

統計数理研究所共同研究リポート 457

統計教育実践研究

第 14 卷

2022 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 457

統計教育実践研究

第 14 卷

2022 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

当研究所では

Annals of the Institute of Statistical Mathematics

統計数理

Computer Science Monographs

Research Memorandum

統計数理研究所研究レポート

統計計算技術報告

統計数理研究所研究教育活動報告

統計数理研究所共同研究レポート

を発行している．共同研究レポートは，統計数理研究所共同研究およびそれに準ずる共同研究による報告書として刊行される．この共同研究レポートは，2021 年度統計数理研究所公募型共同利用（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（研究代表者：末永勝征，課題番号 2021-ISMCRP-5008）による研究成果報告書として作成されたものである．

統計教育実践研究 Vol. 14

目次

I. ICT を活用した教育方法と評価

1. ICT・オンライン教材を活用した公開講座における統計教育の試み
松本智恵子（福井大学） 1
2. 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
及川久遠（大和大学）・深澤弘美（東京医療保健大学） 4
3. ICT を活用した PPDAC サイクルの実践事例研究
姫野哲人（滋賀大学）・渡部靖司・福澤純治・新海孝則（愛媛県立松山南高等学校）
・服部裕一郎（高知大学） 6
4. 【特別講演】これからの社会実装のあり方について～DX と組織と人材育成～
小泉誠（慶應義塾大学 システムデザイン・マネジメント研究所、デジタルリテラシー協議会） 8

II. 2021 年度日本統計学会統計教育賞受賞受賞講演

1. データサイエンス教育における多様性
小野陽子（横浜市立大学、WiDS TOKYO@Yokohama City University） 15
2. 高等学校におけるデータサイエンス教育の現状と展望
林宏樹（兵庫県立姫路西高等学校、帝京大学大学院、
JDSSP 高等学校データサイエンス教育研究会事務局長） 17

III. オンラインによる統計教育の展開

1. データサイエンスを対象としたオンライン PBL の課題と展望
竹内光悦（実践女子大学） 22
2. オンライン FESTAT の2年間を通して見える探究発表会の可能性
石井裕基（香川県立観音寺第一高等学校） 25
3. バーチャル空間でのオンライン教育の動向や可能性について
葉袋友花里（oVice 株式会社、PR マネージャー） 27

IV. 統計教育の研究および授業実践事例 I

1. 中学校における批判的統計リテラシーを育成するための教材開発
細田幸希（筑波大学大学院人間総合科学研究科）
・大脇和志（筑波大学大学院人間総合科学研究科） 28
2. パスカルの三角形の回帰分析
高木和久（高知工業高等専門学校） 34
3. 「中高生・スポーツデータ解析コンペティション—2021—」
ポスター作品審査報告および受賞作品の紹介
竹内光悦（実践女子大学、中高生・スポーツデータ解析コンペティション事務局） 38

V. データサイエンス人材育成と質保証

1. データサイエンティスト検定TM リテラシーレベル
—「データの力を解き放つ」基礎を問う試験の目的と今後の展望について—
菅由紀子（データサイエンティスト協会スキル定義委員、株式会社 Rejoui 代表取締役、
広島大学客員教授） 40

2. 人材育成における JDLA の取り組み	
岡田隆太朗（日本ディープラーニング協会理事／事務局長）	41
3. 情報処理学会におけるデータサイエンス教育の取り組み	
萩谷昌己（情報処理学会データサイエンス戦略推進委員会委員長、東京大学）	46
4. 求められる問題解決力のための知識を測る	
平岡靖敏（日本規格協会 品質管理検定センター所長）	50
5. データサイエンス数学ストラテジストのご紹介	
瀬良智也（日本数学検定協会）	52
6. データサイエンスの能力とその測定方法	
～統計検定データサイエンス CBT 発展級・エキスパート級の設計を通じて	
佐藤彰洋（横浜市立大学）・竹村彰通（滋賀大学）	58
7. CBT によるデータ対話型アナリティクススキルの測定と評価	
～統計検定データサイエンス基礎試験における実装～	
渡辺美智子（立正大学）	66

VI. 統計データ分析コンペティションを用いた統計教育

1. 統計データ分析コンペティション 2021 の概要とデータ分析実践事例	
堀川顕子（統計センター）・太佐美結（フェリス学院高等学校）	72
2. 統計データ分析コンペティション 2021 に参加して	
橋本三嗣（広島大学附属高等学校）・村澤舞・山家里穂（広島大学附属高等学校）	74
3. 統計活用奨励賞及び特別賞の授業実践事例（統計データ分析コンペティション 2021）	
馬場国博（慶應義塾湘南藤沢高等部）	
・谷優輝・河又杏香・遠藤沙恵（慶應義塾湘南藤沢高等部）	76

VII. 統計教育の研究および授業実践事例 II

1. 家庭や地域の実態調査を取り入れた統計的問題解決の実践について	
ー総合的な学習の時間や SDGs を関連させた取組ー	
加藤孝児（愛知県豊田市立寿恵野小学校）	78
2. 問題解決までの見通しをもち、適切なグラフや分析結果を問題解決に活用しようとする子の育成	
佃匠悟（童子山小学校）	89
3. 数学的モデル化を遂行する力を評価するフレームについての研究	
ー中学校第 1 学年「データの活用」領域における授業実践ー	
三上泰生（福井大学教育学部附属義務教育学校）・松本智恵子（福井大学）	94
4. 課題研究で使えるデータサイエンスの授業実践	
都丸希和・佐藤健太（名古屋大学教育学部附属中・高等学校）	100
5. 高校生のためのデータ駆動型の授業デザインと実践	
ー地域の課題から気づきが生まれるデータサイエンス教育ー	
松宮敬広（滋賀県立虎姫高校）・和泉志津恵（滋賀大学）・畑山満則（京都大学）	102
6. 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ー仮説検定授業の実践を踏まえてー	
ー地域の課題から気づきが生まれるデータサイエンス教育ー	
稲葉芳成（立命館宇治中学校・高等学校）	108
7. 神戸大学のデータサイエンス分野における高大連携の取り組み	
林兵馬（神戸大学附属中等教育学校）	
・齋藤政彦（神戸大学数理・データサイエンスセンター長）	110

VIII. 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	111
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	112

ICT・オンライン教材を活用した公開講座における統計教育の試み

松本智恵子（福井大学 教育・人文社会系部門教員養成領域）

連絡先：c-matumo@u-fukui.ac.jp

1. はじめに

本報告は、福井大学で行われている公開講座「体験ふむふむ数学クラブ」にて取り組んだ、統計・データサイエンスに関する実践「データサイエンティストになろう！」についての実践報告である。公開講座の内容と公開講座当日の様子について報告した後、ICT機器やオンライン教材を活用した統計・データサイエンス教育における課題について考察する。

教育分野においてPCやICT機器類を活用することに関しては、学習指導要領や「AI戦略2019」においても推奨されている。統計・データ分析にはPCが不可欠であるため、ICT機器が学校現場に浸透することは統計教育にとっては喜ばしいことである。しかしながら、ICT機器を利用した授業を展開する際には、ICT機器を効果的に用いて子どもたちの思考力・判断力・表現力を伸ばす教材開発やカリキュラム開発などが必要である。本実践事例における考察を通して、ICT機器を活用した統計・データサイエンス教育を行う際に注意すべき点や効果的な指導方法について考察したい。

2. 体験ふむふむ数学クラブ

「体験ふむふむ数学クラブ」は、福井大学教育学部数学教室が平成19（2007）年度より実施している公開講座である。一般の社会人や小中高校生（小学校第4学年以上）を対象に、体験的なグループ学習によって数学を楽しんで学び、参加者の数学に対する興味・関心を高めるとともに、数学が現代社会の様々な場面において果たしている役割について理解してもらうことを目標としている。平成26（2014）年度からは、福井大学の教員と学生・大学院生が共同で運営を行う形で実施しており、上記の目標を実現するカリキュラムを学生・大学院生が作成・実践する場ともなっている（西村他，2018）。折り紙やパズルを用いた図形学習・幾何学に関する内容を取り扱うことが多いが、確率・統計に関する内容を扱ったこともある（松本他，2012）。参加者は、小学生とその付き添いの成人という組み合わせが多い。

3. 「データサイエンティストになろう！」実践内容

今回の公開講座「データサイエンティストになろう！」は、令和3（2021）年11月6日に福井大学にて実施した。本実践内容の計画、テキスト作成、講座

の進行は松本が中心となって務めており、また、当日は学部生や大学院生、大学教員の合計6名がファシリテーターとして受講生の活動を支援した。講習時間は途中1回の休憩を入れて3時間である。

本公開講座は、福井大学にあるPC室の一つである「情報処理演習室」で開催している。本PC室は、PCが5～6台並ぶ固定された長机が並ぶ構成となっており、「体験ふむふむ数学クラブ」が掲げるグループ学習を行いにくい環境となっているが、長机1つをグループとし、グループ一つにファシリテーターが一人張り付く形で各班の支援を行った。

配布テキストによる本講座の内容は次の通りである。

- (1) 「画像認識」をやってみよう！
- (2) 「データ」を探してみよう！
- (3) 「データ」を「グラフ」にしてみよう！

本公開講座の目標である「数学、特に統計・データサイエンス分野に対する興味・関心を高めるとともに、現代社会における統計・データサイエンス分野の広がりについて理解」できる講座となるよう、進行役とファシリテーターとが協議を重ねた内容となっている。

3.1 「画像認識」をやってみよう！

公開講座の題名に掲げている「データサイエンティスト」の定義は多岐に渡っている。また、データを分析し、問題解決を図るためには、「データ」を得、得たデータを整理して「特徴」を掴む作業が必要不可欠であるが、上記の作業は非常に地味であり、小学生も多く受講している本公開講座では受講生の興味を引くことができない。それゆえ、本公開講座では、現代社会におけるデータサイエンスの典型的な分析手法であり、受講生にデータサイエンスについて興味を持たせるための導入として、まず、Googleが提供する、誰でも短時間で簡単に機械学習モデルを作成できるウェブベースの画像認識ツール「Teachable Machine (<https://teachablemachine.withgoogle.com/>)」を用いて「画像認識」を体験してもらった。「Teachable Machine」はWebブラウザであるGoogle Chromeと訓練データ用の画像さえあれば誰でも簡単に画像を分類する深層学習モデルを作成することができるため、「機械学習（深層学習）とはこういうものなのか」ということを手取り早く体験でき、受講生も、講師側が用意した画像データを用いて簡単に「画像認識とは何か」を理解できていたように思われる。

「Teachable Machine」で画像認識を体験してもらった後、「画像」とはどのような「データ」なのか、「画像」に写っているものが何であるのかを判断する際、「Teachable Machine」はどこを「見て」判断しているのか、「画像」に写っているものについて判断するために何が必要なのか、その仕組み（深層学習）について簡単に説明した後、「画像認識」の精度を上げるにはどうすれば良いのかについて考えてもらってから、「データ」が大切だということについて説明を行った。

3.2 「データ」を探してみよう！

「データサイエンティスト」は、「データ」から問題を解決することができる専門家である。しかしながら、データを分析するためには良いデータがなければならない。公開講座の中盤では、「データ」が大切であるという認識のもと、データを用いて問題解決を図るにはどのようなデータが必要なのかについて考えてもらった後、「データ」の例として、e-StatやRESAS、気象庁のサイトにあるデータをダウンロードしてもらい、「データ」とはどのようなものなのかについて理解を深めてもらった。

問題解決を図ることができる良いデータを得るためには、どのような問題を解決したいのかという「目的」を持つことが必要だが、「目的」部分については受講生に一任したため、データを探したりダウンロードしたりする活動については難しさを感じている受講生もいた。例えば、e-Statでは、「キーワード検索」欄にキーワードを入力することで目的に応じたデータを探し出すことができるが、「目的」を上手に言語化できなければ検索機能を利用することは難しい。この「言語化」部分については、もう少し時間を取って説明した方が良かったのではないかと考える。

3.3 「データ」を「グラフ」にしてみよう！

「画像認識」について理解するためには大学数学の知識（線形代数や微分・積分）が必要であるが、「データの整理」と「統計図表の作成」だけでも、データを説明することは可能である。このことを説明するために、3.2節でダウンロードしたデータを用いてグラフ作成を行い、作成したグラフから何が分かるかについて考察を行った。グラフ自体はExcelの操作ができれば簡単に作成できるが、グラフからの考察を行う際には、目的に応じた適切なグラフを描くことが必要であり、また、「比較」を行うことで考察を行うことが簡単になる。木村(2021)が作成した「福井県のデータを利用した統計授業用資料」をもとに以上のことを説明した上で、あらかじめ講師が用意していたデータや、3.2節でダウンロードしたデータを用いてExcelでグラフを作成し、できあがったグラフにグラフから分かるこ

とを書き添えて、大学のPCにインストールされている授業補助アプリを用いてミニ発表会を行った。総務省が作成した「統計表における機械判読可能なデータ作成に関する表記方法 (https://www.soumu.go.jp/main_content/000723626.pdf)」に則ったデータ表が作成されつつあることもあり、ダウンロードしたデータからグラフを作成する作業は意外にすんなりとできていたように思う。棒グラフを作成した受講生が多かったが、2021年7月の福井市の日別降水量のデータを用いて道路が冠水した日の降水量がダントツに多かったことを示す棒グラフを作成するなど、良い考察ができるグラフを作成することができていた。しかしながら、特に気象庁のサイトから気象データをダウンロードしてグラフを作成した参加者の中には、ダウンロードしたデータが「csv形式」のデータであることに気付かなかった参加者がおり、グラフ作成後の保存の際にグラフが消えてしまうという残念なことも起こってしまった。csv形式のデータはExcelで普通に開くことが可能であるため、データの形式に関しては、説明を更に厳格に行うと同時にファシリテーターにも目を配ってもらう必要があるだろう。

4. 考察と課題

4.1 公開講座アンケートより

Q8. この公開講座の内容はいかがでしたか？

12件の回答

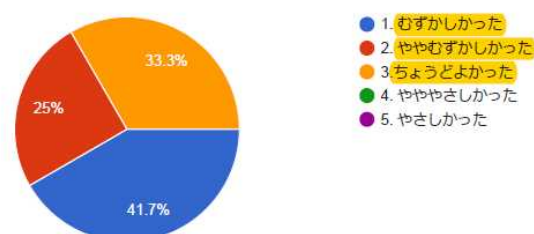


図1 公開講座の難易度

Q10. あなたは、この公開講座にどの程度満足されましたか？

12件の回答

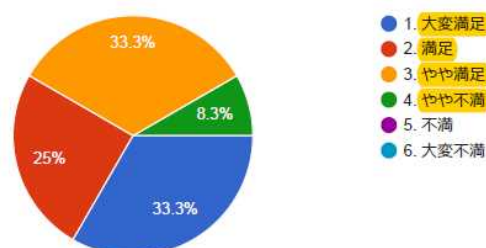


図2 公開講座の満足度

福井大学における公開講座に関する事務を担当している福井大学地域連携推進課地域連携担当が公開講座後に実施したアンケート結果を以下に示す（回答者数

12名／参加者17名)。アンケート回答者の内訳を見ると、10歳以下と10代前半が2/3を占めており、普段の「体験ふむふむ数学クラブ」の時と同様に、小学生とその付き添いの成人の方が殆どであることが分かる。

講座の難易度は「むずかしかった」が41.7%、「ちょうどよかった」が33.3%となっている(図1)。「むずかしかった」を選んだ理由について、「PCの使い方が難しかった」、「始めてパソコンでグラフを作ったから」という意見が挙がっており、また「ちょうどよかった」を選んだ理由として「エクセルやwebの使いかたを大体知っていたから」という意見が挙がっていたため、PCの習熟度によって公開講座の難易度に差が付いていたように思われる。また、満足度については、33.3%が「大変満足」、25%が「満足」と回答しており、難しめではあるが、AIやデータサイエンスに興味を持たせることができた公開講座であったと考える(図2)。

4.2 講師側からの考察と課題

今回、講師として携わって感じたことは二つある。一つ目は、理論的な部分は難解である「機械学習(深層学習)」について、参加者は意外と楽しく作業を行っていたことに驚きを覚えたことである。今回、理論的な部分については、「数理・データサイエンス教育強化拠点コンソーシアム(<http://www.mi.u-tokyo.ac.jp/consortium/>)」が提供している応用基礎レベルのeラーニング教材を利用し、画像を多用した説明を行ったが、「Teachable Machine」を用いた画像認識については、もう少し長めに時間を取って作業を行った方が良かったのではないかと思うほど、参加者は熱心に取り組んでいた。この、理論は難解だが結果は直感的で分かりやすい部分が、深層学習による画像認識が現代社会に浸透している理由の一つなのではないかと思う。

二つ目は、受講生によってPCの操作の習熟度にばらつきがあり、指導に難しさを感じたことである。これは、同じ環境のPCを用いて多変量解析を教えている、高校生以上が対象となっている福井大学公開講座「統計・データサイエンスへの誘い」(の前身である医療統計講習会)では感じなかったことである。AI戦略2019やGIGAスクール構想により、学校現場では一人一台のPCの導入が進んでいる。しかしながら、今回用いた福井大学のPCはマウスとキーボード(ローマ字入力)で操作するものであり、学校現場で用いられているタッチパネルやタッチペンで操作するPCとは操作環境が異なっていたため、小学生達は戸惑いを覚えたのではないかと考える。タッチパネル形式は入力方法としては直感的で手軽ではあるが、データサイエンスやプログラミングを本格的に行うためにはキー

ボードの方が便利な部分もあるため、バランス良く指導することが大切であろう。ただし、「体験ふむふむ数学クラブ」の場合、殆どの小学生参加者には保護者が付き添っており、マウスとキーボード操作に慣れている大人が、小学生が出したアイデアをPC上で操作することでデータを集めたりグラフを完成させたりする場面が数多く見られた。実際の授業においても、グループ学習を上手に利用し、PC操作については子どもたちがルールに則って互いに教え合うことにより、PC操作に関する習熟度のばらつきは解消できるのではないと思われる。

今後の課題としては、データサイエンス・統計をより身近に感じることができるよう、PCの操作や習熟度を考慮したカリキュラムを作成し、データサイエンス・統計の面白さや現代社会との繋がりについて小学生にも理解できるようにすることが挙げられる。また、ICT機器とグループ学習を取り入れた統計・データサイエンスのカリキュラムを作成できる教員を育てるために何が必要なのかについても更なる考察を行いたい。

