続をさらに意識するようにし、これまでに扱わなかった等分散性の検定などについても、追加で講義内容に含めることとした。また、専門の統計学の科目がない学部（法学部、人文学部）の受講生も一定数いることから、こちらの学生への配慮も必要であった。

今後、高等学校の新学習指導要領においては、数学Bは「数列」、「統計的な推測」、「数学と社会生活」から構成され、推測統計学を勉強する高校も今後は多くなってくることが想定される。このことにも配慮し、共通教育の統計学の中で扱う内容も少し見直しをしていく時期に来ていると思われる。

3.2 専門教育における教育の改善の実践

専門教育の統計学では、経営学部専門科目の関連の強化と数学のかかわりをより説明することを心掛けた。

専門科目とのかかわりについては、例えば確率の基礎の箇所では、数理的意思決定も含めることで心理学とのかかわりを説明した。質的データの単元では協定連携企業と共同でのアンケート調査の実施(グループワークを実施した年度もあり)[6]、マーケティングにおけるクロス集計表の活用を取り上げた。

次いで、共通教育に対する配慮として、数式の詳しい証明などについては共通教育の受講を進めた。また、相関係数の箇所では線形概念も取り上げつつ、対数関数などの非線形関数を紹介した。

4. 統計学に関するアンケート調査と今後の展望

共通教育と専門教育との連携の成果を模索の第1段階として、統計学II、経営統計学概論IIにおいて2023年1月に「統計学に関するアンケート」を実施した。これは2013年度の調査と比較することも兼ねている。調査項目と集計結果の詳細は付録2を確認されたい。

設問1における履修理由において「数学が得意だから」というわけではないことが分かった。また設問2では半数近くが数学を受験勉強で使っている。設問5において、2013年度の専門教育では記述統計に興味を持ったという回答が目立ったが、2022年度では減少した。これらから高校で統計学教育の一定の普及がみられると言える。

設問6では、2022年度の共通教育では確率の基礎が難しいと答える学生が目立った。これは順列・組合せ・標本空間・数学的確率・統計的確率といった概念をさらに丁寧に扱ったが、これらに苦手意識を持つ学生がこちら側の想定よりも実際は多かったことが考えられる。一方で、専門教

育では区間推定、検定、t分布といった推測統計が難しいという回答が多かった。

設問7,8ともに専門科目では統計学について違いを感じるという回答を得た。この違いとは設問7での高校数学の授業との違いを示している。ただし、設問8における「違いを感じる」は統計学、データサイエンスにおける一連の活動で重視している個所の違いによるものに起因するものである必要がある。そのためには、双方の科目が一連の活動の下での取り組みであることをより意識させる必要もあろう。

参考文献

- [1]渡辺美智子(2011)「21世紀型ワークスキルとしての統計教育改革-スキャンズレポートとデミング博士-」、『統計教育実践研究』第4巻
- [2]北海道大学数理・データサイエンス教育研究センターWebサイト
<https://www.mdsc.hokudai.ac.jp/>
(最終閲覧日 2023年2月25日)
- [3]数理・データサイエンス・AI教育(文部科学省Webサイト)
https://www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/00001.htm
(最終閲覧日 2023年2月25日)
- [4]大島純・野島久雄・波多野誼余夫(2006)『教授・学習過程論』、放送大学教育振興会
- [5]関哲人・速水孝夫・米田力生(2014)「どのように共通教育と専門教育の「統計学」の橋渡しをするか?-共通教育の統計学の講義における実態調査に基づく考察-」、『統計教育実践研究』第6巻
- [6] 関哲人・鈴木雄大・山岡英二・油谷弘洋・久保田佳輝・並木翔太(2019)「経営学部における統計学講義で実践した企業との連携協定に基づくグループワークの成果報告」『統計教育研究実践研究』特別号

付録1：本学における数学の受験状況

本学における一般入試において、人文学部を除く文系学部においては、「数学」が選択教科になっている。数学の試験では、「数学Ⅰ・A」から必修問題を、さらに「数学Ⅰ・A」、「数学Ⅱ」、「数学B」の中から選択問題(試験時に選択する)を出題している。選択問題については、やはり「数学Ⅰ・A」の問題を選択する傾向が強いようである。以下に示すように、一般入試の選択科目の試験において、数学を選択する率が一番高かった(おおむね40%程度)。