引用文献

- 木村彰真(2021), 福井県のデータを利用した「データの活用」に関する教材開発, 福井大学教育学部卒業論文。
- 西村保三, 大石美咲, 田島祥大, 松本智恵子, 間庭彰郎, 河端尚輝(2018), 折り紙を用いた図形学習の実践事例—H29 体験ふむふむ数学クラブ「折り紙で図形を作る」の実践報告—, 福井大学教育実践研究 43, pp. 63-68.
- 松本智恵子, 大久保裕介, 佐分利豊, 竹澤康宏, 坪川武弘, 西村保三, 福田浩之, 松田立行, 山下敏明(2012), 公開講座における問題解決型統計学習の実践〜「体験ふむふむ数学クラブ」の取り組みより〜, 統計教育実践研究 4, 126-129

謝辞

本研究は、科研費21K02600の助成を受けたものである。

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

発表者：大和大学 及川久遠

共同研究者：東京医療保健大学 深澤弘美

連絡先：oikawa.hisao@yamato-univ.jp
h-fukasawa@thcu.ac.jp

1. はじめに

新型コロナ感染拡大により、教育現場においてもDX化が喫緊の課題となり、オンラン授業が多く実施されるようになった。動画配信型等のオンライン授業は、繰り返しの学習が可能であり、知識や操作の習得に効果的な側面も認められ、教育現場では効果的な教育の方法としてコンピュータを活用する場面がさらに多くなっている。しかしながら、コンピュータを活用した試験については、不正の防止など課題が多く、試験問題の準備及びその実施は簡単ではない。中間試験や期末試験など成績に直結する試験は、対面での実施を余儀なくされることも多い。オンライン上で実施する試験は、これまでの紙ベースの試験をそのままコンピュータ上に置き換えられないことも多く、コンピュータの特性を生かし、より適切な評価方法を確立することが必要である。

近年、統計検定をはじめ統計学教育の評価がコンピュータ上で実施される機会も少なくない。そこで、本研究では、特に統計リテラシー教育に焦点を当て、その評価をコンピュータ上で実施する際の留意点を整理し、コンピュータ適応型テスト実装の可能性を検討したい。

2. コンピュータ上で行う試験の現状と課題

(1) 学習支援システム(LMS)を活用した試験

WebClass等の学習支援システムでは、オンライン上で試験を実施するための機能が豊富に用意されている。上図に示すように、複数選択肢の中からひとつの答えを選ぶ「単数選択式」、複数選択肢の中から複数の答えを選ぶ「複数選択式」、語句や数値を答える「単語数値入力」、自由記述で回答する「記述式」な

どの設問を設定できる。設問をあらかじめ多めに用意しておいて、その中からランダムに出題することや、設問や選択肢をランダムに並び替えて不正行為を防止するための指定をすることもできる。しかしながら、オンライン上で他者から解答を得るなどの機会を完全に封鎖することはできず、不正を完全に防止することはできない。

Figure 1 displays three examples of question formats available in WebClass:

- 単数選択式 (Single Choice):** A question "今日の天気は?" (What is the weather today?) with five radio button options: 1. 雨 (Rain), 2. 曇り (Cloudy), 3. 晴れ (Sunny) - selected, 4. 雪 (Snow), and 5. その他 (Other).
- 複数選択式 (Multiple Choice):** A question "世界のGDPのランキングで上位2ヶ国を選びなさい" (Select the top 2 countries by world GDP ranking) with five checkbox options: 1. イギリス (UK), 2. 日本 (Japan), 3. アメリカ (USA) - checked, 4. ドイツ (Germany), and 5. 中国 (China) - checked.
- 単語/数値入力 (Word/Number Input):** A question about the first Prime Minister of Japan. It has two input fields: (1) with the text "伊藤博文" (Ito Hirobumi) and (2) with the number "44".

図1 WebClassの出題形式

(資料：<https://www.datapacific.co.jp/>)

(2) Zoom等によるビデオを活用した試験

ZoomMeetingにおいてビデオカメラを活用することにより、本人確認および試験中の挙動を教員が観察しながら試験を行うことができる。2020年新型コロナウイルス感染症拡大初期においては、カメラを2台接続して顔や手元を確認できるようにするなどの工夫をして試験を行った学校も多かったが、それでも不正を完全に防ぐことは難しく、評価方法を変更するなど余儀なくされている。

(3) 試験会場に常設された PC を用いた試験

統計検定等の資格試験では、試験会場を設置し、常設の PC を使って試験をうける体制を提供している。統計検定は自宅からのオンラインで受験することができないため、会場にて本人確認が可能であり、PC も試験用の設定になっているため不正行為を防止できる。

近年では、オンラインでの受験を検討とする資格試験もあり、オンライン試験を監視するシステムの開発も進んでいる。また、試験問題そのものを工夫して、これら従来の問題点を克服できる試験方法としてコンピュータ適合型試験がある。次章以降に概要をまとめる。

3. コンピュータ適合型テスト

コンピュータ適合型テストとは、項目反応理論を用いて受験者一人一人の能力に合わせた問題を提示する試験である。一般的な紙ベースのテストは、公平性を担保するために同時刻に同一問題を出題することを前提に設計されているため、受験者の能力によっては解けるはずのない難しい問題や易しすぎる問題もすべて回答する必要があった。しかし、情報技術の発展によりテストの形態は多様化し、好きなときに好きな場所でコンピュータを用いて受験することも可能となった。さらには、コンピュータ適合型テストを導入することにより、少ない問題で評価が可能であること、試験時間を短縮できること、不正行為を防止できることなどもメリットとしてあげられる。

4. 項目反応理論

紙ベースの試験では、通常 0 から 100 などの点数で評価するが、試験問題の難易度が異なれば、例えば 80 点という評価点が常に同じ能力を示しているとは言えない。物差しや体重計で長さや重さを測るときは、通常誰が図っても同じ長さや重さを示すが、試験では出題される問題が異なる場合の点数では、それが問題の難易度と受験生の能力のどちらに起因しているか判断できない。従来の試験では、項目困難度や項目識別力といった項目に関する指標が、受験者集団に依存して相対的に決定され（標本依存性）、試験の点数という受験者の能力に関する情報が試験に含

まれる問題（項目）に依存する（項目依存性）。この問題を解決するには、試験問題（テストに含まれる項目）のむづかしさ（項目困難度）と受験者の能力を別々に推定し、同じものさし上で評価できる項目反応理論が必要となる。項目反応理論は、近年コンピュータの性能が向上したことに伴い、欧米を中心に標準的な方法として広く利用されるようになった。TOEFL、情報処理技術者試験の一つである「IT パスポート試験」には項目反応理論が導入されている。

項目反応理論においては、試験を構成する個々の問題について、項目困難度や項目識別力といった問題の特徴を表すパラメータの値がすでに明らかになっていることと、異なる試験について、それらの得点を比較できるようにするために等化することが必要である。項目反応理論のメリットは、受験者のレベルにあった問題を出題できることである。コンピュータを活用することにより、受験者の解答パターンをみながら、その受験者に最適な問題（受験者の能力レベルにあった難易度と識別力を持つ問題）を出題できる。個に応じたテストを提供でき、紙ベースの試験より少ない問題で正確な能力測定が可能である。試験時間も短縮できる。

異なるテストの得点を比較可能にするためには、等化が必要で、データが IRT モデルに完全に適合していれば、受験者集団によらない難易度や識別力を得られる。等化により標本依存性を克服できる。

5. まとめ

統計基礎リテラシーを評価するうえでは、知識を問う問題以外に統計的思考力を問う問題が必須となる。このような問題に対して、IRT をいかに活用し評価を行うかについては、今後さらに検討し、出題分野ごとの難易度・識別力の設定や知識問題のみ IRT を適用してリテラシーレベルを評価したうえでそのレベルにあった思考力問題を提示するなど具体的な評価方法を確立していきたい。

【参考文献】

佐藤貴彦、「項目反応理論における予測分布を用いたデータ解析」、東京理科大学修士論文

本研究は統計数理研究所の重点型研究（テーマ 2「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」）として助成を受けています。

ICTを活用したPPDACサイクルの実践事例研究

姫野哲人* (滋賀大学データサイエンス学部)

渡部靖司 (愛媛県立松山南高等学校)

福澤純治 (愛媛県立松山南高等学校)

新海孝則 (愛媛県立松山南高等学校)

服部裕一郎 (高知大学教育学部)

*連絡先 〒522-8522 滋賀県彦根市馬場 1 丁目 1-1

E-mail : tetsuto-himeno@biwako.shiga-u.ac.jp

1. Introduction

この2年間、新型コロナウイルス感染拡大の影響で、教育現場でオンライン教育の環境が急速に整ってきた。大学生のパソコン所有率は2019年には59%だったが、2020年は80%、2021年は91%と増えてきた(株式会社カティサーク、【2021年度版】大学生のSNS利用率・PCスマホ普及率アンケート調査結果)。数年前まで文系の学生はそれほどパソコンを所持していなかったが、今では学部問わずパソコンを所持するようになってきている。

一方、高校でのパソコン普及率は各種調査で2021年段階でも50~60%程度となっているが、ICTの活用の必要性を認識している高校は全体の9割を超える状況となっている(旺文社、【2021年度版】全国の高等学校におけるICT活用実態調査)。

これらのことから、高校教育、大学教育において、ICTを活用した教育は必要不可欠であることが分かる。本発表では、大学教育、教員養成課程、高校教育の3つの視点から、ICTの活用とPPDACサイクルの活用について紹介する。

2. 大学教育でのICTの活用

大学における数理・データサイエンス・AI教育の普及に向けて、文部科学省が教育プログラム(リテラシーレベル・応用基礎レベル)の認定制度を実施しており、これらの中で「実データ、実課題を用いた演習」が必要となっている。この認定制度への申請にあたり、多くの大学でPPDACサイクルの実践体制を整えることが急務となっている。

滋賀大学データサイエンス学部では、1年後期、2年後期に、それまで学んできた知識を生かして課

題を解決するPBL型授業があり、3、4年次ではゼミにおいてPPDACサイクルにより様々な課題解決を実践している。これらは、学生自らデータを収集して課題解決を行うケース、各種コンペティションへ参加するケース、購入データや企業から提供されたデータを基に課題解決を行うケースなど様々な形があり、例えばコンペティション参加のケースでは、2022年度は統計データ解析コンペティション、WiDS Hiroshimaのアイデアソンなどでの入賞を果たしている。

3. 教員養成課程でのICTの活用

新学習指導要領が小学校では2020年度、中学校では2021年度より全面実施、高等学校では2022年度から順次実施となり、今回の改訂では「知識・技能」、「思考力・判断力・表現力等」、「学びに向かう力、人間性など」の資質・能力育成に注力されており、統計やプログラミング教育が重要視されている。

これらの改訂により、教員にもICTの活用がより求められるようになってきている。高知大学教育学部の取り組みとしては、子供たちの統計的問題解決能力を育む授業力向上のため、ICT授業においてSimpleHistやSimpleBox(宮崎大学 藤井良宣教授による開発)を用いた教材開発を行ったり、Geogebra(<https://www.geogebra.org/?lang=ja>)を用いて生徒が主体的に数学的性質を探究する教育指導を行っている。

4. 高校教育でのICTの活用

愛媛県立松山南高校ではICTを積極的に活用し

た PPDAC サイクルの実践を数多く行っている。以下では、これらの一部を紹介する。

(1) 教科「情報」における購買データの活用

新学習指導要領「情報Ⅰ」(データの活用)では、「データの収集、整理、分析及び結果の表現の方法を適切に選択し、実行し、評価し改善すること」と述べられている。課題研究「データサイエンス」への接続を円滑に行うため、教科「情報」で生徒に身近な購買データを用いた授業を実施している。それにより、「データの活用」を身近に感じさせ、教科「情報」と課題研究「データサイエンス」との連携で相乗的効果が期待できる。

教科「情報」における購買データの活用として、True Data が提供する ID-POS データである、ウレコンや Dolphin Eye を用いて購買動向を分析し、可視化した結果から考察する課題研究を行った(図1)。



図1 「情報」の授業の様子

(2) 課題研究の取組

スーパーサイエンスハイスクール第5期において、普通科では課題研究を学校設定科目「データサイエンス(DS)」で、1年生1単位、2年生1単位、3年生1単位で行っている。1年次(DSⅠ)は、その後、RESAS(地域経済分析システム)の使い方とテキスト課題探究メソッド(啓林館)を用いて課題研究について学び、全国大会のコンテストで入賞した先輩方の作品を参考にして、データを利活用した課題研究について学ぶ。2学期以降は、「愛媛の未来をデータで考える」をテーマに、各班独自の切り口で研究を行い、10月の中間発表会でプレゼンテーション発表を行う。3学期は、課題研

究のまとめとして、ポスターを制作し、A4用紙で4枚程度の論文にまとめる。2年次(DSⅡ)では、e-Stat(政府統計ポータルサイト)や自治体などのオープンデータなどを用いて、各々が興味を持ったことをテーマに本格的な課題研究を行い、プレゼンテーション発表やポスター発表を行う。また、3年次(DSⅢ)は2年次に行った課題研究をA4用紙で8枚程度の論文にまとめる。また、各種コンテストへの応募や発表に関わる活動を行う。

・データサイエンスⅠ(DSⅠ)の取組

第5期の指定を受けた1年生から普通科全体で行う課題研究を学校設定科目「データサイエンス(DS)」で実施している。1年生は「データサイエンスⅠ(DSⅠ)」で、プレ課題研究を行っている(図2)。



図2 「データサイエンスⅠ」の授業の様子

・データサイエンスⅡ(DSⅡ)の取組

2年生は「データサイエンスⅡ(DSⅡ)」では、6つの分野(人文科学系、社会科学系、保健・福祉・スポーツ系、自然科学系、情報・統計系、データマーケティング)に分かれて課題研究を行っている(図3)。



図3 「データサイエンスⅡ」の授業の様子



デジタルリテラシー協議会 事務局
慶應義塾大学 システムデザイン・マネジメント研究所 研究員
日本ディープラーニング協会 プロジェクトアドバイザー

<略歴>
慶應義塾大学卒業後、株式会社リクルートに入社。Eコマース・アドテクノロジー・業務支援SaaS・キャッシュレス決済等、随々の新領域で経営企画・事業開発・戦略設計を行い、20を超えるサービスを担当。同社、Airレジの事業企画責任者を経たのち、経済産業省へ入省。商務情報政策局 情報経済課にて産業横断でのデータ利用とAIの社会実装推進を担当し、AI戦略、デジタル人材育成、スタートアップ支援、新たなデジタル市場の取引基盤、モバイル・スマートシティ、デジタルID、デジタルアーキテクチャ・デザインセンターの立ち上げ等、省庁横断でのデジタル産業政策を推進。
2021年に退職後、産学官を横断して活動中。

デジタルリテラシー協議会の活動について

Di-Lite 3



デジタルリテラシー協議会とは？

Di-Lite 4

「日本のデジタル人材育成を加速させる」
ための宜民連携の会議体です。

協議委員メンバー (50名順)		参加団体 (50名順)	
	安宅和人 慶應義塾大学 環境情報学部教授 データサイエンス CSO データサイエンスアドバイザー ト協会代表理事		
	草野 隆史 データサイエンス ト協会代表理事		
	富田 達夫 独立行政法人情報機構 推進機構 理事長		
	西山 圭太 東京大学 未来ビジョン研究センター 専任教授 日本ディープラーニング協会 専任理事		
	松尾 豊 東京大学 大学院工学系研究科 教授 日本ディープラーニング協会 専任理事		

＜デジタル＞

+

＜リテラシー＞

- ・ リテラシーとは、便利なもの・役立つものに「アクセスし、活用する能力」のことである。
- ・ それには、対象がどんな性質、これまでにない特性を持っているかを理解していて、どのようなことに有効かを知っていることが大事。

例：ヘルスリテラシー、金融リテラシー

＝

＜デジタル リテラシー＞
デジタル技術にアクセスし
目的のために使う能力

「全員」に、「全体」を。

「全員」に、「全体」を。

Di-Lite 7

▼「全員」とはなにか？

- ・ DX時代のビジネスは「デジタルを使う」「デジタルを作る」両輪の人材がいることで推進する

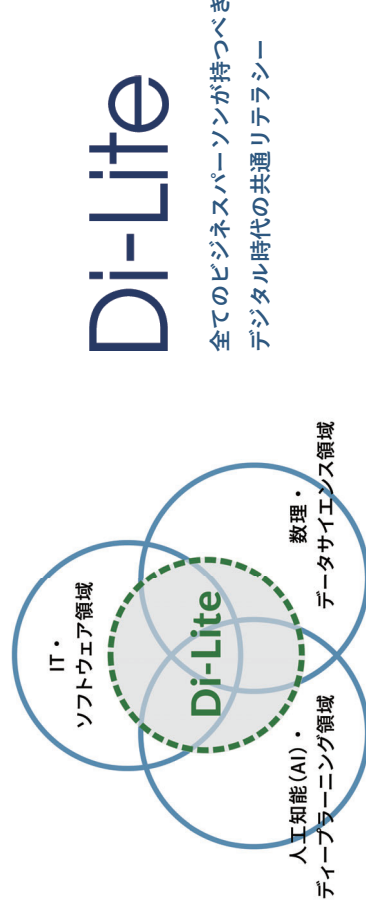


「全員」に、「全体」を。

Di-Lite 8

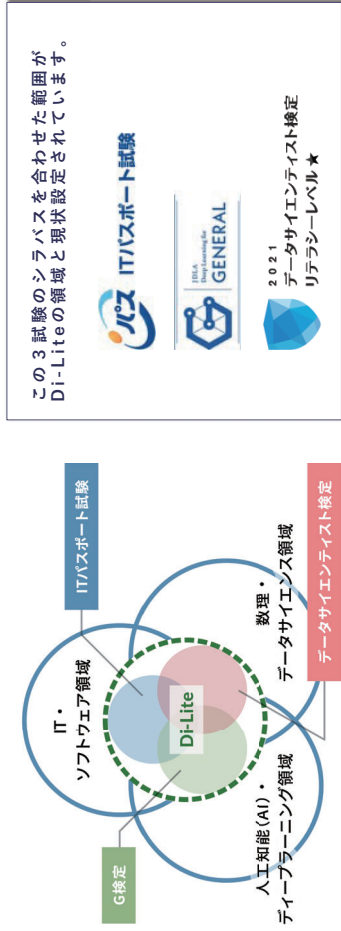
▼「全体」とはなにか？

- ・ おおよその「全体」＝IT・ソフトウェア×数理・データサイエンス×AI・ディープラーニング



▼「全体」とはなにか？

「Di-Lite」習得のためのCertificate Program



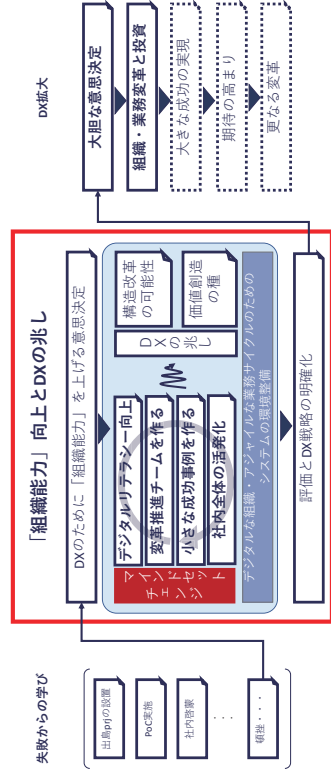
今後この3領域に留まらず、重要と考えられる要素(知識・スキル・マインド)を順次追加を検討