- | |
|---|
| <p>・選択科目の選択率(2013年入試)</p> <p>経営学科 数学 40.9%, 政治経済 27.1%,
日本史 20.7%, 地理 8.6%, 世界史 2.7%</p> <p>経営情報学科 数学 38.5%, 政治経済
28.7%,
日本史 23.9%, 地理 6.6%, 世界史 2.3%</p> <p>・選択科目の選択率(2022年入試)</p> <p>経営学科 数学 43.9%, 政治経済 21.3%,
日本史 21.7%, 地理 8.3%, 世界史 4.7%</p> <p>経営情報学科 数学 35.7%, 政治経済
32.7%,
日本史 22.4%, 地理 6.7%, 世界史 2.9%</p> |
|---|

付録2: 統計学についての質問紙調査(アンケート調査)

本調査は「統計学」について実態を問うており、左段は2014年1月20日頃、右段は2023年1月20日頃に共通教育、専門教育のそれぞれの授業で質問紙調査(アンケート調査)を行ったものである。多少、選択肢を変えたところもあったが、前回との比較がしやすいように、概ね同一の質問項目とした。

設問1. 「統計学」の授業を履修したきっかけは？

2013年度版

※複数回答	共通教育	専門教育
統計学の内容に興味があるから	19	51
専門で必須の知識だから	7	175
単位が必要なので	31	162
数学が得意だから	17	21
その他	15	13

2022年度版

※複数回答	共通教育	専門教育
統計学の内容に興味があるから	22	54
専門で必須の知識だから	6	79
単位が必要なので	9	113
数学が得意だから	9	19
時間割の都合	3	91
その他	9	20

設問2. 数学の試験を受験しましたか？

2013年度版

	共通教育	専門教育
受験した	43	118
受験しなかった	33	133

2022年度版

	共通教育	専門教育
受験した	38	141
受験しなかった	20	191

設問3. 大学受験時に数学を勉強していた人へ質問します。数学の受験勉強を行った科目を教えてください。

2013年度版

※複数回答	共通教育	専門教育
数学I	54	142
数学II	46	125
数学A	49	137
数学B	45	123
数学III・C	6	11
数学の受験勉強は特にしていない	回答なし	102

2022年度版

※複数回答	共通教育	専門教育
数学I	42	55
数学II	36	44
数学A	40	14
数学B	36	39
数学III	5	6
数学の受験勉強は特にしていない	9	89

設問4. 共通教育または専門教育の「統計学」に関する授業を履修しましたか？

2013年度版

※複数回答	共通教育	専門教育
専門の統計学を履修した	63	
専門の統計学を履修していない	12	
統計学Iのみを履修した		79
統計学IIのみを履修した		0
統計学I,IIともに履修した		91
どちらも履修していない		81

2022年度版

※複数回答	共通教育	専門教育
専門の統計学を履修した	21	
専門の統計学を履修していない	36	
統計学Iのみを履修した		18
統計学IIのみを履修した		0
統計学I,IIともに履修した		83
どちらも履修していない		223

設問 5.「統計学」の授業内容で興味を持ったことがあれば教えて下さい。

2013 年度版

※複数回答	共通教育	専門教育
確率の基礎	19	
条件付き確率	4	
確率変数、確率分布	12	
正規分布	38	88
記述統計	7	115
点推定	6	
区間推定	31	73
検定	44	77
回帰分析	6	
t 分布・ χ^2 乗分布	6	
t 分布		29

2022 年度版

※複数回答	共通教育	専門教育
確率の基礎	22	77
条件付き確率	13	
確率変数、確率分布	12	
正規分布	21	52
記述統計	5	37
点推定	11	
区間推定	30	45
検定	21	44
回帰分析	8	
t 分布・ χ^2 乗分布	8	
t 分布		19

設問 6.「統計学」の授業内容で難しかった内容があれば教えて下さい。

2013 年度版

※複数回答	共通教育	専門教育
確率の基礎	6	
条件付き確率	10	
確率変数、確率分布	24	
正規分布	11	43
記述統計	32	73
点推定	7	
区間推定	12	86
検定	30	122
回帰分析	19	
t 分布・ χ^2 乗分布	22	
t 分布		97

2022 年度版

※複数回答	共通教育	専門教育
確率の基礎	30	8
条件付き確率	4	
確率変数、確率分布	9	
正規分布	5	25
記述統計	8	37
点推定	14	
区間推定	16	82
検定	21	98
回帰分析	24	
t 分布・ χ^2 乗分布	27	
t 分布		81