© 2021 Digital Literacy Council

DX推進を通して考えてみる

DX推進において重要な、現状議論との差分

- 失敗からの学びを元に、「組織能力」そのものを向上させるプロセスを、徐々にデジタルへと組織の「マインドセットエンジン」を移行し、社内全体が活発化、DXの兆しを掴む。
- その後、大胆な意思決定を経て、組織・業務改革と大規模投資を実行。



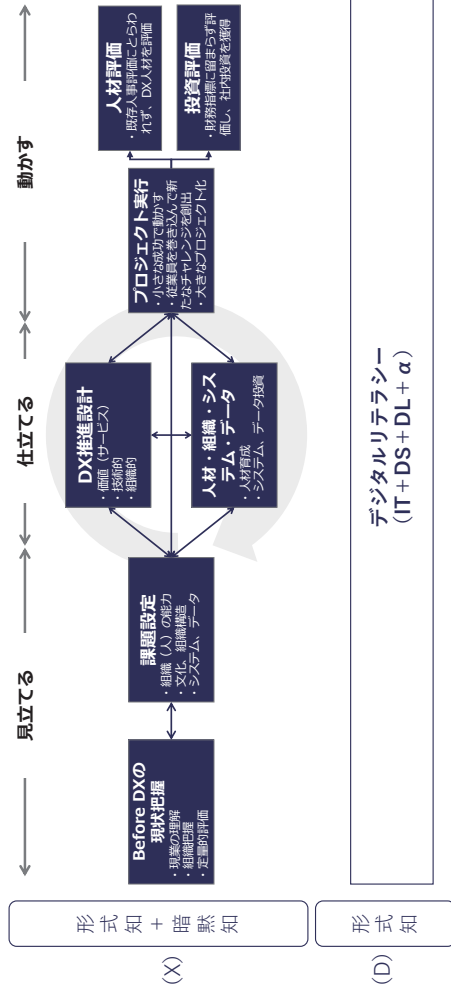
2021年8月～10月にDLIAが実施したインタビューより

© 2021 Digital Literacy Council

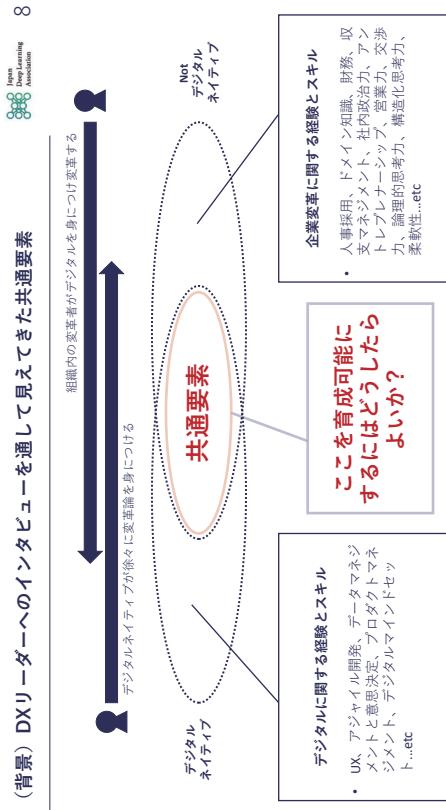
人材育成＝目的達成のために、組織が人を対象として行う活動（≡投資）

デジタル (D) に加えてDX推進に必要となる「変革スキル (X)」

- この要素は、全社を動かす“大きな動き”の場合も、部門やサービス単位で動かす“小さな動き”の場合でも同様が必要。
- それぞれの要素をスペシャリストレベルに身につけるのではなく、この全体を捉えて各所各人と連携して実行できるスキルレベルが重要



【問い】 こういったものをどう育成するのか？



13

ここ数年私がやってみたこと

14

いつも言っていたこと

AI人材不足の背景にある、構造的な問題。
それを超えたい。



Copyright © 2020 AI Quest All rights reserved. AI Quest

7

16



AI Quest

From 2019

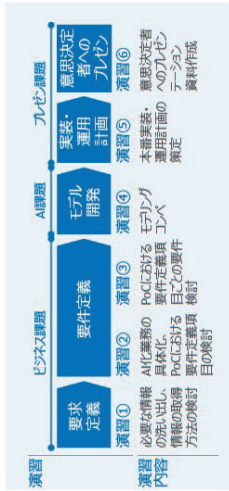
15

ということか？

人材不足の構造的課題

- AI自体が教育体系化できるほど成熟しておらず、且つ変化のスピードが速い。
- 必然的に前線でAIを推し進めている人材を講師にしないと実践的な教育が困難。
- 講師の確保がままならず、講師不足が起こる。

構造的な問題を、超える



教材作成

架空の企業へのAI導入プロジェクトを一気通貫で疑似体験できるよう設計された独自教材



“場”の構築

受講生ポータル画面



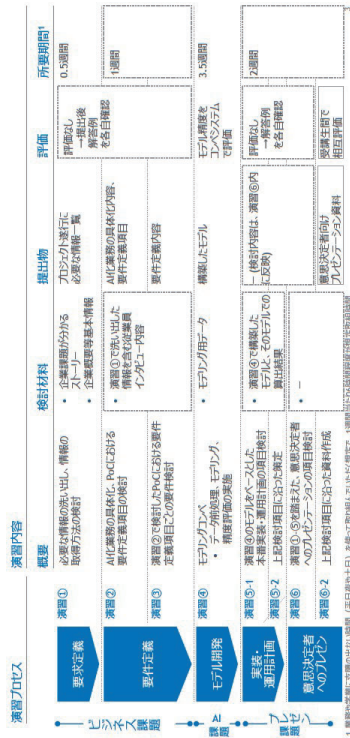
テレビ会議システム*



コミュニケーションツール



1ヶースタディーあたり約1.5ヶ月
1週あたり約6時間目安のプログラム



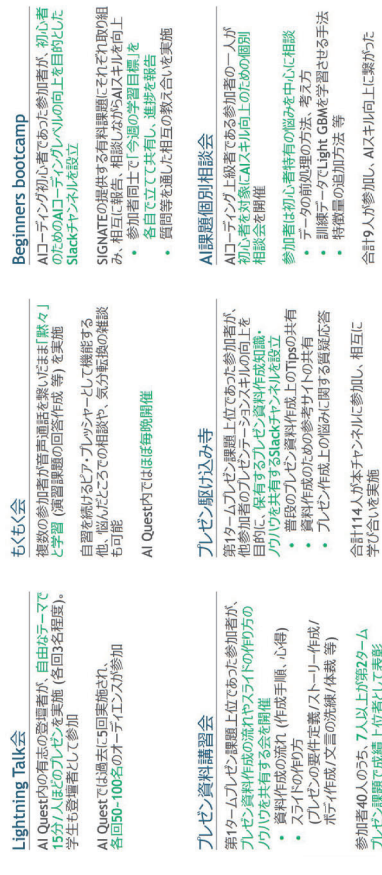
プロフェッショナルの提言



基礎づくりの基礎



□ ニテイの自発的な動北



3年間で約2000名が参加

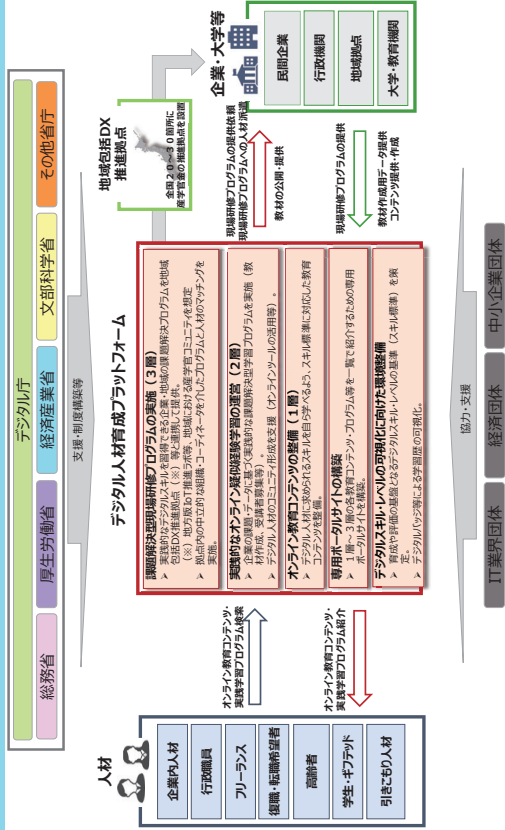


関係資料

- デジタルリテラシー協議会
<https://www.dilite.jp/>
- 経済産業省 デジタル時代の人材政策に関する検討会
https://www.meti.go.jp/shingikai/mono_info_service/digital_jinzai/index.html
- 経済産業省 実践的なAI人材育成について(AIQuest)
https://www.meti.go.jp/policy/it_policy/jinzai/AIQuest.html
- AIQuest公式ウェブサイト
<https://aiquest.meti.go.jp/>
- AIQuest 参加者インタビュー
https://www.meti.go.jp/policy/it_policy/jinzai/AIQuest_Alumni_PR.pdf
- AIQuest 教材サンプル
https://www.meti.go.jp/policy/it_policy/jinzai/AIQuestText_Sample.pdf

デジタル人材育成プラットフォーム 概要イメージ

- デジタル人材の不足に対応し、地域の企業・産業のDXを加速させることで、デジタル田園都市国家構想を実現するためのプラットフォームを構築する。



データサイエンス教育における多様性

小野 陽子・横浜市立大学
WiDS TOKYO@Yokohama City University
〒236-0027 横浜市金沢区瀬戸 2 2-2
Email : onoyk@yokohama-cu.ac.jp

1. はじめに

VUCA の時代、従前の社会的枠組み条件は大きく変動し、不確実で複雑、曖昧なものとなっている。こうした中、特定のディシプリンの限界性が明らかとなり、それを乗り越えるための新たな価値創造の前提として、多様性と包摂性の重要性が認識されてきている。データサイエンスに、こうした状況をデータから「見える化」し、新たな価値創造につなげるものとして大きな期待が寄せられていることに異議を唱えるひとは少ないだろう。その一方で、内閣府・年次経済財務報告[1]に見られるように、データサイエンス技法ではなく、従前の統計学的手法の適切な利用も、多様性の推進において重要な役割を担っている。データから統計的な推測を行うことで、また、これまでにその存在を意思決定の基盤となるデータの中に取り込まれてこなかった人々のデータを適切に取り入れる上で、データ収集・分析技術の進歩のみならず、多様な価値を持つ人々が集い、検討するという極めて基本的な行動が、変化の早い社会において必要とされている。本発表では、多様性を支えるデータサイエンス人材の育成活動の例を報告する。

2. 多様性への取り組み

Fleming[2]は、17000 件の特許を分析し、プロジェクトチームの専門領域の多様性と、特許によるイノベーションの創出した価値の大きさを分析した。この結果は、「価値のあるイノベーションには多様性に富んだチームが必要だ」として引用されることが多い。また、企業経営においても、リーダーシップと多様性の関係が研究されるなど[3]、多様性と価値創造の関係について議論がなされている。しかしながら、多様性という言葉の多義性から、一度定義をし、固定すれば永続的に変更されないというものではなく、分野、文脈により多様性の定義が異なることも見受けられる。多様性を一意に定めることが難しいのであれば、我々はどうすべきか。社会の変化をしなやかに受け止め、そのあるべき姿について「自分ごと」として取り組む上での指針を、常に多様な価値による討議を経て、アップデートすることができるかどうか問われている。

日本における多様性の取り組み例として、経済産業省による「新・ダイバーシティ経営企業 100 選」[4]が挙げられる。「新・ダイバーシティ経営企業 100 選」の前身である「ダイバーシティ経営企業 100

選」では、「多様な人材を活かし、その能力が最大限発揮できる機会を提供することで、イノベーションを生み出し、価値創造につなげている経営」として「ダイバーシティ経営」を捉え、競争優位を構築するための経営戦略と位置付けている[5]。ここでの多様な人材とは、性別、年齢、国籍、障がいの有無などだけでなく、キャリアや働き方などの多様性も含まれている。

教育における多様性への取り組みの例としては、文部科学省による「ダイバーシティ研究環境実現イニシアティブ 女性研究者研究活動支援事業」[6]が挙げられる。これは、「研究環境のダイバーシティ実現に関する目標・計画を掲げ、優れた取組を体系的・組織的に実施する大学や独立行政法人等を選定し、その取組を重点的に支援するもの」であり、女性研究者の活動を支援したものである。

一方、海外の大学教育における取り組みの現状はどのようなになっているだろうか。アメリカ大学協会のウェブサイトでは、DEI (Diversity, Equity and Inclusion: 多様性、公平性、包摂性)に関する大学としての声明をいかに記すか[7]など、多様性のみならず、包摂性、公平性について、大学がどのような姿勢で取り組むのかを明らかにしている。そこには、日本の多様性取り組みに多く見られるような、女性活躍は見受けられない。また、常にそれらの声明のアップデートが見て取れる(参考[8])。

日本の多様性への取り組みは、多様性(diversity)の類義語である variety (同じ種類の中の違い)に近く、diversity(そもそも種類が違うこと)に対処できていないように思われる。そのような日本において、データサイエンス人材教育を行う上で、第一に検討することは、この分野に多様な人材を取り込むことである。そしてその取り掛かりの一つは女性を中心とした活動を行うことではないだろうか。

3. 日本における WiDS (Women in Data Science) の広がり

Women in Data Science(WiDS)とは、2015 年に始まった米国スタンフォード大学を中心とした世界的なプロジェクトである。データサイエンス分野で活躍する女性を中心に、データサイエンスの意義、面白さ等を伝え、この分野に男女を問わず多くの人材をいざない(inspire)、教育の機会を提供し(educate)、また支援する(support)ことを目指している。スタンフォード大学で年 1 回開催されるシンポジウム(WiDS Global)には、毎回 1000 人以上が参

加、その動画はこれまで 10 万回以上再生されている。また、世界での地域イベントの開催は、60 か国、150 地域以上で行われている。

日本での WiDS は、3 人のアンバサダーを中心に産官学連携の下で行われている。2018 年度以来、横浜市立大学、広島、日本 IBM 株式会社が合計 8 回のシンポジウムを継続的に開催した。この他、教育啓蒙活動の一環として日本で開催したワークショップは計 19 回、アイディアソンは計 4 回行われている。この活動には、統計・データサイエンスの専門家のみならず、中高校生や専門分野外からの参加者も見られ、既に教育面で広がりのある活動となりつつある。

2018 年夏に Stanford の WiDS Global ディレクターへ、日本初となる開催希望を横浜市立大学から連絡した際に、ディレクターから開催目的を問われた。その際に、我々が開催する上での柱となる事柄として述べたことは次の 3 点である。

1) データサイエンス領域に参画する人材をいざなう上で、多様性・包摂性の観点から教育活動に貢献したいこと

2) 緩やかでしなやかな産官学連携の形を作りたいこと

3) 学会発表とは異なる、“Data Science for Social Good”という観点で社会をデザインすることを自由闊達に議論し発信していく場としたいこと

こうして始まった日本での取り組みは、WiDS Global において高く評価されている。横浜市立大学シンポジウムにおけるデザイナー・エマ理永氏が全世界での WiDS 登壇者約 4000 人の中から、Best of WiDS Speakers[9]としてアジア太平洋地域の素晴らしい発表 9 人の一人に選ばれたことは、その顕著な例と言えるであろう。

データサイエンスを一過性のお祭り騒ぎとせず、日本らしい多様で、しなやかな価値を生み出すために、これからも日本の WiDS を進める所存である。

4. 今後の課題

多様性とは、研究のみならず、気づきを与え、知識を伝える教育の現場においても重要な事柄である。自分の意見に囚われることなく、自分とは異なる価値観や理念を持つ他者の視点を、感情的にではなく知的に理解する力、すなわちエンパシーを育み、他者の視座から物事を捉えることの重要性を伝達することが必要だからである。

未来は更に多様化するが、分断されるのではなく、更なる繋がりを、実空間のみならず仮想空間までもに広げることになるであろう。そして、未来の問題は、国境や文化を越え、あらゆるニーズに対応することが求められることであろう。その際に、データサイエンスが果たす役割は、現在よりも更に深く多岐に渡ることは容易に想像できる。そして、データサイエンス研究・教育に関係する人々は皆、未来の

ために多様な視座を育むことが求められる。

謝辞

日本における WiDS の取り組み (WiDS HIROSHIMA 菅由紀子氏, WiDS TOKYO @ IBM 西戸京子氏, WiDS TOKYO @ Yokohama City University 著者) が、2021 日本統計学会統計教育賞という名誉ある賞をいただけたましたこと、大変光栄に存じます。ご推薦頂きました統計教育分科会主査 宮崎大学 藤井良宣先生をはじめ、立正大学 渡辺美智子先生、関係の皆様のご支援に、心よりお礼を申し上げます。

データサイエンスは誰かが取り組むものではなく、「自分ごと」として共感すること、取り組む仲間を信頼し、誠実にことに当たるということ。日本の WiDS 活動として私どもが発信してきたこれらのことを、教育活動の観点からご評価頂きましたこと心より嬉しく思います。

参考文献

[1] 内閣府(2019)令和元年度年次経済財務報告、付注 2-9 多様性と生産性の推計

https://www5.cao.go.jp/j-j/wp/wp-je19/h09_fchu0209.html (最終確認日: 2022.3.8)

[2] Fleming(2004)Perfecting Cross-Pollination, Harvard Business Review, Vol.82, Issue 9.

[3] Chrobot-Mason, D., Ruderman, M. N., & Nishii, L. H. (2014) Leadership in a diverse workplace. In D. V. Day (Ed.), *The Oxford handbook of leadership and organizations* (pp.683-708). Oxford University Press.

[4] 経済産業省 新・ダイバーシティ経営企業 100 選

<https://www.meti.go.jp/policy/economy/jinzai/diversity/kigyol00sen/> (最終確認日: 2022.3.8)

[5] 経済産業省 ダイバーシティ経営企業 100 選

<https://www.meti.go.jp/policy/economy/jinzai/diversity/kigyol00sen/entry/pdf/h27betten.pdf> (最終確認日: 2022.3.8)

[6] 文部科学省 ダイバーシティ研究環境実現イニシアティブ 女性研究者研究活動支援事業

https://www.jst.go.jp/shincho/josei_shien/index.html (最終確認日: 2022.3.8)

[7]<https://www.aacu.org/article/how-do-we-write-better-diversity-statements> (最終確認日: 2022.3.8)

[8] <https://diversity.berkeley.edu> (最終確認日: 2022.3.8)

[9]<https://www.widsconference.org/ww-best-of-wids-speakers.html> (最終確認日: 2022.3.8)



第19回統計教育の方法論 ワークショップ

理数系教員授業力向上研修会(東京)



高等学校におけるデータサイエンス教育の 現状と展望

兵庫県立姫路西高等学校
林 宏樹

兵庫県立姫路西高等学校について

2020年4月～

文部科学省スーパーサイエンスハイスクール (SSH) 指定校

高度な「知」を有するグローバルサイエンティストの育成
～AI時代を切り拓く～

全校生徒対象 「データサイエンスを基盤とする課題研究」

「知」の育成プロジェクト ... データサイエンス教育を受け、探究力を養う。

「知」の検証プロジェクト ... 生徒ひとりの自己分析力を養う。

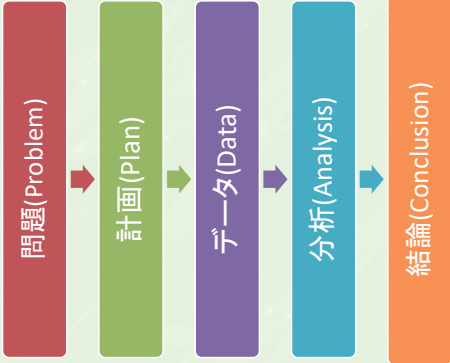
「知」の連携プロジェクト ... 大学・企業等との連携より、発信力を養う。

データサイエンス教育の概要

	1 年 次	探 究 準 備 期 間	2 年 次	探 究 実 践 期 間	3 年 次	探 究 展 開 期 間
学校設定 科目	データサイエンス研究 (4 単位) データサイエンス探究 (2 単位)	データサイエンス研究 (3 単位) データサイエンス探究 (2 単位)	データサイエンス探究 (2 単位)	データサイエンス探究 (1 単位)	グローバル研究 (2 単位) グローバル探究 (1 単位)	
概 要	データサイエンス教育の基礎・ 基本に関する知識・技能の定着 を学ぶ。 ・データ分析による課題発見力を 伸ばす。 ・データ解析による問題解決力を 伸ばす。 ・新たな価値を生む創造性のメカ ニズムを理解する。	データサイエンスを基盤とした 研究実践の実施 ・自分の興味関心のあるテーマ の研究を行う。 「データ活用実践」 「情報開発実践」 「自然科学実践」 の3つの類型に分かれて研究 を実践する。	データサイエンスの 学びの振り返り ・研究論文を作成する。 ・発表会等で外部発信を行う。 ・後輩の研究指導を行う。 ・自らのデータサイエンスの 学びを振り返り、知識・技能 の定着を認知する。			
実践内容						

データサイエンスと探究活動

統計的探究プロセスに、
PPDACサイクルというものがある。
確かな**問題**と**計画**の立案があつてこそ、
データ収集や**分析**を行うことができ、
問題解決へ導くことができる。
研究フレームワークである。サイクルであるから、結論フェーズで
新たな問題が発見されれば、
次のPPDACサイクルを回す。



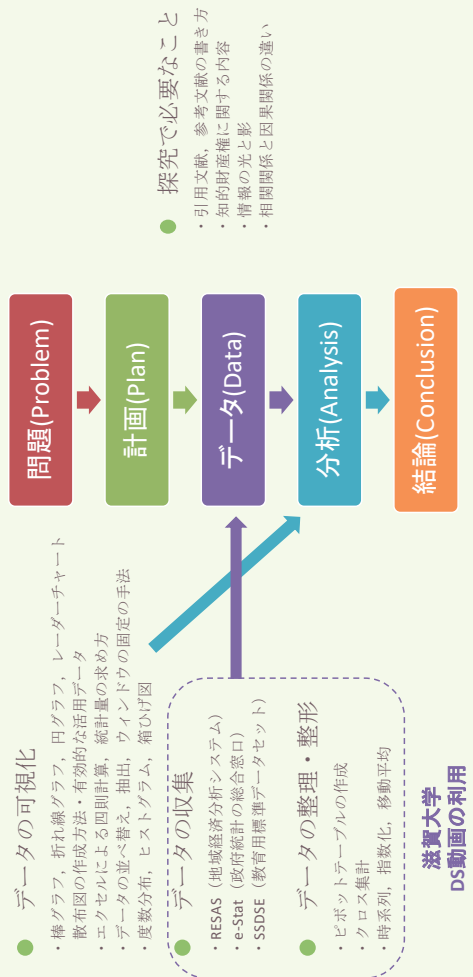
データサイエンスを基盤とした探究活動

統計的探究プロセス (PPDACサイクル)

総合的な探究の時間



DSにおける基礎・基本の学び



学校設定科目「データサイエンス探究」

1年普通科 (2単位) 対象人数：240名

社会と情報・総合的な探究の時間による代替



単回帰分析による分析モデルの構築

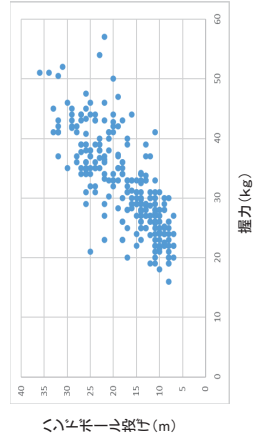
質問：ハンドボール投げの結果を予測したい。
ハンドボール投げの結果と関係のありそうな種目は何でしょうか。

予測したいもの → 「目的」 ハンドボール投げ(目的)を予測するための種目
ハンドボール投げ → 「目的変数」 「説明変数」を見つけよう！
「目的変数」に対して、「1つ」だけの「説明変数」との関係 ⇒ 単回帰分析

握力の測定値によって、ハンドボール投げが予測できないか？



質問
握力が30kgであった。では、過去のデータから
ハンドボール投げはどうか予測できますか？



データサイエンスで目指すこと

「現在は情報化社会」と言われ、データサイエンスの世界では、
「扱えるデータが容易に得られる社会」と解釈ができる。

データサイエンスは、ビッグデータから**有益な情報をあぶり出す**ことです。
多変量のビッグデータからあぶり出す解析を多変量解析という。

創造性

分析(A2 データ解析)

データサイエンスで活用できる手法例

第2学年の学び

回帰分析法(単回帰分析・重回帰分析)

- ・未知な事柄を予測するための手法 キーワード: 回帰式・回帰係数・決定係数・寄与率・p値・予測モデル標準化

- ・データ群を、平均0、標準偏差1にする手法

仮説検定(平均の差の検定)

- ・データ群とデータ群を比較し、平均の差が有意であるか判断する手法

クラスタリング(k-means法など)

- ・統計的にデータをグループ分けする手法

主成分分析法

- ・変数間の関係性から共通する因子(潜在変数)を導き出し、多くのデータを少ない変数に縮約する分析手法

pythonを活用した
データ分析

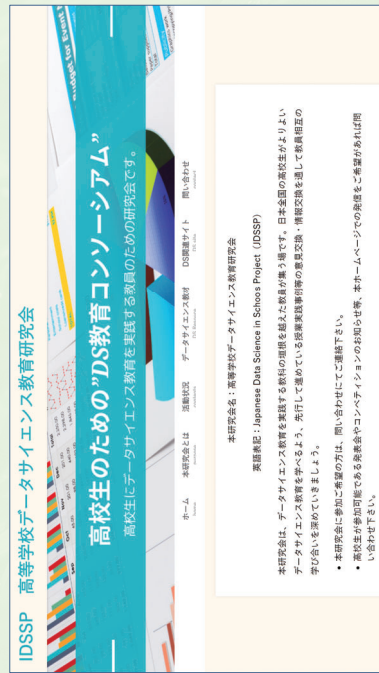
データサイエンス教育の普及



生徒の研究成果一覧

- (データ解析) データ分析コンペティション
令和2年度 優秀賞, 特別賞, 学校賞 令和3年度 特別賞, 学校賞
- (スポーツ分析) 中高生スポーツデータ解析コンペティション
令和2年度 最優秀賞 (バドミントン) ・奨励賞 (サッカー) ・敢闘賞 (野球)
- (地方創生) 和歌山県データ活用コンペティション
令和2年度 データ活用賞 令和3年度 大賞
- (健康ヘルス) Wids HIROSHIMA アイデアソン
令和2年度 優勝
- (統計) 兵庫県統計グラフコンクール
令和2年度 入選・佳作 令和3年度 特選・入選・佳作

JDSSP高等学校データサイエンス教育研究会



<http://ds-education.com/>

JDSP高等学校ターサイエンス教育研究会

研究会名：高等学校データサイエンス教育研究会

英語表記: Japan Data Science in Schools Project (JDSSP)

[illegible]

- ・本研究会に追加ご希望の方は、問い合わせはてご連絡下さい。オンラインで実施しておりますので、気軽に。
- ・参加費は無料です。
- ・高校生が参加可能な分野である音楽・英語・数学・物理・化学・生物・地学・天文学・宇宙科学・環境科学・情報科学・経済学・社会学・法学・医学・工学・農学・芸術・スポーツ・健康科学・その他分野の発信をご希望があれば問い合わせ下さい。

● 最新情報

2022年7月19日	日本文学	第1回国際教育の方向性ワークショップ・国際教育委員会の招待会（東京）
2022年11月7日	初年度会	国際系教員長・ラーニングエッセンス向上研修会（愛媛）
2023年12月21日	05847	小学5・6年向け算数教材
2024年1月19日	初年度会	大学学校教員向け統計授業（神戸大学経済データサイエンスセンター）
2024年11月7日	05850	【12月1日】第1回 JSSSP-データサイエンス教育研究会
2024年10月30日	初年度会	【11月1日】第1回 JSSSP-データサイエンス教育研究会
2024年10月21日	05847	統計と情報の専門刊「エントラレー」への寄稿
2024年10月21日	05847	日本統計学会月刊「統計」10月号

鯨一博の会研究

年に5・6回の研究会を
オンラインで実施しています。

研究会の構成

- 高等学校の実践事例
- 専門家のDS事例

JDSSP高等学校データサイエンス教育研究会

[研究会の記事一覧](#)[illegible]

JDSP高等学校タータサイエンス教育研究会

DS教材の一覧

- ・高等学校の教員が活用できる教材の紹介
- ・外部発表における生徒の成果作品

ホームページ | DS研究

DS教材の記事一覧

小5 5、6年向け組技教材 2023年12月21日 DS教材	統計と機械の専門誌「エストレーラ」への寄稿 2023年10月12日 DS教材
小5 5、6年向け組技教材 2023年12月21日 DS教材	データイエンス関連書籍教材（基礎講座印刷版講習会等） 2023年10月4日 DS教材
小5 5、6年向け組技教材 2023年12月21日 DS教材	DS研究季刊「SLP International winners」(英訳版付) 2023年9月17日 DS教材
小5 5、6年向け組技教材 2023年12月21日 DS教材	日本統計学会月刊「統計」8月号 2023年9月10日 DS教材
小5 5、6年向け組技教材 2023年12月21日 DS教材	「データ活用」に載せる保護モデル、教材 2023年9月3日 DS教材
小5 5、6年向け組技教材 2023年12月21日 DS教材	日本統計学会月刊「統計」9月号 2023年10月12日 DS教材
小5 5、6年向け組技教材 2023年12月21日 DS教材	「DS研究季刊」全知事年度DS活用研究発表会（発表要旨） 2023年7月13日 DS教材
小5 5、6年向け組技教材 2023年12月21日 DS教材	日本統計学会月刊「統計」7月号 2023年7月4日 DS教材
小5 5、6年向け組技教材 2023年12月21日 DS教材	2023年7月4日 DS教材

1 2 3

JDSP高等学校タータサイエンス教育研究会

論文・報告書の検索・研究

 第1回 日本学術会議 第3 回研究会 2023.04.26	 第1回 日本学術会議 第3 回研究会 2023.04.26	 第1回 日本学術会議 第3 回研究会 2023.04.26	 第1回 日本学術会議 第3 回研究会 2023.04.26	 第1回 日本学術会議 第3 回研究会 2023.04.26	 第1回 日本学術会議 第3 回研究会 2023.04.26	 第1回 日本学術会議 第3 回研究会 2023.04.26
--	--	--	--	--	--	---

JDSSP高等学校データサイエンス教育研究会

高校生が参加できる発表会

- ・発表会名, 申込締切, 開催日, 主催などの情報を掲載

● 外部発表会一覧

高校生が参加できるコンテスト・発表会です。
本ページへの掲載を希望の方は問い合わせください。

発表会名	申込締切	開催日	主催
FESTAT (全国統計探究発表会)	2021年5月31日	2021年7月18日等	香川県立綾歌高等学校
第1回 中学生・高校生データサイエンスコンテスト	2021年6月30日	発表要綱等参照 2021年9月7日	神戸大学
第13回 データビジネス創造コンテスト	2021年8月3日	2021年9月11日等	慶應義塾大学SFC研究所 データビジネス創造・ラボ
統計データ分析コンペティション	2021年8月10日	論文提出 2021年9月10日	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
第5回 昭和歌山県データ活用コンペティション	2021年10月22日	2021年12月18日	和歌山県
WOMAN IN DATA SCIEN	2021年9月30日	資料提出	広島県 広島大学 ひろしま自動車産業官博



ご清聴ありがとうございました。

JDSSPデータサイエンス教育研究会のホームページに掲載できる内容がありましたら、問い合わせせてください。

アドバイザーボードとして、指導助言いただける専門家の方もお声掛けください。ご協力よろしく願っています。

統計教育・データサイエンス教育の発展のために！

今後の展望

JDSSP高等学校データサイエンス教育研究会

- ・教育実践例の集積
- ・SSH校、教員方の作成教材の掲載
- ・コンテスト入賞などの研究紹介
- ・外部のコンテストなど申込要項の集積
- ・人材の橋渡し（指導員・専門家）

困ったときに
立ち寄れる研究会

データサイエンス教育を
実践する教科の垣根を
越えた教員が集う場

統計教育・データサイエンス教育を指導する教員
学ぶ生徒が活用できる集積サイト作り

はじめに

▶ COVID-19 における授業の変化

- ▶ オンライン授業の実施と課題
 - ▶ オンライン授業：オンデマンド型、双方向型、…
 - **オンデマンド型**の課題：リアルタイムに質問不可、学生の理解度は課題の達成度と学生からの質問以外は不明。**何度も繰り返し直すこともあり**、教員の負担は例年の数倍？教員が楽をしていると学生から言われることも…
→ **しかしながら**、1) 繰り返しができる（最適な授業ができる）、2) 何度でも密講生の都合で繰り返し返してみることができる、3) 時間や場所の制約を気にせずに授業展開が可能、などの利点も…
 - **さらに**、文系学生が苦手意識を持ち、PC 教室での演習も行う「統計・データサイエンス」においては、1) 情報スキルの個人差への対応、2) PC 室の座席による受講制限の撤廃、の点でも利点有り
 - 双方向型の課題：ネットワーク環境の課題、受講生の表情や動作の観察不可、…
 - **しかしながら**、感染予防時の「熱量のある」グループワークの悩み
 - **さらに**、遠隔地との PBL の実施も可能、遠方地の企業との連携、海外との連携も可能、…**複数のキャンパスで同時同授業の開催も可能**

いろいろな課題や改善方法が議論・提案されつつある

オンライン授業時代のデータサイエンス授業の展開

→ アフターコロナを見据えた新しい統計教育に関する授業法の可能性

▶ 2

JCOTS2022

今後の大学での統計・データサイエンス

- ▶ 小中高における学習指導要領における統計拡充
 - ▶ 小中高で統計教育に関する内容の拡充、体系的な学びの展開
 - ▶ 高校情報科では情報Ⅱで「データサイエンス」を導入
 - ▶ 「探求学習」では統計・データサイエンスに関する内容の活用の期待。統計的問題解決（PPDACサイクル）の学修
- ▶ 大学においても文理問わずデータサイエンス教育の導入
 - ▶ データサイエンス学部・学科、コース、専攻、研究所の開設
 - ▶ 「数理・データサイエンス・AI教育認定制度」の導入

数理・データサイエンス・AI 教育の必修化を目指すなら、（必要性は理解しても）苦手意識をもつ学生の不安、指導する側の不安も増大？

▶ 3

JCOTS2022

余談：もうひとつの裏の悩み

- ▶ Excel はとりあえず、SPSS はいつまで…
 - ▶ Excel はとりあえず使えた方がよい。とはいえ、大規模データや多変量解析など、高校の情報科の内容次第で大学でもレベル上げは必要？
- ▶ オンライン授業では市販ソフトウェアを前提とした授業は難しい（SPSS を前提とした授業は厳しい）
- ▶ R や Python を（高校でもしているので）導入したいが、文系では学生の抵抗もあり難しい？

オンライン授業のため、1) フリーソフト（例：R）を前提とした統計授業の展開（卒論指導も SPSS でできなくなったので…）、2) PC 教室の人数制限撤廃はもちろんのこと、3密を避けることにも貢献、などもあり。
→ Excel と R で展開し、いかに使いやすいかを伝えられるかが課題
→ PBL のグループワークでも、オンラインでも画面を見せあえる

▶ 4

JCOTS2022

実施したオンデマンド型授業について

- ▶ 朝日ネットの manaba および respon 社の respon を利用
 - ▶ 100 分×14 回授業。ただし演習時間を加味して、従来から実践演習→トピックA&ミニ演習→トピックB&ミニ演習と、授業自体は 30 分×2 つ、のこりは演習時間
 - ▶ 授業動画を 7-15 分を一つとして細かく。資料は公開まま
 - ▶ 課題は respon で提出。次回の講義日前日まで。
 - ▶ 本来の授業時間中に学習の振り返りと質問・相談の respon あり
- ▶ その他
 - ▶ オンラインのため休講は一切なし
 - ▶ 中間テストは小テスト機能で実施
 - ▶ 中間レポートも一度あり

▶ 5

JCOTS2022

実践女子大学における仮想空間を活用した産学連携 PBL 型授業の取り組み

- ▶ 対象授業：
 - ▶ 情報リテラシー活用
 - ▶ 共通教育科目
 - ▶ 2 年生以上対象・後期
 - ▶ 2 つのキャンパス同時開催
- ▶ 利用プラットフォーム
 - ▶ 2020 年度 zoom
 - ▶ 2021 年度 oVice
 - ▶ Microsoft Office 365
 - ▶ 発表スライドの協働作成
 - ▶ LMS で分析データの共有



- ▶ 協力企業
 - ▶ インターテックノスフィア様
 - ▶ 課題とデータ
 - ▶ 課題を踏まえた自社調査データを貸与
 - ▶ キックオフ、中間発表、最終発表に参加

▶ 7

JCOTS2022

双方向授業の展開－バーチャル教室の活用

- ▶ Zoom 等のツールの利点と欠点の例
 - ▶ ○ 遠隔地からの接続も可能であり、ブレイクアウツを利用するとグループワークも可能
 - ▶ ...
 - ▶ × グループ内の活動は見えるが、他チームの活動が見えない
 - ▶ × 全体ではビデオオフにすることも多く、本来の授業の感覚とは少し違和感あり
 - ▶ ? 産学連携 PBL を行う際にも、人数の感覚が教室でしている感が起きない...