設問 7.「統計学」の授業を通して、高校生の時の数学の勉強と比べてギャップがあると感じましたか？

2013 年度版

	共通教育	専門教育
かなり感じた	18	99
少し感じた	34	119
あまり感じない	16	17
ほとんど感じない	5	12
分からない	2	4

2022 年度版

	共通教育	専門教育
かなり感じた	13	103
少し感じた	33	137
あまり感じない	10	53
ほとんど感じない	2	15
分からない	0	23

設問 8.「統計学」の授業を通して、高校の授業と比べて内容の説明の仕方に違いがあると感じましたか？

2013 年度版

	共通教育	専門教育
かなり感じた	25	84
少し感じた	31	126
あまり感じない	16	27
ほとんど感じない	3	6
分からない	1	8

2022 年度版

	共通教育	専門教育
かなり感じた	10	112
少し感じた	28	142
あまり感じない	16	49
ほとんど感じない	4	17
分からない	0	12

仮説検定・統計解析によるアンケート分析手法の実践報告

東京都立大島海洋国際高等学校 片江 康裕

E-mail: ykdef0825a@ac.auone-net.jp

1 実践概要

仮説検定・統計解析によるアンケート分析の成果は、令和元年度に実施された第74回関東甲信静数学教育研究大会にて発表しており、対象とする試験（計算）問題の自己評価の数値化に関する考察等を示した。以後、数学教育学会並びに日本数学教育学会で発表を継続し、2022年度における数学教育学会春季年会では、神戸市の学校で実施した成果を通じて、数年間の実践に関する総括を行っている。過去の実践では、計算問題における難易度の自己評価アンケートとして行ってきたが、本年度は、「情報Ⅰ」の学力試験に対する自己評価のアンケートをCBT方式で実施した。本実践を通じてこれまでの成果を再確認し、Webシステム化に関して、今後の方針を示す。

2 これまでの経緯

過去の実践では、計算能力に対する自己評価アンケートを継続的に行い、評価の方法について考察してきた。残差の分析により、対象とする試験の各問題の傾向に関する考察が可能であることに特に注目してきたが、仮説検定を棄却できず、問題群にまとめて問題群の傾向のみを確認する場合も幾度かあった。この課題は、「易しい」、「普通である」の回答に関して正答者の回答のみ有効とみなし、各問題の正誤とアンケート結果の整合性を考慮することにより解決を図った。

アンケート結果の解析には表計算ソフトを用いているが、初期の頃は紙面で行っていたため、目視で数えて集計する作業を伴った。その後、GoogleFormsを活用し、結果の入力方式を電子化したアンケートを実施することが可能となった。出力に関しては、表計算ソフトによる解析処理を手動で行う作業は伴うが、解析処理の効率を上げることには成功している。

本年度の実践では、入力側はGoogleFormsとし、出力側に関してもWebシステム化を試みた。Web化

に伴い、GoogleFormsから出力されるデータの編集を最小限に留め、Web上で解析処理が可能なプログラムを組み込む等を試みることにより、作業の効率化を図っている。図1は、GoogleFormsで作成した試験の入力画面である。



図1 入力画面 (GoogleForms)

3 仮説検定・統計解析に基づくアンケート分析

3.1 アンケートの実施方法

一般的なテスト問題を解く際に、各問題の難易度の自己評価を回答するように促す。本実践では図2のように、問題解答の後に設問を用意し、各々の問題に対する難易度を「簡単だった」、「普通である」、「難しい」の3段階で回答させる択一式としている。アンケートは試験と同時にCBT方式で実施しており、全ての問題の解答並びにアンケートの回答を要求し、問題を解くと同時にアンケートに対して必ず回答しなければならない仕組みとなっている。本実践では、「情報Ⅰ」の試験問題をアンケートの対象として、分析を行っている。



図2 アンケート出題形式

過去に実施したアンケートでは、難易度を5段階としていたこともある。

アンケート結果は、統計学における独立性の検定に基づくクロス集計により分析を行う。集計したデータを解析し、残差の分析に注目して各々の問題の難易度の傾向が把握できるようになっている。