Zoom などでも十分に遠隔地授業もできるが、何かさらにできないか...

▶ 6

JCOTS2022

対面時のインテリジェクノスフィア社とのPBL

21年度 PBL 自主企画調査テーマ

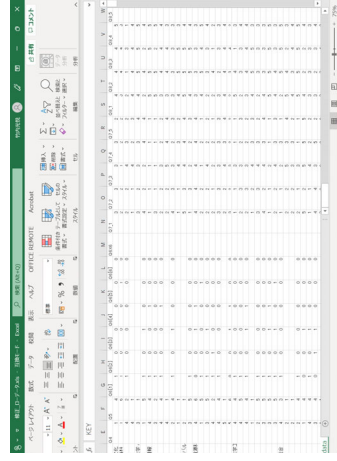
リサーチデータをもとに、

大学生の視点で、大学生が感じる

“ストレス”を解消するアプリを企画する
(※ディバイス問わず)

＜プロセス＞
リストアップする調査の分析
・データ解析の結果
・アプリのコンセプト
・アプリの活用シーン
・ユーザーの感情を一言にまとめる
・アプリの要素「コンテンツ」、「ディバイス/プラットフォーム」、「マネタイズ」を含める

＜アウトプット：企画書＞
・データ解析の結果
・アプリのコンセプト
・アプリの活用シーン



調査名：大学生が抱えるストレスに関する自主企画
調査期間：2021/06/04～2021/06/07

▶ 8

JCOTS2022

バーチャル教室を利用した産学連携 PBL

oVice でのバーチャル教室

校種名 竹内光俊

2021-08-19 08:00

本講義では渋谷と日野キャンパスの両キャンパスで同時開講します。そのため、オンラインが前提になります。oVice というバーチャル教室を使っています。

初回では zoom で最初に説明しますので、zoom アサイン、アサイン、oVice を初回に説明します。2回目以降は oVice で授業時間になったらログインし oVice

1. oVice へのログイン URL は <https://isenjins>

2. oVice へのログインの仕方(会員登録しない) <https://www.notion.so/oVice-981b06167b904b2>

3. oVice の基本操作の説明 <https://www.youtube.com/watch?v=C8r02gYDA>

不明点はコングレションで質問してください。

バーチャル教室と会議室



使い方の説明

9

JCOTS2022

まとめ

- ▶ オンライン授業におけるバーチャル教室の活用
- ▶ 教員も利用者の不慣れだったため、改善の余地あり
- ▶ 機能も充実してきたことから、オンラインの利点を活かした授業開発に期待
- ▶ 遠隔地授業、バーチャル教室ならではの特色ある授業を検討

11

JCOTS2022

受講者アンケートより

受講者

oVice と zoom



9少	8ど	7い	6他	5視	4画	3VO	2ハ	1総
----	----	----	----	----	----	-----	----	----

10

- 1.ほとんどのグループワークは対面でやるけれど、oViceとかのインターネットを介してやるのはあまりなかったから手こずったところがあった。だけど、みんなで協力して上手くやれたと思う。
- 2.生きたデータ活用から、実際の企業に提案できる貴重な機会を得られたととても身になりました。oVice自体も、私自身は初の試みでしたがすぐに慣れて十分に活用することが出来たので安心しました。先生方にも多くのサポートをしていただき、大変ありがとうございました。
- 3.この授業は履修登録期間の最後に追加した授業でしたが、最もやりがいのある授業でした。オンラインでのグループワークや企業の方・
- 4.今学期で1番学んだことが多かった授業だと思いました。1限はつまらなかつたですが、履修してよかったです。半年間ですが、お世話になりました。インテジーさんのフィードバックは、中間も最終プレゼンも自分のために視点をとても参考になりました。おふたりの先生方もこの授業で初めてでしたが、わかりやすい資料や講義、また、雑談などもきいていて頭に入りやすく、また、実践ばかりだったので、面白い授業でした。ありがとうございました。
- 5.この授業は履修登録期間の最後に追加した授業でしたが、最もやりがいのある授業でした。オンラインでのグループワークや企業の方・

オンライン FESTAT の 2 年間を通して見える探究発表会の可能性

石井 裕基・香川県立観音寺第一高等学校 数学科 教諭
〒768-0069 香川県観音寺市茂木町四丁目 2 番 38 号
TEL : 0875-25-4155 FAX : 0875-25-4145
Mail : ap0170@kagawa-edu.jp

1. はじめに～FESTAT（全国統計探究発表会）とは

本校が、2019 年 4 月から 2022 年 3 月までの 3 年間指定された SSH 科学技術人材育成重点校（広域連携）の取り組みの一つとして実施している統計分析を活用した探究発表会を「FESTAT（フェスタット）」という。その最大の目的は、「互いに学び・互いに繋がる場を創出する」ことにある。そこに参加している高校生や大学生、教員、統計専門家などが交流を通して繋がり（図 1）、AI・データサイエンスに興味関心をもつ若者が増えることを目指している。

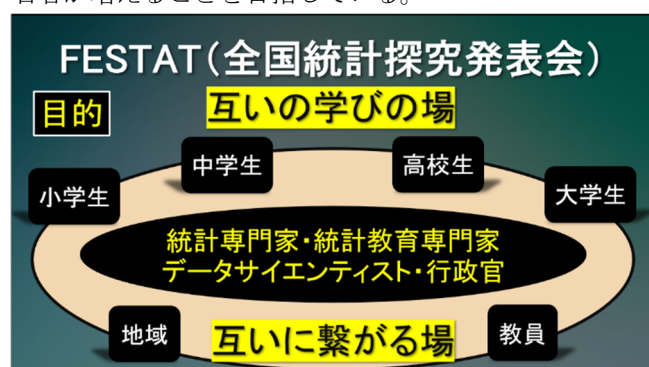


図 1 FESTAT イメージ図

2. FESTAT2019～2021 の報告

(1) FESTAT2019（開催方法 対面とオンライン）

①場所 ハイスタッフホール（観音寺市民会館）

滋賀大学大津キャンパス

②日程と内容 司会 本校生

7 月 28 日（日） 9 : 30～16 : 30

ポスターセッションと口頭発表

講演

③指導助言者 アドバイザー 11 名

④発表参加校

・中学校 1 校 1 チーム

観音寺市立観音寺中学校

・高等学校 9 校 21 チーム（本校含む）

長野県上田高等学校、愛知教育大学附属高等学校、

愛知県立旭丘高等学校、福井県立若狭高等学校、

山口県立下関西高等学校、徳島県立脇町高等学校、

愛媛県立松山南高等学校、熊本県立宇土高等学校

・大学 2 校 12 チーム

（内、10 チームがポスターのみ）

滋賀大学、香川大学

⑤発表タイトル（一部）

「東御市の人口減少を止めたいー地域内経済循環を目指し、企業誘致によって今までにない雇用を創造するー」「バドミントンにおいて有利なのはサーバーかレシーバーか」「我売洋菓子者也」「交流人口増加による愛媛県の活性化」「うどんの消費の

統計分析」「駐車待ち時間を短縮する駐車場ナビゲーション」「香川県の交通事故数を減らすのに有効な方法は何か」「『無難に外角一辺倒』を統計的に検証する」

⑥生徒の振り返り

・人によって発表の仕方というのが変わっていて、見聞きしていてとても面白かった。アイデアに関して自分の知見が深まったと感じました。本当に来て良かったです。

・途中経過という状態だったけど、グラフに関してや今後のことについて様々なアドバイスを頂けて、来て良かったと思った。

(2) FESTAT2020（開催方法 オンライン）

①場所 YouTube, ZOOM

②日程と内容 司会 大学生

7 月 24 日 FESTAT キックオフイベント (ZOOM)

大学生メッセージ、参加生徒の決意、講演

9 月 5 日～11 日 事前動画視聴 (YouTube)

講演動画視聴 → 事前質問入力

研究発表動画視聴 → コメント入力

9 月 13 日 FESTAT 本番 (ZOOM)

講演への事前質問への回答

4 つの分科会に分かれての指導助言・交流

ファシリテーターを大学生に依頼

交流会

外部アドバイザーと高校教員間での統計教育等に関する意見交換

③指導助言者 アドバイザー 8 名、大学生 8 名

④発表参加校

・高等学校 17 校 32 チーム（本校含む）

群馬県立高崎高等学校、長野県上田高等学校、愛

知県立旭丘高等学校、愛知教育大学附属高等学校、

福井県立若狭高等学校、洛星高等学校、大阪府立

四條畷高等学校、奈良学園高等学校、神戸大学附

属中等教育学校、兵庫県立姫路西高等学校、徳島

県立脇町高等学校、高松第一高等学校、愛媛県立

松山南高等学校、愛媛県立宇和島東高等学校、山

口県立下関西高等学校、鹿児島県立国分高等学校

⑤発表タイトル（一部）

「人の流れを見える化する」「統計的視点から見た

ヒット曲の法則」「21 世紀型商店街 ～私たちが

統計学を用いて考えた新しい形の商店街～」「県別

交通事故調査」「百舌鳥古市古墳群の可能性」「芸

能人の広告の影響力ー高校生のファン心理に着目

してー」「バドミントンが強くなるスポーツ分析」

『ちくわ』と『かまぼこ』売上げの差」「早めに

追い込めは正しいのか？」「国連本部の最適所在地

について」「黄金比の可能性 ～美しさに秘められた能力～」
「観音寺の観光業の低迷の原因特定と改善策の思案 ～新型コロナを乗り切るために～」

⑥生徒の振り返り

- ・はじめての発表で学ぶことが多かった。先生方にアドバイスをいただけたので研究の進め方やスライドの作り方を学ぶことができ良かった。また、他の高校生の発表を聞き質問をすることになり、発表をより理解しようとしながら聞いた事で発表を聞くのが面白かった。
- ・事前にオンラインで先生方の話を聞いてさらに当日詳しく聞けたのでオンラインならではの点だと思いました。オンラインだったからこそ先輩方や講師の方々、たくさんの高校生と交流でき、様々な意見を頂けて良かったです。『スライドを作ったら一旦親に見せてみるといいよ』『研究内容も大事だけど何よりストーリー性はとても大切だよ』といったこれからも活用できる意見を聞くことができて参考にできているのでいい経験になりました。
- ・全国のさまざまな高校の研究を知ることができて、とても良かった。面白いことに着目して研究していたり、斬新なところに着目して考えていたりしていたところもあって面白かった。また、質問をしたり、それに答えたりするという経験を積めたので良かった。

(3) FESTAT2021 (開催方法 オンライン)

①場所 ZOOM, oVice

②日程と内容 司会 大学生

7月18日 FESTAT キックオフイベント

大学生メッセージ、講演、(ZOOM)

大学生による課題研究発表 (oVice)

8月21日 FESTAT 本番 (ZOOM, oVice)

講演 (ZOOM)、研究発表 (oVice)

交流会 (ZOOM)

③指導助言者 アドバイザー 20名、大学生 9名

④発表参加校

- ・高等学校 18校 40チーム (本校含む)

宮城県仙台第一高等学校、群馬県立高崎高等学校、愛知教育大学附属高等学校、愛知県立旭丘高等学校、福岡県立若狭高等学校、滋賀県立膳所高等学校、洛星高等学校、大阪府立四条畷高等学校、奈良学園高等学校、西大和学園高等学校、神戸大学附属中等教育学校、兵庫県立姫路西高等学校、広島大学附属高等学校、徳島県立脇町高等学校、高松第一高等学校、愛媛県立西城高等学校、愛媛県立松山南高等学校

⑤発表タイトル (一部)

「ヒト感染症の新たな出現について、動物との接し方から考える」「投票立の変化の原因を探る」「新型コロナウイルスによる経済状況の分析」「祭りの必要性和地域力の関係について」「電車の運行の解析～特急列車を走らせるには～」
「高校生によるスクールカウンセラーの認識及びニ

ーズについて」「COVID-19～過去から学ぶ、コロナウイルスの収束～」
「リモートセンシング技術を用いた伝統農法の効果の検証」

⑥生徒の振り返り

- ・データ分析と現実のギャップをどうやって埋めていくかということが最大の課題になるのだと思いました。社会に実際に貢献することのできる研究の実現の難しさを学ぶことが出来ました。
- ・身近なことに疑問を抱き、突き詰めて調べているのがとてもよく伝わってきた。
- ・自分自身、最初は研究をすることにあまり興味がなかったけれど、いざやってみるととても楽しくまたやりたいと思ったから、これから FESTAT に参加する人が増えたら、研究をすることに興味を持ってくれる人が増えていくんじゃないかなと思いました。

⑦オンライン発表会の振り返り

- ・聴衆の反応が見れなかったのが、発表する側として少し不安でした。チャットを使ったアクションがもっと盛んになれば良いと感じます。
- ・とても便利でしたが、質疑応答は相手の質問内容や質問の理由が分かりづらく戸惑いました。もう少し経験を積む必要がありました。
- ・オンラインという形であったが、逆にオンラインだったからこそたくさんの学校が参加でき、午後から予定があった私も途中までではあるが、参加することができた。
- ・良い経験になったと思います。oVice を使ったイベントには初めて参加しましたが、自由に動くことができ、自分の聞きたい発表を聞くことができるなど、良い点がとても多かったです。
- ・こんなにも多くの高校生がデータを使った研究をしていると知り、自分たちも研究を頑張ろうと思えました。特に大学生の皆さんの研究発表が参考になりました。ありがとうございました。
- ・対面でなく、オンラインということであまり慣れていなかったが、上手く使えばより効果的で対面と同じような効果になるのではないかなと思った。

3. 最後に

コロナ禍の中、対面に近い探究発表会ができないか考えたどり着いたのが、ZOOM と oVice の融合によるオンライン開催である。「慣れ」という壁を壊すことができれば、キーワード「融合」の元、新しい探究発表会が見つかることが予測される。

【参考文献・参考 URL】

1. 日本統計学会統計教育分科会等 (2020), 第 17 回統計教育の方法論ワークショップ・理数系教員授業力向上研修会 (東京) および第 9 回科学技術教育フォーラム, 課題研究における数理・データサイエンス教育の実践と SSH 重点枠の記録
2. 日本統計学会統計教育分科会等 (2021), 第 18 回統計教育の方法論ワークショップ・理数系教員授業力向上研修会 (東京), コロナ禍でのチャレンジ, オンライン全国統計探究発表会 (FESTAT)

バーチャル空間でのオンライン教育の 動向や可能性について

oVice(オヴィス)株式会社 PRマネージャー
薬袋友花里(ミナイユカリ)



バーチャル空間でアバターを使って交流しているイメージ

◆発表内容

さまざまなもののオンライン化が急速に進み、教育機関においてもオンラインで授業や学生同士の交流を行う場面が増えてきました。ポストコロナにおいて、どのようにオンラインを活用すべきか、悩まれている方も多いのではないのでしょうか。ここでは教育現場でのオンラインツールの活用状況や、今後の可能性についてお話いたします。

◆アバターで交流するバーチャル空間「oVice(オヴィス)」について
ウェブ上で自分のアバターを自由に動かし、相手のアバターに近づくことで簡単に話しかけられる2次元のバーチャル空間です。2020年8月のサービス開始以降、テレワークにおけるバーチャルオフィスやオンラインイベント、オープンキャンパスなど、さまざまな場面での活用が進み、2022年2月時点で発行スペース数は2万件を突破しました。