独立性の検定に基づくクロス集計では、仮説検定により、アンケートの有効性を検証すると同時に、相関分析の傾向も確かめている。過去の実践では、仮説検定の帰無仮説を棄却することができず、幾つかの問題を1つの問題群とした分析も行うことがあったため、「易しい」と回答した人数と正答人数の相関によりアンケートの妥当性を確かめる必要があった。この課題については、日本数学教育学会第103回全国大会で発表した実践より、問題の正誤との整合性を考慮し、妥当性が見込めないアンケート結果を排除することで解決を図っている。[5][6]

3.2 理論・結果の分析方法

独立性の検定に基づくクロス集計、及び相関分析による統計解析を行った。推測統計学的な解析に対して、仮説検定により妥当性を確認した上で、結果を示している。クロス集計表の縦（難易度） A_i ($i = 1 \sim 3$) の項目と横（各問） B_j ($j = 1 \sim 23$) の項目における期待度数 E_{ij} を求め、統計量 χ^2 の分布を計算して仮説検定を行っている。

$$E_{ij} = \frac{\sum X_i \sum X_j}{\sum X_{ij}}, \quad \chi^2 = \sum_{i=1}^3 \sum_{j=1}^{23} \frac{(X_{ij} - E_{ij})^2}{E_{ij}}$$

検定を行うために関数「CHIINV」を用いて、自由度に応じた有意水準5%の統計量を求め、統計量の値が棄却域にあるかどうかを確かめた。帰無仮説が棄却され、関連性があることが分かることにより、問題の難しさの自己認識と各問に関係性があることが分かる。さらに、調整済み残差分析を行うことにより、各項目について有意な差を比較することができる。各項目における調整済み残差は以下のように計算する。

$$Z_{ij} = \frac{X_{ij} - E_{ij}}{\sqrt{E_{ij} \left(1 - \frac{X_i}{\sum X_i}\right) \left(1 - \frac{X_j}{\sum X_j}\right)}}$$

アンケート対象とする回答の選別については、単に正答した回答だけでなく、「簡単だった」については正答のみ対象とし、「普通である」と「難しい」については、正答・不正答に関わらず解析の対象とした。

3.3 相関分析

相関係数の仮説検定では、フィッシャーのZ変換を用いて統計検定量を求める。有意水準は5%である関数「NORMINV」により検定を行う。（標本数を n としている）

$$\text{相関係数: } R_{xy} = \frac{S_{xy}}{\sqrt{S_x^2 S_y^2}} \quad \text{帰無仮説: } \rho$$

$$Z = \frac{1}{2} \log \frac{1+R_{xy}}{1-R_{xy}}, \quad \zeta = \frac{1}{2} \log \frac{1+\rho}{1-\rho}$$

$$\text{検定統計量: } Z_0 = \sqrt{n-3}(Z - \zeta)$$

4 結果・考察

4.1 分析結果

過去の実践と同様に表計算ソフトによる統計解析を行い、独立性の検定に基づくクロス集計の帰無仮説を棄却し、調整済み残差分析の結果を得た。また、「易しい」と回答した人数と正答人数との相関としていたが、本実践では、正答人数と「簡単だった」と回答した正答人数の相関をとっている。アンケートの単純集計結果と調整済み残差分析の結果、及び相関分析の結果は以下のとおりである。

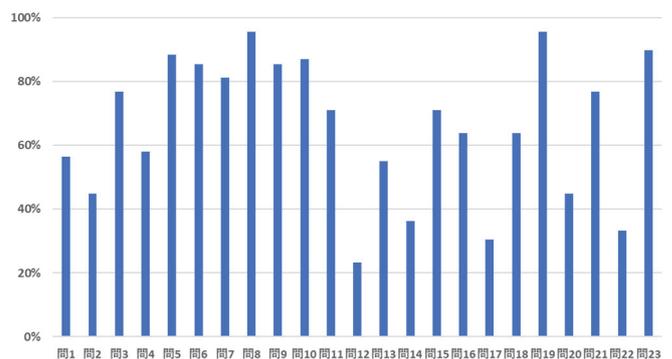


図3 単純集計結果（問題別正答率）

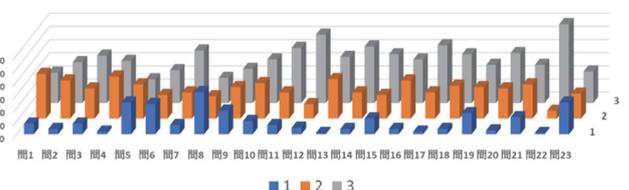
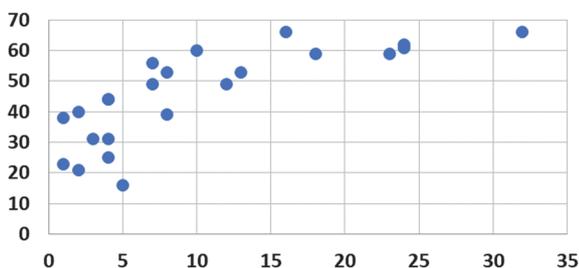


図4 単純集計結果（問題別難易度の回答）

表 1 調整済み残差分析の結果

	問1	問2	問3	問4	問5	問6	問7	問8	問9	問10	問11
簡単	-0.64	-2.02	-0.73	-2.80	4.77	4.59	-1.08	7.54	2.69	-0.19	-1.17
普通	3.21	1.98	0.08	2.57	0.78	-1.15	-0.71	-1.58	0.26	0.85	-0.87
難しい	-2.59	-0.43	0.45	-0.43	-4.15	-2.19	1.44	-3.90	-2.17	-0.67	1.67

	問12	問13	問14	問15	問16	問17	問18	問19	問20	問21	問22	問23
簡単	-1.82	-3.15	-2.13	0.66	-2.09	-2.80	-2.09	1.93	-2.37	0.95	-3.18	4.85
普通	-3.15	2.04	-0.71	-1.24	1.77	-0.62	0.71	0.17	0.36	0.78	-4.40	-0.97
難しい	4.29	0.33	2.19	0.70	-0.18	2.59	0.83	-1.54	1.36	-1.42	6.44	-2.55



	簡単のみ	簡単+普通
標本相関係数	0.77	0.94
帰無仮説	$\rho = 0.5$	$\rho = 0.6$
対立仮説	$\rho > 0.5$	$\rho > 0.6$
標本数	23	23
z	1.02	1.70
ξ	0.55	0.89
検定統計量	2.15	4.52
棄却域	1.64	1.64

図 5 相関分析の結果

(「簡単」の正答人数と正答人数の相関)

4.2 考察

分析方法については、アンケート結果について、「簡単だった」と回答した場合は正答のみ有効とするようにした。アンケートに対する回答が不正確である場合も考えて、「簡単だった」と回答した正答人数と正答人数の相関分析も行うことによって、アンケートの妥当性を裏付けている。

2022年度数学教育学会春季年会における発表では、「普通である」と回答して問題が不正解だった場合も無効と判定したが、本実践では正誤に関係なく、「普通である」を有効としている。「普通である」の回答に問題の正誤を考慮するかどうかは、アンケートの結果に応じて判断すべきと考える。[6]

残差分析により、各問題に対して数値が示され、プラス方向が「有意に多い」、マイナス方向が「有意に少ない」となる。各々の選択肢の回答数が有意に多いか少ないかで、対象としている問題に対する集団の意識の傾向を読み取ることができる。

「簡単である」の残差がプラス方向に高い値であれば「得意である」と考える人が多く、マイナス方向に高い値であれば「得意である」と考える人は少ないと判断できる。「普通である」、「難しい」についても

同じように捉えることで、単純な正誤だけでは読み取れない問題の解答者やその集団の意識の傾向を各問題からくみ取り、どの分野に重点を置いて学習指導を行うべきかを診断することも可能となる。

解析結果から、問題の解説は問 12, 問 14, 問 17, 問 22 を重点的に行えばよいことがわかる。特に問 22 は、「難しい」が極端にプラス方向に向かい、「簡単だった」と「普通である」もマイナス方向に高い値が出ているので、問題に対して丁寧な事後指導を要することが一目で分かる。この傾向は、図 4 のグラフからも読み取れ、結果を数値化することによって他の問題と傾向を比較しやすくなり、優先順位を付ける際に役立つ。

各問題の残差分析結果は、アンケートの項目ごとの分布の傾向と関係があることも読み取れる。表 2 には、本実践における類似傾向をまとめている。

表 2 類似性の比較

	問5		問6		問8		問9		問23	
	人数	残差	人数	残差	人数	残差	人数	残差	人数	残差
簡単	24	4.77	23	4.59	32	7.54	18	2.69	24	4.85
普通	26	0.78	18	-1.15	17	-1.58	24	0.26	19	-0.97
難しい	18	-4.15	25	-2.19	19	-3.90	26	-2.17	24	-2.55