中学校における批判的統計リテラシーを育成するための教材開発

筑波大学大学院人間総合科学研究科（日本学術振興会特別研究員）・細田幸希

筑波大学大学院人間総合科学研究科・大脇和志

〒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

今日のデータ駆動型の社会において、個人が新聞やインターネット等で日常的に目にする統計情報を適切に解釈し、批判的に考察しながら意思決定を行う能力を身につける必要がある。統計教育研究では、そのような能力を統計リテラシー (Statistical Literacy) と定義し、従来からその重要性が指摘されている (例えば, Gal, 2002)。我が国においても統計リテラシーに関する研究が数多くなされてきている。しかし、そのほとんどがデータを収集して統計的分析を行う活動が中心であり、統計情報を批判的に解釈する活動に焦点を当てた研究は十分に行われているとはいえない状況である (古賀, 2019)。

そこで、本研究では、批判的市民性 (critical citizenship) の育成を重視した Weiland (2017) の批判的統計リテラシー (Critical Statistical Literacy) の概念に着目する。批判的統計リテラシーとは、日常生活における社会的政治的 (sociopolitical) な問題や文脈を考慮しながら統計情報を批判的に解釈し、我々が生きている現実世界をよりよく理解するために統計をどのように活用できるかについて検討する能力である。このような批判的統計リテラシーの議論を踏まえれば、学校教育において、生徒が社会の中で用いられている統計情報を解釈し分析するだけでなく、その統計情報に関する背後の文脈まで考察し問い直すことを意図した教材の開発が必要であると考えられる。

したがって、本研究の目的は、批判的統計リテラシーの育成を意図した教材を開発するとともに、その意義と課題を明確化することである。この目的を

達成するために、まず、先行研究による理論的検討から批判的統計リテラシーの育成を目指した教材開発の要件を提示する。次に、提示した要件に基づいて具体的な教材を開発し、教材の教育的価値を考察する。

2. 理論的枠組み

(1) Weiland の批判的統計リテラシーの枠組み

Weiland (2017) は、従来の統計教育で重要視されている統計リテラシーの育成に、批判的市民性の側面を取り入れて育成を目指すべきだと指摘する。ここで想定されているのは、「地域社会や政府に積極的に参加できるようにエンパワーされ、地域社会や政府の中で不正な状況を生み出している構造を問い直し、その不正を（再）生産する構造の変革に積極的に取り組む」(Weiland, 2017, p.35) 市民像である。このような批判的市民性の観点を取り入れることで、統計を用いて社会に内在する多元性への認識を促し、多様な価値観をもつ人々に不利益をもたらす構造に疑問を投げかけ、明日のより公正な世界の創造に貢献する市民を育成できると述べている。

また、Weiland (2017) は、統計的探究で中心となる文脈の考慮の必要性も指摘している。ここでの文脈とは学校生活といった身近な文脈ではなく、政治や環境、人種などの社会的政治的な文脈であり、生徒がそれらの文脈について考慮し複雑な社会的政治的な問題を理解することも、統計リテラシーの育成にとって重要な側面であると述べている。以上より、Weiland はこれらの側面を反映した批判的統計リテラシーの理論的枠組みを次のように示した (表 1)。

表1 批判的統計リテラシーの枠組み (Weiland, 2017, p. 41 をもとに筆者らが訳出)

読解 (Reading)	記述 (Writing)
1. 言語や統計の象徴体系 (symbols system) を理解し、多様な文脈で遭遇する統計情報やデータに基づく議論を批評 (critiquing) することで、社会で作用しているシステム構造に関する認識を深めること。 2. データに基づく議論により形成・強化される社会構造を特定し、問い直す (interrogate) こと。 3. 自己の社会的立場、主観性 (subjectivity)、政治的文脈を理解し、自己の社会歴史的、政治的知識を持ち、それが情報の解釈にどのような影響を与えるかを理解すること。 4. 統計情報の情報源、収集、報告、そして、それらが著者（ここでは情報の発信者）らの社会的立場、社会的政治的、歴史的レンズによってどのように影響されるかについて評価すること。	1. より公正な社会を目指して、不当な構造を揺さぶり (destabilize)、再構築するための努力の一環として、統計的な情報や議論を伝えるために統計調査を活用すること。 2. 社会的政治的な不当な課題を軽減し解決するために統計調査を活用すること。 3. 統計的な問題、データ収集、分析方法を策定する際に、社会的な弁証法的葛藤 (dialectical tensions) を克服し、統計調査の結果でそのような葛藤を浮き彫りに／強調すること。 4. 統計調査の結果を報告する際に、自己の社会的立場、主観性、政治的文脈を他者に伝え、それが世界に対する自分の意味づけをどのように形成するかを説明すること。

上記で示された批判的統計リテラシーの枠組みは、ばらつきのある偏在性を前提とした上で、①統計を読解、記述することに関する重要性の強調、②言語を読解、記述しながら世界を読解、記述することによる批判的意味の強調、③社会的政治的な文脈の考慮に関する重要性の強調の3つの観点で構成されている。まず、一つ目の統計の読解、記述では、言語と統計の象徴体系を理解し、それらを用いて自身の考えを説明できることが求められている。これは、社会の多様な文脈において出会う統計情報やデータに基づく議論を批評し伝えるために必要であるとされている。また、二つ目の言語の読解、記述とあわせた世界を読解、記述では、データに基づく議論を読み解き、その議論によって形成または強化される社会構造を特定し問いただすことが求められている。このことは、今日では公的データ等が豊富にあるという背景から、統計を用いて社会的政治的な問題を分析し、解決のための新たな指針を提供することに繋がるとされている。

さらに、三つ目の社会的政治的な文脈の考慮では、統計的議論を読解、批評、評価する際に自身の社会

的位置や主観性、政治的文脈が情報の解釈にどのように影響を与えるのかを理解することが求められている。これは、統計的議論で記述される世界を読み解くにあたって、個人によって形成されるフィルターによる影響や、データを収集し情報を発信する他者の立場を考慮する必要があるとされている。そのための方法として、調査や実験におけるバイアスの原因に着目する重要性を指摘している。

(2) 本研究における教材開発の要件の設定

Weiland (2017) は、前節で示したように批判的統計リテラシーの概念を理論化しているが、その育成を目指した実践への課題についても指摘している。具体的には、批判的統計リテラシーの視点を数学教育カリキュラムに導入する困難性や、数学教師が扱えるリソースの不足を挙げている。そこで、これらの課題を考慮した上で、本研究では、批判的統計リテラシーの枠組みに基づき、次のような教材開発の要件を設定した。

第一に、批判的市民性の育成を目標とした教科横断的なカリキュラムを構想することを念頭において、

その中に位置づけられる教材とすることである（要件1）。この要件を設定したのは、従来の統計リテラシー教育を拡張する批判的統計リテラシー概念を踏まえると、日本の学校教育で断片的に行われている統計に関する教育の連携が不可欠であると考えたからである。

第二に、統計情報やデータに基づく議論で用いられる統計的知識や概念の理解を促すことである（要件2）。この要件を設定したのは、Weiland（2017）が指摘するように、統計リテラシーに必要な統計的概念や知識をおさえることが必要であると考えたからである。具体的には、前章で述べたばらつきの偏在性と関連して、物事の測定方法や標本抽出の方法に着目することが挙げられる（Weiland, 2019）。

第三に、統計情報の送り手、受け手が置かれた社会的政治的文脈に注意を促し、異なった文脈に立てば異なった解釈が可能であるような余地を教材に持たせることである（要件3）。この要件を設定したのは、自己の立場や既存の社会構造を問い直すためには、現実の争点を含んだ社会的政治的な課題を教材化する必要性があると考えたからである。

3. 教材開発

(1) 開発の手順

本研究では、前章で設定した教材開発の要件に基づいて次のように教材開発を行った。

一つ目の要件は、「批判的的市民性の育成を目標とした教科横断的なカリキュラムを構想することを念頭において、その中に位置づけられる教材とすること」であった。これに対し、本研究では、現行の学習指導要領に含まれる統計リテラシー育成に関連すると考える学習内容を検討し、中学校段階における数学科と社会科の連携を軸に構想を行った。まず、数学科においては、中学校第三学年のデータの活用領域において、「標本調査の方法や結果を批判的に考察し表現すること」（文部科学省, 2018b, p.156）を意図した活動が求められている。この活動では、新聞や

インターネットから得られる実際の標本調査の結果を解釈する方法について検討することが行われる。ここでは、統計情報で用いられている統計的手法の妥当性を検討するだけでなく、その背後にある文脈を明確に捉え考慮する必要があると考える。そのためには、統計的知識や概念を理解しているだけでは不十分で、教科を問わない幅広い文脈的知識についても把握しておかなければならないと考える。したがって、本研究では、教科横断的なカリキュラム構想の一つの視点として社会科との連携を図った。社会科においては、「平和で民主的な国家及び社会の形成者に必要な公民としての資質・能力」の育成を目標としている。ここでの「公民としての資質・能力」は、批判的的市民性に再概念化が可能である。また、社会科で育成する資質・能力のうち「知識・技能」は「社会的事象等に関する理解などを図るための知識」と「社会的事象等について調べまとめる技能」に整理されており、「社会的事象に関する様々な情報を効果的に収集し、読み取り、まとめる技能を身に付ける学習活動を重視する」ことが配慮事項に挙げられている（文部科学省, 2018a, pp.186-187）。実際に、社会科では地理、歴史、公民いずれの分野においても、様々な統計資料を活用して、現代社会の諸課題を捉え、多面的・多角的に考察する学習が展開されている。ここから、社会科と数学科が連携することで、統計情報の文脈に対する理解をより充実させることができると考えられる。

次に、二つ目の要件は、「統計情報で用いられる統計的知識や概念の理解を促すこと」であった。本研究では、第一の要件で述べた標本調査の学習場面を想定して教材開発を行う。ここで扱われる統計的概念として、母集団や標本、標本抽出に関わる概念が挙げられる。また、バイアスが生じる要因に関しては、Hirsch et al.（2015）によるバイアスが生じる要因の分類¹⁾に基づいて取り扱うこととした。

さらに、三つ目の要件は、「統計情報の送り手、受け手が置かれた社会的政治的文脈に注意を促し、異

なった文脈に立てば異なった解釈が可能であるような余地を教材に持たせること」であった。これに対し、本研究では、バイアスへの理解をより深めるために、実際の世論調査を教材化する。具体的には、報道各社が行う世論調査間の内閣支持率の差を扱う。内閣支持率は、毎月の世論調査で報道各社が尋ねており、その高低は内閣退陣や解散総選挙のバロメーターとなるなど、大きな社会的政治的影響力をもっている（岩本，2018）。世論調査は統計的な手法で行われるが、質問の仕方や調査日などによるバイアスを含み、調査結果には「ずれ」が生じる。しかし、この調査手法については十分に理解されず、「ずれ」は誤差としてではなく、報道各社のスタンスと結び付けて理解されることがしばしばある。結果として、社会には世論調査に関する様々な言説が飛び交っている。このような世論調査の内閣支持率を取り上げることで、バイアスを吟味させ、現状の世論調査の解釈のされ方を問い直し、多様な解釈を促すことにした。

(2) 開発した教材

上記の手順で開発した「岸田内閣の支持率」の教材に関する概要を示す。本教材は、世論調査の内閣支持率の差が生まれた要因は何かについてバイアスに着目して探究することを目的とした教材である。

【問題を定式化して、世論調査の仕組みを把握する】

初めに、教師が岸田内閣発足時に行われた世論調査の記事を提示し、報道各社によって内閣支持率が異なることを認識させる（10月4日から17日までの間に実施された10の世論調査による内閣支持率は、40.3%から63.2%までの幅があった）。そして、具体的な問題を以下のように示す。

〔問題〕

岸田内閣の支持率の差はなぜ生まれたのだろう。

また、教師が作成した報道各社による世論調査に関する一覧表²⁾をもとに、生徒は世論調査の方法の共通点と相違点についてワークシートにまとめ考察する。その後、教師が考察した共通点と相違点の全

体で共有し、標本調査の学習内容を踏まえながら世論調査の構造やバイアスの意味について解説を行い理解させる。

【バイアスが生じる要因に関する分類表をもとに、世論調査に含まれるバイアスを予想する】

次に、生徒は、上記で考察した共通点・相違点とバイアスの生じる要因の分類表を比較しながら、関連すると考えられるバイアスの要因の対応づけを行う。そして、教師は関連づけた事柄とバイアスが生じる要因を全体で共有する。例えば、生徒は「調査日が平日か土日か」という相違点や報道各社による質問の仕方の差異（例：重ね聞きの有無）、標本に占める固定電話と携帯電話の割合の差などに着目しながら、生じうるバイアスについて検討することが期待される。最後にこれまでの活動を振り返り、世論調査とバイアスの関係について、以下の点を踏まえてながらまとめて授業を終える。

〔まとめ〕

- 世論調査は、標本調査の手法を用いて実施することで、短期間で有権者の意見を集約することができる。
- 調査日や調査方法、調査票の構成の仕方などが各社で異なるため、結果にはバイアスがかかっている。
- 標本調査の結果にはバイアスが含まれていることを前提に、統計情報を批判的に読み解く能力（統計リテラシー）を持つことが大切。

4. 開発した教材の教育的価値

本研究で開発した教材の教育的価値は次の二点である。第一に、世論調査を通して、社会で用いられている標本調査の仕組みの理解を深めることができる点である。これは、標本や母集団の概念や無作為抽出の方法に関して、実際の内閣支持率の世論調査に基づいて具体化しながら学習できるということである。また、調査方法に関する共通点・相違点とバイアスの生じる要因との関連づける活動を位置付け

たことで、生徒が学習した標本調査とバイアスの概念を用いてそれぞれの調査方法を批判的に考察することが容易になると考える。さらに、これらを通して、統計情報を的確に把握し活用する態度を育成することに繋がると考える。

第二に、世論調査が標本調査の仕組みで行われることを詳しく学習することで、世論調査が生み出す統計情報を、多様な文脈を考慮して捉え直すことができる。すなわち、内閣支持率という数値の生み出され方を批判的に検討することで、内閣支持率とは何か、世論とは何か、世論をいかに捉えるべきか、などと問い直す契機を統計教育の中で確保することができる。Weiland (2017) の表現を借りるならば、統計情報をレンズとして、世界を読解し、記述する（あるいは、し直す）ことにつながる。例えば佐藤 (2016) は、国民の「公的意見」を「輿論」(よろん)、「国民感情」を「世論」(せろん)と呼んで区別し、今日の世論調査が計測した結果を「日々の国民投票」とみなすことに注意を促す。本教材を活用する場合にも、「世論調査で測定されている内閣支持率は、本当に『世論』と呼んでいいものなのか」と発問することで、佐藤が提起するように、内閣支持率という統計情報がもたらす社会的政治的問題状況を認識し、世論(調査)の意味づけを再考することを促すことが期待できる。

5. おわりに

本研究の目的は、生徒の批判的統計リテラシーの育成を意図した教材を開発することであった。そのために、先行研究に基づいて教材開発の要件を設定し、世論調査に含まれるバイアスを考察できる教材を開発したとともに、教科横断的な視点で批判的統計リテラシーの育成に繋がると期待しうる教材であることを示した。

今後の課題は、授業実践による実証的な考察を通して本教材の有効性を検証することである。

注

1) 本研究では、アメリカの高校生対象の数学と統計学の教科書である、Hirsch et al. (2015) *Transition to College Mathematics and Statistics* で取り上げられている「sample selection bias」と「response bias」の概念に基づいて分類表を作成した(表2)。

2) 表3を確認。

付記

本研究の一部は、JSPS 科研費(課題番号: JP20J10388)の助成を受けて行われた。

引用参考文献

- Gal, I. (2002). Adults' statistical literacy: Meaning, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Hirsch, C. R. et al. (2015). *Transition to College Mathematics and Statistics: Student Edition*. McGraw-Hill Education.
- 岩本裕 (2018). 民意を測る世論調査 吉田徹 編「民意のはかり方—『世論調査×民主主義』を考える—」, pp.25-59, 法律文化社.
- 古賀峻也 (2019). 高校生の「読解の文脈」での統計的リテラシーにおける批判的思考の育成. 日本科学教育学会 科学教育研究, 43(2), 154-164.
- 文部科学省 (2018a). 中学校学習指導要領(平成29年度告示) 解説社会編, 東洋館出版社.
- 文部科学省 (2018b). 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.
- 佐藤卓巳 (2016). 世論調査の「よろん」とは?——世論観測から輿論 2.0 へ. 放送メディア研究, 13, 25-40.
- Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, 96, 33-47.
- Weiland, T. (2019). Critical Mathematics Education and Statistics Education: Possibilities for Transforming

表2 バイアスが生じる原因

標本選択によるバイアス	① 標本の大きさ：人は標本の大きさが大きな標本を信じる傾向がある。 😊 標本の大きさが大きければ信頼できる標本といえますか？ ② 回答が自発的なものである：調査に協力するかどうか（標本に入るかどうか）を回答者自身が決められると、その問題について強い意見を持った人達で標本が構成される。 😊 この調査に回答しない人（したくない人）はどのような人か？ ③ 集めやすい標本：すぐに収集可能な標本を用いると、母集団を代表しない考え（意見）をもつ標本となってしまう。 😊 ○○生に聞いた結果は、△△県の中学生に聞いた結果と同じ？ ④ 誤った母集団：標本が関心のある集団から選択されたものではない。 😊 どのような母集団を想定しているか？
	① 無回答（の割合）：標本として選ばれた人に連絡がつかなかったり、その人が回答を拒否したりする。 😊 無回答の割合が増えるとうなる？ ② 質問紙（の内容）：質問に、ある特定の答えに誘導するような文言や、回答する人が誤解してしまうような文言が入っている。 😊 質問文や選択肢はどうなっていますか？ ③ 不正確な回答：正確に覚えていなかったり、調査員に自分の本音を知られたくなかったりする理由で、回答者が誤った回答をする。 😊 ホンネを言いたくないことってありませんか？ ④ タイミングの悪さ：調査を行う時期が調査結果に影響を及ぼす。 😊 最初はみんな期待するよね？ ⑤ 測定誤差：結果を集めたり分析したりする時に生じるミス。例えば、調査員が電話での回答を誤って記録したり、アンケート用紙に誤った回答を記入したりしてしまっていること。 😊 集計や処理方法は適切かな？