	問2		問4		問13		問16		問18		問20	
	人数	残差	人数	残差	人数	残差	人数	残差	人数	残差	人数	残差
簡単	4	-2.02	2	-2.80	1	-3.15	4	-2.09	4	-2.09	3	-2.37
普通	29	1.98	32	2.57	30	2.04	29	1.77	25	0.71	23	0.36
難しい	31	-0.43	32	-0.43	35	0.33	33	-0.18	37	0.83	38	1.36

5 Web システムについて

本実践より、Web システム化を見据えて、結果の表示の自動化・電子化を試みている。入力 Form に関しては、既に GoogleForms を活用したアンケートを実施しており、データ回収・データ編集の作業効率向上を実現している。解析処理・結果表示には、GoogleColaboratory を活用しており、作業の効率化を図り、調整済み残差の分析結果だけでなく、集計結果等を表示させる Web システム化にも着手した。現段階では、入力 Form と出力画面は別の Web システムとなり、出力されたデータの受け渡しの際に、人的作業を介して出力側にデータを読み込ませる必要がある。図 6 のように、GoogleForms と解析・画面表示の 2 つのシステムに分かれており、編集を加えた CSV ファイルを読み込ませる人的作業を伴うシステムとなっている。今後は 2 つの Web システムを統合

したシステムの構築を目指すことを考える.解析・画面表示については, Web 上における表計算ソフトの解析処理のプログラム化の着想等を, 文献[7]で解説している.

GoogleColaboratory により, グラフ表示も可能となったが, 入力 Form との一体化を見据えたシステム化を計画的に進める. 開発業務に関しては, 協力者等にある程度委託することを検討している.

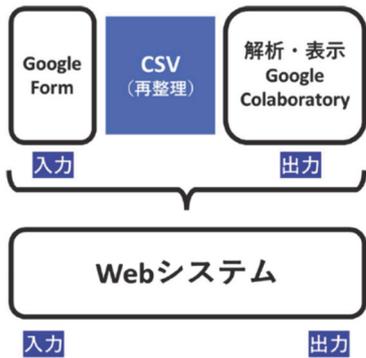


図 6 システムの一体化 (イメージ)

図 7, 図 8 は, Google Colaboratory を活用して作成した Web システムの表示画面である.

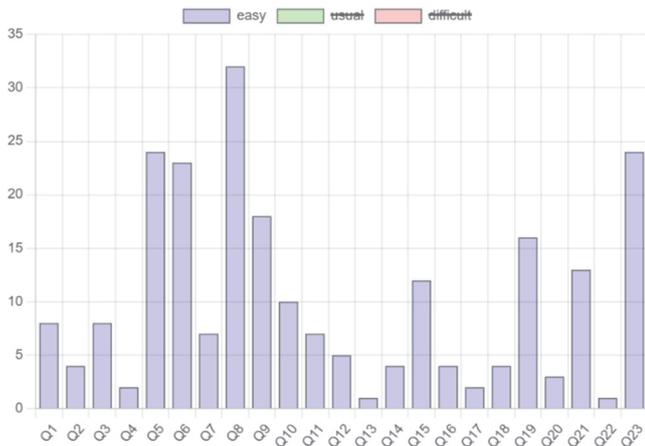


図 7 出力側の画面①
(問題別難易度の回答・簡単/正答のみ)

分析結果を表示します。ファイルを選択して読み込みを行ってください。

欄: 変換 (解したい), 最終行: 検定統計量, 変換率, 判定 (KKから変換)

欄: 問題 問題1→最終問題, 残差 (値) > [1.96]で検定 (有意に多い・少ない)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	-0.63	-2.01	-0.73	-2.79	4.766	4.589	-1.08	7.539	2.686	-0.18	-1.17	-1.82	-3.14	-2.12	0.663	-2.09	-2.79	-2.09	1.927	-2.37	0.953	-3.17	4.852
2	3.212	1.978	0.084	2.567	0.781	-1.15	-0.70	-1.57	0.256	0.853	-0.87	-3.15	2.035	-0.70	-1.23	1.769	-0.62	0.706	0.166	0.359	0.781	-4.40	-0.97
3	-2.58	-0.43	0.445	-0.42	-4.15	-2.19	1.444	-3.90	-2.16	-0.67	1.665	4.285	0.325	2.194	0.695	-0.17	2.590	0.828	-1.53	1.357	-1.42	6.442	-2.55
4	272.8	60.48	KK																				

図 8 出力側の画面②
(調整済み残差分析の結果)

参考文献

[1] 片江康裕: 計算能力と学習意欲の関係性に関する一考察 主体的な学びを促す方法の確立を目指して, 第 74 回関東甲信静数学教育研究大会千葉大会要項, p.148 (2019)

[2] 片江康裕: 計算能力及び学習意欲のアンケート結果における統計解析の手法に関する一考察 主体的な学びを促すための実践方法を求めて, 数学教育学会 2020 年度春季年会予稿集, (2020)

[3] 片江康裕: 仮説検定・統計解析によるアンケート分析手法 クロス集計・相関分析結果の傾向とその考察, 第 102 回全国算数・数学教育研究 (茨城) 大会発表要旨集, p.366 (2020)

[4] 片江康裕: 仮説検定・統計解析によるアンケート分析手法の電子化, 数学教育学会 2021 年度春季年会予稿集, (2021)

[5] 片江康裕: 数学を楽しむ授業の工夫とその考察, 第 103 回全国算数・数学教育研究 (埼玉) 大会要旨集, p.424 (2021)

[6] 片江康裕: 仮説検定・統計解析によるアンケート分析手法に関する考察, 2022 年度数学教育学会春季年会予稿集,(2022)

[7] 片江康裕: 仮説検定・統計解析によるアンケート分析手法の Web システム化の提案, 情報処理学会 CLE 第 39 回研究会, (2023)

[8] 岩佐英彦・宿久洋: 授業評価・市場評価のための「アンケート」調査・分析ができる本, 株式会社秀和システム, (2009)

[9] 掌田津耶乃: Colaboratory でやさしく学ぶ JavaScript 入門, 株式会社マイナビ出版, (2022)

[10] 東京大学教養学部統計学教室編: 基礎統計学 I 統計学入門, 東京大学出版会, (2020)

[11] 森下悦生: Excel で学ぶ流体力学, 丸善出版株式会社, (1999)

[12] Google 社: Google アカウント, <https://www.google.com/>, (2023.02.27)

[13] Google 社: Google Colaboratory, <https://colab.research.google.com>, (2023.02.27)

[14] チャート (グラフ) :Chart.js, <https://www.chartjs.org/>, (2023.02.27)

「中高生・スポーツデータ解析コンペティション」について

日本統計学会スポーツデータサイエンス分科会他が主催で中高生向けスポーツデータ解析コンペティションを開催しています。中学生・高校生に対して、実際の日本のプロ野球、プロサッカーおよびプロバスケットボールなどのデータ（開催年で多少変わります）を貸与しますので、各チームでテーマを自由に設定し、統計的な方法でデータ分析し、その内容をポスター発表形式で競います【規定部門】。また貸与データを使わずに各チームが独自で収集したスポーツデータを用いた分析を競う【自由部門】もあります。スポーツやデータ分析に興味のある人はぜひ参加ください。参加に際しては、学校の先生を代表者として申し込んでください。なお外部データ利用申請により、活動成果をSSH研究会や統計グラフコンクールなどに発表することも可能です。

参加費は無料です。

■スケジュール

8月：申込受付、データの配布開始。※9月末まで受付

チーム（1チーム2人～5人をお願いします。一人でも構いません）で、貸与データやオープンデータを利用して、新しい事実の発見、戦略提案、などの分析を行います。学校の先生（チームの代表、責任者）を通じてお申し込みください。申し込み後、参加条件等を確認し、貸与データをお送りいたします。

1月：締切までに分析の仕上げ、ポスター作成

各チームは締切までに作品として、【B2 サイズ】のポスターを一部作成してください。提出はPDFで事務局サイトにて、【1月末、23:59】までにご提出をお願いします。

2月：「統計・データサイエンス教育の方法論ワークショップ」内で審査結果発表

2月下旬から3月上旬に開催予定の「統計・データサイエンス教育の方法論ワークショップ」内で、審査結果の発表、および作品の紹介を行います。ワークショップ終了後、授賞チームに賞状（最優秀賞、優秀賞、奨励賞には盾も）を送付いたします。3月末をもって貸与データの削除をお願いします。またメールおよびウェブでも結果は告知します。