表3 報道各社による世論調査に関する一覧表

No.	調査主体	内閣支持率	調査日	調査方法	ワーディング
1	読売新聞 NNN（日本テレビ）	56%	10月4日(月) ～5日(火)	コンピューターで無作為に作成した固定電話と携帯電話の番号にかけたRDD（Random Digit Dialing）方式で18歳以上の有権者を対象に実施。固定では有権者住が判明した787世帯の中から443人、携帯では応答のあった1618人の中から681人、計1124人の回答を得た。回答率は固定56%、携帯42%。	あなたは、岸田内閣を、支持しますか、支持しませんか。 支持する 56 支持しない 27 その他 2 答えない 16 小数点以下四捨五入。
2	共同通信	55.7%	10月4日(月) ～5日(火)	全国の有権者を対象に4日午後から5日にかけて、コンピューターで無作為に発生させた番号に電話をかけるRDD（ランダム・デジッ・ダイアリング）法で実施した。固定電話は、福島県の一部地域を調査対象から除いた。実際に有権者がいる世帯にかかったのは665件、うち535人から回答を得た。携帯電話は、電話がかかったのは1687件、うち552人から回答を得た。	あなたは、今回発足した岸田内閣を支持しますか、支持しませんか。 支持する 55.7 支持しない 23.7
3	毎日新聞 社会調査研究センター	49%	10月4日(月) ～5日(火)	携帯電話のショートメッセージサービス（SMS）機能を使う方式と、固定電話で自動音声の質問に答えてもらう方式を組み合わせ、携帯725件・固定310件の有効回答を得た。	岸田内閣を支持しますか。 支持する 49 支持しない 40 答えない 11
4	朝日新聞	45%	10月4日(月) ～5日(火)	コンピューターで無作為に電話番号を作成し、固定電話と携帯電話に調査員が電話をかけるRDD方式で、4日夕から5日夜にかけて全国の有権者を対象に調査した。固定は有権者がいると判明した684世帯から375人（回答率55%）、携帯は有権者につながった1330件のうち597人（同45%）、計972人の有効回答を得た。	今度できた岸田文雄内閣についていかがですか。あなたは、岸田内閣を支持しますか。支持しませんか。 支持する 45 支持しない 20 その他・答えない 35 数字は%。小数点以下は四捨五入。
5	日本経済新聞 テレビ東京	59%	10月4日(月) ～5日(火)	日経リサーチが全国の18歳以上の男女に携帯電話も含めて乱数番号（RDD方式）による電話で実施した。回収数854、回答率43.1%。有権者判明件数1982。	あなたは岸田内閣を支持しますか、しませんか。 支持する 49 支持しない 21 いえない・わからない 30 （「いえない・わからない」と回答した方に）お気持ちに近いのはどちらですか。 支持する 35 支持しない 12 いえない・わからない 53 ※結果、「支持する」59、「支持しない」25、「いえない・わからない」16
6	NHK	49%	10月8日(金) ～10日(日)	全国の18歳以上を対象にコンピューターで無作為に発生させた固定電話と携帯電話の番号に電話をかける「RDD」という方法で行った。調査の対象となったのは3328人で、58%にあたる1921人から回答を得た。	あなたは、岸田内閣を支持しますか。それとも支持しませんか。 支持する 49% 支持しない 24%
7	JNN（TBS）	58.6%	10月9日(土) ～10日(日)	（調査方法） 固定・携帯電話による聞き取り（RDD方式） 全国18歳以上の男女 2348人（固定1002人、携帯1346人） 有効回答 1225人(52.2%)（固定612人(61.1%)、携帯613人(45.5%)）	岸田内閣についてどう思いますか。次の4つの中から1つだけ選んで下さい。 非常に支持できる 5.0% ある程度支持できる 53.6% あまり支持できない 29.0% まったく支持できない 5.8% 答えない・わからない 6.6% （小数点2位以下は四捨五入）
8	時事通信	40.3%	10月8日(金) ～11日(月)	全国の18歳以上の男女2000人を対象に個別面接方式で実施。有効回収率は64.0%。	支持 40.3% 不支持 18.8% 分らない 40.8%
9	産経新聞 FNS	63.2%	10月9日(土) ～10日(日)	調査エリアごとの性別・年齢構成に合わせ、電話番号を無作為に発生させるRDD（ランダム・デジッ・ダイアリング）方式で電話をかけ、算出した回答数が得られるまで調査を行った。電話の割合は「固定電話4：携帯電話6」。内閣支持率のみ回答が不明確な場合には、「どちらかといえば」と再度質問して回答を得た。調査対象は全国の18歳以上の男女1016人。	岸田文雄内閣を支持するか 支持する 63.2% 支持しない 27.4% 他 9.3%
10	ANN（テレビ朝日）	43.4%	10月16日(土) ～17日(日)	【調査日】 2021年10月16・17日（土・日曜日） 【調査方法】 電話調査（RDD方式） 【対象】 全国18歳以上の男女1960人 【有効回答率】 53.6% 回答者 1051人	あなたは、岸田文雄連立内閣を支持しますか、支持しませんか？ 支持する 43.4% 支持しない 28.3% わからない、答えない 28.3%

パスカルの三角形の回帰分析

高木和久・高知工業高等専門学校

高知県南国市物部 ktakagi@ge.kochi-ct.ac.jp

1. はじめに

高等学校の学習指導要領が変わり、情報Ⅰでは単回帰分析、情報Ⅱでは重回帰分析を取り扱うことになった。単回帰分析は通常は偏微分を用いて回帰直線を決定する。しかし、第2章に述べるように回帰直線が点 $(\bar{x}, \bar{y})$ を通ることを仮定すれば常微分のみを用いて回帰直線を決定することができる。

私の勤務校は工業高等専門学校であり、第2学年までに1変数の微分積分の大半を履修する。データサイエンスの授業は3年次にあるので回帰分析について授業で取り扱うときは第2章に述べる内容を学生に示して理論的に回帰直線の方程式を導出する。社会における回帰分析の利用例としてはスマート農業の事例を取り上げている¹⁾。これは衛星画像を元に正規化植生指数(NDVI)を算出し、小麦の穂の水分含有率を目的変数、NDVIを説明変数として回帰分析を行うものである。

説明変数の数が増えるだけなので、3年次のデータサイエンスの授業では重回帰分析も取り扱うが、その理論について、数学を用いて解説することはない。なお、私が担当する専攻科1年(大学3年相当)の必修科目では第3章に述べる内容で重回帰分析の原理の説明を行っている。

2. 回帰直線の方程式の導出

n 個のデータ (x_i, y_i) ($i = 1, 2, \dots, n$)に対し

$$\bar{x} = \frac{\sum x_i}{n}, \bar{y} = \frac{\sum y_i}{n}, S_{xx} = \frac{\sum (x_i - \bar{x})^2}{n}$$

$$S_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n}, S_{yy} = \frac{\sum (y_i - \bar{y})^2}{n}$$

と定める。ここでは $S_{xx} > 0$ とする。

さて、点 $(\bar{x}, \bar{y})$ を通り傾き m の直線の方程式は

$Y - \bar{y} = m(X - \bar{x})$ である。 (X, Y) に (x_i, y_i) を代入したとき、必ずしも $y_i - \bar{y} = m(x_i - \bar{x})$ が成り立つとは限らない。

$$f(m) = \sum \{y_i - \bar{y} - m(x_i - \bar{x})\}^2$$

とおき、 $f(m)$ が最小になるときの m の値を求める。

式を変形して

$$f(m) = \sum \{(y_i - \bar{y})^2 - 2(x_i - \bar{x})(y_i - \bar{y})m + (x_i - \bar{x})^2 m^2\}$$

$$\begin{aligned} \frac{f(m)}{n} &= \frac{1}{n} \sum (y_i - \bar{y})^2 - \frac{2}{n} \sum (x_i - \bar{x})(y_i - \bar{y})m + \frac{1}{n} \sum (x_i - \bar{x})^2 m^2 \\ &= S_{yy} - 2S_{xy}m + S_{xx}m^2 \end{aligned}$$

$$\therefore f(m) = n(S_{xx}m^2 - 2S_{xy}m + S_{yy})$$

$$f'(m) = 2n(S_{xx}m - S_{xy}), f''(m) = 2nS_{xx} > 0$$

$f'(m) = 0$ とおくと

$$-2S_{xy} + 2S_{xx}m = 0 \quad \therefore m = \frac{S_{xy}}{S_{xx}}$$

$f''(m)$ は常に正であるから、 $m = \frac{S_{xy}}{S_{xx}}$ において $f(m)$

は最小値をとる。

傾き $m = \frac{S_{xy}}{S_{xx}}$ を持つ直線 $Y - \bar{y} = \frac{S_{xy}}{S_{xx}}(X - \bar{x})$ を

y の x に対する回帰直線と呼ぶ。さて、 u_i, v_i を

$$u_i = \frac{x_i - \bar{x}}{\sqrt{S_{xx}}}, v_i = \frac{y_i - \bar{y}}{\sqrt{S_{yy}}}$$

と定める。 u_i, v_i はそれぞれ x_i, y_i を標準化した変数である。そして u_i と v_i の共分散 S_{uv} は

$$S_{uv} = \frac{\sum (u_i - \bar{u})(v_i - \bar{v})}{n}$$

である。ここで

$$\bar{u}_i = \frac{(\bar{x}_i - \bar{x})}{\sqrt{S_{xx}}} = \frac{\bar{x}_i - \bar{x}}{\sqrt{S_{xx}}} = 0, \bar{v}_i = \frac{\bar{y}_i - \bar{y}}{\sqrt{S_{yy}}} = 0$$

であるから、

$$\begin{aligned} S_{uv} &= \frac{\sum u_i v_i}{n} = \frac{\sum \frac{(x_i - \bar{x})}{\sqrt{S_{xx}}} \frac{(y_i - \bar{y})}{\sqrt{S_{yy}}}}{n} \\ &= \frac{1}{\sqrt{S_{xx}}\sqrt{S_{yy}}} \cdot \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}} \end{aligned}$$

となる。この値を x_i と y_i の相関係数と呼ぶ。

$$\text{また、} R^2 = \frac{\sum \{m(x_i - \bar{x})\}^2}{\sum (y_i - \bar{y})^2} \text{ を決定係数と呼ぶ。}$$

式を変形すると

$$R^2 = \frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2} m^2 = \frac{S_{xx}}{S_{yy}} \left(\frac{S_{xy}}{S_{xx}} \right)^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}}$$

右辺は x と y の相関係数の2乗に等しい。

3. 重回帰分析の理論

$$u_0, y, \vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}, \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \text{ を変数、}$$

$y_1, \dots, y_n, x_{k1}, \dots, x_{kn} (1 \leq k \leq m)$ をデータとする。

このとき、

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ki}$$

$$S_{yk} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})(x_{ki} - \bar{x}_k)$$

$$= \left(\sum_{i=1}^n y_i x_{ki} \right) - \bar{y} \cdot \bar{x}_k$$

$$S_{kl} = \frac{1}{n} \sum_{i=1}^n (x_{ki} - \bar{x}_k)(x_{li} - \bar{x}_l)$$

$$= \frac{1}{n} \left(\sum_{i=1}^n x_{ki} x_{li} \right) - \bar{x}_k \cdot \bar{x}_l$$

と定める。 y を $u_0 + \sum_{l=1}^m u_l x_l$ で近似する。

$$f(y, \vec{x}) = \left(y - u_0 - \sum_{l=1}^m u_l x_l \right)^2$$

とおくと

$$\frac{\partial f}{\partial u_0} = -2 \left(y - u_0 - \sum_{l=1}^m u_l x_l \right)$$

$$\frac{\partial f}{\partial u_k} = -2 x_k \left(y - u_0 - \sum_{l=1}^m u_l x_l \right)$$

$$= -2 \left(y x_k - u_0 x_k - \sum_{l=1}^m u_l x_l x_k \right)$$

が成り立つ。次に y と $\vec{x}$ にデータを代入して平均を取った値 Q :

$$Q = \frac{1}{n} \sum_{i=1}^n f(y_i, \vec{x}_i)$$

について Q を最小にする $u_0, \vec{u}$ の組み合わせを求める。

極値では

$$\frac{\partial Q}{\partial u_0} = 0, \quad \frac{\partial Q}{\partial u_k} = 0 \quad (1 \leq k \leq m)$$

が成り立つ。まず $\frac{\partial Q}{\partial u_0} = 0$ より

$$\frac{1}{n} \sum_{i=1}^n \left(y_i - u_0 - \sum_{l=1}^m u_l x_{li} \right) = 0$$

$$\frac{1}{n} \sum_{i=1}^n y_i - u_0 - \sum_{l=1}^m u_l \left(\frac{1}{n} \sum_{i=1}^n x_{li} \right) = 0$$

$$\bar{y} - u_0 - \sum_{l=1}^m u_l \bar{x}_l = 0$$

$$\therefore u_0 = \bar{y} - \sum_{l=1}^m u_l \bar{x}_l \quad \dots (\star)$$

次に $\frac{\partial Q}{\partial u_k} = 0$ より

$$\frac{1}{n} \sum_{i=1}^n \left(y_i x_{ki} - u_0 x_{ki} - \sum_{l=1}^m u_l x_{li} x_{ki} \right) = 0$$

$$\frac{1}{n} \sum_{i=1}^n y_i x_{ki} - u_0 \cdot \frac{1}{n} \sum_{i=1}^n x_{ki} - \sum_{l=1}^m u_l \left(\frac{1}{n} \sum_{i=1}^n x_{li} x_{ki} \right) = 0$$

$$\frac{1}{n} \sum_{i=1}^n y_i x_{ki} - u_0 \bar{x}_k - \sum_{l=1}^m u_l \left(\frac{1}{n} \sum_{i=1}^n x_{li} x_{ki} \right) = 0$$

($\star$)を代入して

$$\frac{1}{n} \sum_{i=1}^n y_i x_{ki} - \left(\bar{y} - \sum_{l=1}^m u_l \bar{x}_l \right) \bar{x}_k - \sum_{l=1}^m u_l \left(\frac{1}{n} \sum_{i=1}^n x_{li} x_{ki} \right)$$

$$= 0$$

$$\frac{1}{n} \sum_{i=1}^n y_i x_{ki} - \bar{y} \cdot \bar{x}_k = \sum_{l=1}^m u_l \left(\frac{1}{n} \sum_{i=1}^n x_{li} x_{ki} - \bar{x}_l \cdot \bar{x}_k \right)$$

$$\therefore S_{yk} = \sum_{l=1}^m u_l S_{kl}$$

$1 \leq k \leq m$ の範囲の全ての k についてこの式が成り立つので、行列を用いて表すと

$$\begin{pmatrix} S_{11} & \dots & S_{1m} \\ \vdots & \ddots & \vdots \\ S_{m1} & \dots & S_{mm} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} S_{y1} \\ \vdots \\ S_{ym} \end{pmatrix}$$

となる。

4. パスカルの三角形の回帰分析

データサイエンスを学校で教えるに当たって、どのようなデータをどのようにして得るか、が重要である。本校は工業高専であるので専門学科の先生にお願いして、この先生が大学院生の時代にとった脳波のデータを利用して貰った。マハラノビス汎距離を計算して線形判別関数を自作し、被験者が目を閉じているか、それとも目を開けているかを脳波のデータから判別する事例を作って授業で紹介した。

しかし学生にデータを渡して加工させることは難易度が高く、実現できなかった。3年生（高等学校3年生に相当）の全員が取り組む課題としては、もっと簡単に入手できるデータで、しかも前処理を必要としないものが必要である。

図1はパスカルの三角形と呼ばれるもので、高校生にはおなじみのものである。

				1					
				1	1				
			1	2	1				
		1	3	3	1				
	1	4	6	4	1				
1	5	10	10	5	1				
1	6	15	20	15	6	1			
1	7	21	35	35	21	7	1		
1	8	28	56	70	56	28	8	1	

図1. パスカルの三角形

この三角形を斜めにデータが並んでいるものとみなす。1,2,3,4, … を a_n 、1,3,6,10, … を b_n 、1,4,10,20, … を c_n 、1,5,15,35, … を d_n とおく。明らかに、 a_n の一般項は n である。

そして、 $n = 1, 2, 3, 4$ のデータを用いて、 a_n を説明変数 x 、 b_n を目的変数 y として回帰分析を行った。得られた回帰直線の方程式は $y = 3.5x - 3.5$ であった。 b_n は常に整数値をとる数列であるから、この式は明らかに b_n の一般項を表していない。そこで、 $\frac{b_n}{a_n}$ を目的変数 y 、 a_n を説明変数 x として回帰分析をさせた。

$$\bar{x} = \frac{5}{2}, \bar{y} = \frac{7}{4}, S_{xx} = \frac{5}{4}, S_{xy} = \frac{5}{8}$$

であるので、 y の x への回帰直線の方程式は

$$y = \frac{1}{2}x + \frac{1}{2} \text{ となり、}$$

$$\therefore \frac{b_n}{a_n} = \frac{a_{n+1}}{2} \quad \therefore b_n = \frac{a_n(a_{n+1})}{2}$$

$a_n = n$ であったから、 $b_n = \frac{n(n+1)}{2}$ となり、この式は b_n の一般項を正しく表している。

なお、 b_n の一般項 $\frac{n(n+1)}{2}$ は 1 から n までの整数の和と一致する。これは $\sum_{k=1}^n a_k = b_n$ が成り立っているからである。

次に、 a_n を説明変数 x 、 $\frac{c_n}{b_n}$ を目的変数 y として回帰分析を行った。

$$\bar{x} = \frac{5}{2}, \bar{y} = \frac{3}{2}, S_{xx} = \frac{5}{12}, S_{xy} = \frac{5}{4}$$

であるので、 y の x への回帰直線の方程式は $y = \frac{1}{3}x + \frac{2}{3}$ となり、

$$\therefore \frac{c_n}{b_n} = \frac{x+2}{3} \quad \therefore c_n = b_n \cdot \frac{a_{n+2}}{3} = b_n \cdot \frac{n+2}{3}$$

この式から c_n を n の式で表すと

$$c_n = \frac{n(n+1)}{2} \cdot \frac{n+2}{3} = \frac{n(n+1)(n+2)}{6}$$

となる。この式は c_n の一般項を正しく与えている。

同様に、 a_n を説明変数 x 、 $\frac{d_n}{c_n}$ を目的変数 y として回帰分析を行うと、回帰直線の方程式は $y = \frac{1}{4}x + \frac{3}{4}$ となり、

$$d_n = c_n \cdot \frac{a_{n+3}}{4} = c_n \cdot \frac{n+3}{4} = \frac{n(n+1)(n+2)(n+3)}{24}$$