■審査基準

今回の審査では、以下の基準で評価します。

新規性：データ分析として新しいことにチャレンジしている。

論理性：分析の手続きが論理的に適切であり、分析ストーリーがある。

妥当性：データの処理、分析の方法が適切にできている。

データ利活用力：貸与または持参のデータを効果的・適切に利活用している。

表現力：分かったことを適切にポスターに表現できている。

なお審査員については、統計教育関係研究者・教育者、スポーツデータに関する実務者などに依頼しています。その他の情報につきましては、日本統計学会統計教育分科会のサイト (<https://estat.sci.kagoshima-u.ac.jp/SESJSS/>) でも告知いたしますので、ご参照ください。

■申込方法

ウェブサイト (<https://hs.sports.ywebsys.net/>) から登録

代表者（教員）は登録後に生成される誓約書を印刷し署名及び押印したものをスキャンしてポータル上で提出し、メンバー（生徒）署名をポータルサイト上で登録となります。

日本統計学会統計教育分科会のお知らせおよび入会案内

統計教育分科会 主査

宮崎大学 藤井良宜

今、日本の統計教育は大きく変わろうとしています。

平成 20 年 3 月に小学校、中学校の学習指導要領の改訂が行われました。新しい学習指導要領では、小学校算数において度数分布表の取り扱いが復活し、中学校数学では新しい領域「資料の活用」が新設されました。「資料の活用」では、1 年生から 3 年生まで、毎年統計や確率に関する内容を学習することになっています。さらに、平成 21 年 3 月に告示された高等学校学習指導要領では、必修科目である「数学 I」の内容の中に、「データの分析」が加わり、すべての生徒がこの内容を学習することになりました。このことは、前回の学習指導要領の改訂では統計的な内容が大幅に削減されていたわけですので、大きな方針の転換が行われたことを意味しております。また、現代社会におけるデータの重要性やデータを処理分析するための統計のニーズの高まりが反映されているとも考えられます。

しかし、学習指導要領の改訂だけでは十分な改革は望めません。実際に授業を行う先生方の授業実践こそが重要なのです。日本統計学会では、統計学の研究を通して統計の専門家や統計の教育者を育成するばかりではなく、統計教育の実践も積極的に支援しております。その一つの方策として、平成 14 年 11 月の日本統計学会の評議会において、統計教育の普及・発展に貢献することを目的として活動する統計教育分科会が設置されました。統計教育分科会では、毎年 3 月に行っております統計教育ワークショップをはじめとして、これまで次のような活動を行ってきております。

- 小・中・高校・大学における統計教育、あるいは社会人等を対象とする統計教育の実践、ならびに研究に関する研究会の開催
- 各種統計教育の実践、ならびに研究のための作業グループ、臨時プロジェクト等の活動
- 統計教育に関する他の学会、諸団体との共同プロジェクトの実施
- その他分科会の目的を達成するための適切な作業

活動状況は、公式ウェブページ (<https://estat.sci.kagoshima-u.ac.jp/SESJSS/>) に掲載しています。

また、この分科会では、日本統計学会会員だけでなく、学会会員以外の小・中・高校等の先生方、社会人、行政職員の方など統計教育に関心のある方の入会を歓迎しております。

会員には、電子メールで統計教育に関する情報も適宜お送りしております。会費等も必要はありません。統計教育に関心のある皆様方の参加をお待ちしております。

入会のお申し込みは下記ウェブページへアクセスし、必要事項を入力し送付してください。

<https://statds.k-junshin.ac.jp/sesjss-entry>

ご不明な点につきましては下記事務局までお問い合わせください。

日本統計学会統計教育分科会事務局 (sesjss@stat.k-junshin.ac.jp)

統計教育実践研究 第 15 卷 編集委員会

編集委員長	竹内 光悦	(実践女子大学)
編集委員	末永 勝征	(鹿児島純心女子短期大学)
	藤井 良宜	(宮崎大学)
	渡辺美智子	(立正大学)

The Institute of Statistical Mathematics
Cooperative Research Report 466

Research on best practice in teaching statistics
Vol. 15

March, 2023

The Institute of Statistical Mathematics
10-3 Midori-cho, Tachikawa
Tokyo Japan 190-8562