となる。この式は d_n の一般項を正しく与えている。

5. 回帰分析を利用した部分分数分解

回帰分析を用いて部分分数分解を行うことができる。次の問題は静岡理工科大学の 1997 年の入試で出題された問題である。

問
$$\frac{1}{(x+1)(x+2)(x+3)} = \frac{a}{x+1} + \frac{b}{x+2} + \frac{c}{x+3}$$

が x についての恒等式であるとき定数 a, b, c の値を求めよ。

分析は、R を用いる場合は次のように入力する。

```
x<-1:5
p<-1/(x+1)
q<-1/(x+2)
r<-1/(x+3)
s<-p*q*r
summary(lm(s~p+q+r))
```

出力結果は次のようになる。

```
Coefficients:
              Estimate
(Intercept) -4.965e-17
p             5.000e-01
q            -1.000e+00
r             5.000e-01
```

R の出した予測式は $s = 0.5p - q + 0.5r$ であり、

$$\frac{1}{(x+1)(x+2)(x+3)} = \frac{1}{2} \left(\frac{1}{x+1} - \frac{2}{x+2} + \frac{1}{x+3} \right)$$

のように部分分数に分解されることがわかる。

よって $a = \frac{1}{2}, b = -1, c = \frac{1}{2}$ である。

$$p_k = \frac{1}{x+k}, P_k = \prod_{i=1}^k p_i$$

とおくと P_n は $p_1, \dots, p_n$ の線形結合で表される。

回帰分析を用いて $p_1, \dots, p_n$ の係数を求めよう。

まず、 $x_1 = p_1, x_2 = p_2, x_3 = p_3, y = P_3$ として重回帰分析を行う。 x に 1, 2, 3, 4, 5 を代入すると

$$x_1 = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}$$

$$x_2 = \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}$$

$$x_3 = \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}$$

$$y = \frac{1}{24}, \frac{1}{60}, \frac{1}{120}, \frac{1}{210}, \frac{1}{336}$$

というデータが得られる。

目的変数を $y = P_3$ 、説明変数を $x_k = p_k$ ($1 \leq k \leq 3$)

として回帰分析を行うと

$$y = \frac{x_1}{2} - x_2 + \frac{x_3}{2}$$

が得られる。これより

$$P_3 = \frac{1}{2}(p_1 - 2p_2 + p_3)$$

が成り立つことが予想される。

目的変数を $y = P_4$ 、説明変数を $x_k = p_k$ ($1 \leq k \leq 4$)

として回帰分析を行うと

$$y = \frac{x_1}{6} - \frac{x_2}{2} + \frac{x_3}{2} - \frac{x_4}{6}$$

が得られる。これより

$$P_4 = \frac{1}{6}(p_1 - 3p_2 + 3p_3 - p_4)$$

が成り立つと予想される。一般に、 P_n を $p_1, \dots, p_n$ の線

形結合で表したときの係数の比は図 2 のようになる。

					1	
				1	-1	
			1	-2	1	
		1	-3	3	-1	
	1	-4	6	-4	1	
1	-5	10	-10	5	-1	
1	-6	15	-20	15	-6	1

図 2. 部分分数分解式中の $p_1, \dots, p_n$ の係数の比

6. 終わりに

P_n を $p_1, \dots, p_n$ の線形結合で表したときの係数の比を回帰分析を用いて生徒に予想させ、可能なら証明を考えさせることは探究活動のテーマとして成立すると思われる。これからも実践を行って、改良を重ねてゆきたい。

参考文献

(1) 国立研究開発法人 農業・食品産業技術総合研究機構、衛星画像を利用した小麦収穫作業システムは主要秋播き品種への適用が可能

<https://www.naro.go.jp/project/results/laboratory/harc/2010/cryo10-16.html> (2022.2.25 確認)

(2) 高木和久、パスカルの三角形と部分分数分解、高専教育フォーラム No.2 PP.2-3、大日本図書、2021

中高生・スポーツデータ解析コンペティション1

- ▶ 大学生向けとして、日本統計学会スポーツ統計分科会（現：スポーツデータサイエンス分科会）と情報・システム研究機構統計数理研究所が主催
- ▶ 第3回から中等教育向けの部門を設置
- ▶ 2020年度（第10回）から独立して、名称変更改め
- ▶ 本コンペティションでは、野球・サッカー・バスケットボール・卓球など日本プロスポーツの実際のデータを参加チームに貸与し、そのデータの分析力を競うコンペティション
- ▶ 毎年、約 50-60 チームが参加し、日本統計学会会長・理事長をはじめ、統計学者、数学教育者、スポーツデータ関係会社の方々に審査依頼

▶ 2

JCOTS2022

中高生・スポーツデータ解析コンペティション2

▶ 2つの部門

- ▶ **規定部門**…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のデータなどを貸与。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う。貸与データが主な分析対象とするなら、補助的な追加データ OK
- ▶ **自由部門**…各自でスポーツに関するデータを収集し、そのデータで規定部門と同様に分析等を行う
- ▶ 2021 年度は「規定：自由 = 7 : 8」
- ▶ 規定部門のデータも外部利用申請で活動成果を VVK 研究会や統計グラフィコンクールなどに発表することも可能
- ▶ **参加費は無料**
- ▶ 感謝：統計数理研究所、日本統計学会、など

▶ 3

JCOTS2022

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。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・3月：「統計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。

▶ 4

JCOTS2022

審査基準など

- 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。
- データ活用力：
 - ▶ 賞与または持参のデータを効果的・適切に活用している。
- 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。

そのほかの情報についてはサイトでご確認を

申し込みページ

中高生・スポーツデータ解析コンペティションWebサービス

お知らせお問い合わせポータル

中高生・スポーツデータ解析コンペティション

出場チームのエントリーは締め切りました

コンペティションの詳細はコチラをクリック

発会式及び申込み・誓約書の提出などをオンラインで完結するように変更しました

お知らせ

▶ 2021年8月6日(金) 中高生・スポーツデータ解析コンペティション-2021-の開催

申し込み締切

エントリーを締め切りました

概要

▶ 一覧へ

出場チームの方

発表タイトル登録などはこちらから

https://hs.sports.ywebsys.net/

結果のページ

日本統計学会統計教育分科会のトップページ
(<https://estat.scikagoshima-u.ac.jp/SESJSS/>)

■ 2021-02-28

中高生・スポーツデータ解析コンペティション-2020-の結果が公開されました (リンク)

■ 2020-12-30

第 18 回 統計教育の方法論

中高生・スポーツデータ解析コンペティション-2020-の結果

今回も 60 に近い出品があり、日本統計学会会長・理事長をはじめ、統計学者、統計教育者、数学教育者、スポーツデータ関連識者、など、多くの先生方によって厳正に審選し、以下のように選別いたします。

総評： 今年度は COVID-19 感染防止のため、グループワークがしづいなが、今年も 60 に近い興味深い作品が提出されました。ありがたうでしたが、「規定部門」は例年通り、また例年の作品を上回る素晴らしい作品が多数応募されました。「自由部門」ではバドミントンやレーカーなど、賞与データがない種目の作品もあり、大変興味深い作品が多くありました。個々の作品では、自ら問いを設定（リサーチクエスション）し、必要データを処理し、分析して課題・問題を解決し、それらを他者に上手に伝えるという、分析のストーリーができており、興味深かったです。R や Python を用いた多変量解析が使われている作品も多くあつて、分析手法も効率的に使われていました。賞与データでもデータの前提処理が必要な種目もありましたが、適切に前提処理ができており、こちらも驚きました。新しい学部転身者が中学校で 2021 年度から、高校で 2022 年度から始まりますが、賞与・情報の授業でデータサイエンスに関する内容を履修していきたいと思います。探究学習をさらに進めるかもしれません。これらで得た知識や技能を用いて、ぜひ 2021 年度も興味深い作品に取り組んでいただければと思います。

規定部門			
賞	作品タイトル	受賞者および学校名	作品 (クリックすると PDF が見れます)
最優秀賞	サッカーのシュートは本当にペナルティエリア外からでも精確的に打つべきなのか？	大西諒・福岡健雄・大久保成博・児玉英樹 (奈良高等学校)	
	卓球選手の強さを表す指標の作成	石川樹大・川崎希治・藤村小樹・石井裕基 (静岡県立藤宮高等学校)	
優秀賞	サッカーにおけるクロスからの得点量	大西諒・福岡健雄・大久保成博・児玉英樹 (奈良高等学校)	

まとめ

- ▶ 2022 年度も今年度と同様の予定
 - ▶ ルール変更も可能性もあるため、参加の際に注意
 - ▶ 自由部門は途中辞退を許可？
 - ▶ 部門を増やす可能性も…
 - ▶ コンペ終了時に全チームにコメント
- ▶ 規定部門のデータでも外部利用申請で別の研究発表
 - ▶ 本コンペでの利用が原則
 - ▶ 校内外の研究会での発表も可
 - ▶ SSH 研究会、統計グラフコンクール、ISLP 国際統計ポスターコンペティション、…

データサイエンティスト検定™ リテラシーレベル

“データの力を解き放つ”基礎を問う試験の目的と今後の展望について

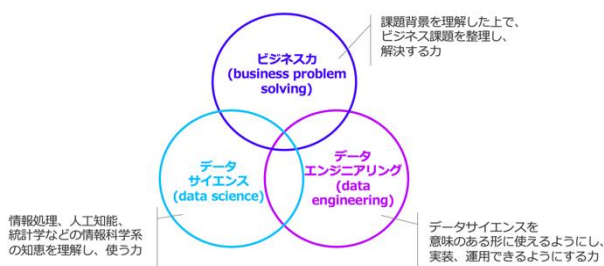
データサイエンティスト協会 スキル定義委員
株式会社 Rejoui 代表取締役/広島大学 客員教授
菅 由紀子 / yukiko@rejoui.co.jp

データサイエンティスト協会設立の背景と目的

データサイエンティスト協会は、社会におけるビッグデータ活用の潮流に伴い、重要視されるデータサイエンティストの育成のため、そのスキル要件の定義・標準化を推進し、社会に対する普及啓蒙活動を行うことを主な目的として 2013 年 5 月に設立されました。国内におけるあらゆる領域でのデータサイエンスの活用や高度専門人材の育成、データ分析業界の健全な発展に貢献するための活動を行っています。

特に設立された 2013 年頃は「データサイエンティスト」という言葉の定義が欠落しており、企業の雇用側の期待に雇用される側のスキルセットが合致しないケースが増加、データ分析業界では大きな課題となっていました。また、若い才能たちが、自分達をどう訓練し、スキルを身に付けていくべきかが分からない、という状況もありました。これらを改善するため、この新しいデータプロフェッショナルに必要とされるスキルセットを定義し、スキル育成と評価のための軸・基準を作る活動を行って来しました。

データサイエンティストに必要な 3 つのスキルセット



データサイエンティスト検定™リテラシーレベル開始の背景

設立当初からデータサイエンティスト協会に対し、データサイエンティストの能力を測る資格試験の要望は多く届いていました。しかし「データから価値を創出し、ビジネス課題に答えを出すプロフェッショナル」であるデータサイエンティストの力量を測ることは、この領域の技術はすぐに陳腐化する進化が加速度的な領域であるため、その設置に対しては幾度となく深い議論が繰り返されてきました。

しかし、大学でのデータサイエンス必修化の流れ・アカデミ

アからの要望、産業界で DS が活躍する場の広がり & 役割の広がり、データサイエンティストに求められるスキル定義の1層の明確化とその普及の必要性から 2020 年始より視覚化の検討が行われ、2021 年 9 月に試験が開始されました。

本試験の対象と目的

本試験は、データサイエンティストスキルチェックリストに定めるアシスタント・データサイエンティスト（見習いレベル：★）のスキル、および数理・データサイエンス教育強化拠点コンソーシアムが公開している数理・データサイエンス・AI（リテラシーレベル）におけるモデルカリキュラムを総合し、実務能力と知識を有するかを問うことを目的としています。そのため、データサイエンティスト初学者やこれから目指すビジネスパーソン・データサイエンティストに興味を持つ学生を主な対象としています。昨年の 9 月の実施では約 1400 人が受験し、約 900 名が合格しました。第二回の試験は 2022 年 6 月、第三回を予定しており、第二回・第三回はスキルチェックリスト Ver. 4 に基づき試験問題を刷新予定です。また、今後の展望としてデータサイエンティスト協会スキル定義委員では独り立ちレベル：★★の資格についても検討を開始しております。

本発表においては、試験の出題対象範囲や各スキルカテゴリにおいて問われる能力などについて具体的に紹介いたします。

人材育成における JDLA の取り組み

岡田隆太郎

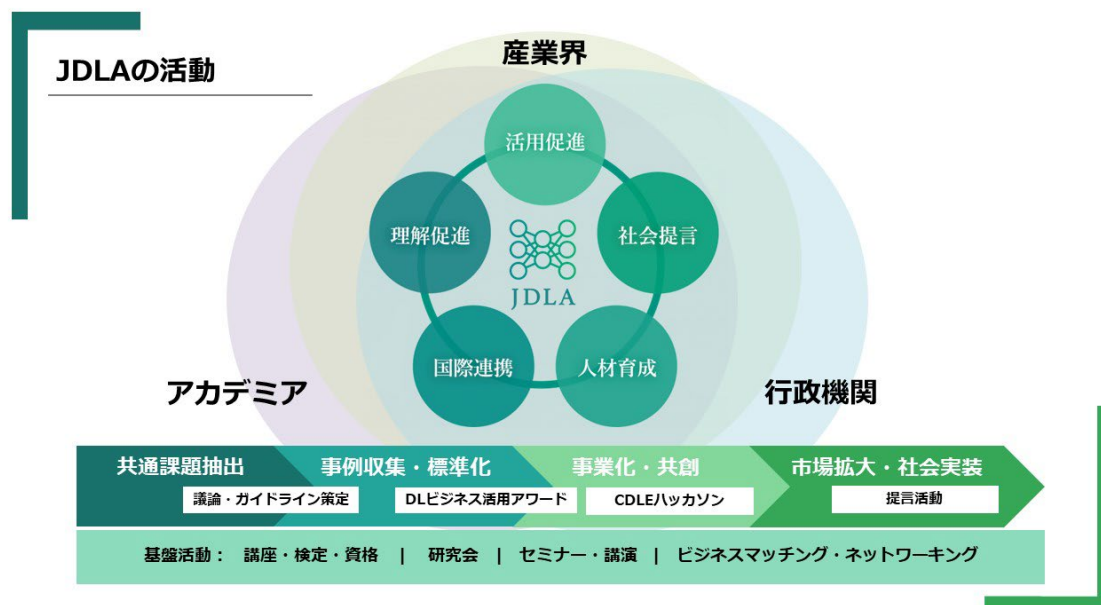
日本ディープラーニング協会 理事/事務局長

東京都港区芝公園 1 - 1 - 1

03-6820-2096・info@jdla.org

1. JDLA とは

日本ディープラーニング協会は、ディープラーニングを事業の核とする企業が中心となり、ディープラーニング技術を日本の産業競争力につなげていこうという意図のもとに設立されたものです。産業活用促進、人材育成、公的機関や産業への提言、国際連携、社会との対話など、産業の健全な発展のために必要な活動を行っています。



2. JDLA の人材育成

2012 年ごろから急速に進展してきたディープラーニング技術は、画像認識や音声認識等で大きな性能の飛躍を遂げています。機械にさまざまな状況が「見える」ようになるため、ものづくりとも相性がよく、また、さまざまな労働力不足の問題を解決する可能性があるなど、日本全体にとって大きなチャンスであると考えます。諸外国のディープラーニングに関する取り組みは急速に進んでいます。日本がグローバルに戦える体制を整えていくには、早期に「ディープラーニング産業」を拡大する取り組みが必要と考えます。

そのために、まず重要なのは人材育成です。ディープラーニング技術を担う人材、また、ディープラーニングの可能性と限界を正しく理解し、うまく事業に活用する人材の両方が必要です。ディープラーニングは決して万能なわけではありません。得意なデータや不得意なデータがありますし、ディープラーニング以外のさまざまな機械学習技術を使うほうが有効な場合も多くあります。人工知能、機械学習、ディープラーニングに関する俯瞰的な理解をもった人材を、特に、ディープラーニングを活用したいと考えるユーザ企業のな

かで増やすことが、産業全体の進展にとって大変重要だと考えます。同時に、ディープラーニングを用いたシステムの開発をできる人材も大きく不足しています。さまざまな領域のエンジニアや研究者、学生などに技術を習得してもらえればと思っています。

現在 JDLA ではディープラーニングに関する知識を有し、事業活用する人材（ジェネラリスト）と、ディープラーニングを実装する人材（エンジニア）の育成を目的に各々に必要な知識やスキルセットを定義し、資格試験を行うとともに、協会が認定した事業者がトレーニングを提供しています。

また、2021 年 5 月より全てのビジネスパーソンに向けた、AI／ディープラーニングについてまず「知る」ための無料エントリー講座の提供も開始しています。

JDLA の人材育成（試験）事業

エントリー向けのオンライン講座	デジタル社会を生きる ビジネスパーソン必携リテラシー	エンジニアをはじめデジタル社会の “創る”職種における登竜門
 AI4E AI for Everyone すべての人のためのAIリテラシー講座 （想定必要学習時間：約6時間程度） ディープラーニングについてまず「知る」ためのビジネスパーソン向け無料エントリー講座。 ディープラーニングを含めた「データ×AI」の理解・活用のための基礎的な概念やリテラシー、事例等を動画レッスンで学習できます。 受講者15,000人突破！	 G検定 JDLA Deep Learning for GENERAL （想定必要学習時間：約30時間程度） ディープラーニングを活用したプロジェクトに関わるすべての人（ジェネラリスト）向けの検定。 プロジェクトの検討・企画・推進のために必要な、実践要素を含むリテラシー習得に関する試験になっています。 合格者44,000人突破！	 E資格 JDLA Deep Learning for ENGINEER （想定必要学習時間：約120時間程度） ディープラーニングエンジニアをはじめ、実際にディープラーニングプロジェクトの開発・推進に関わる人のための資格。 アルゴリズムの構築や実装等開発・実装に関わる理論的な知識やスキルを習得する内容になっています。 合格者 約3,500人突破！

<G 検定>

ディープラーニングの基礎知識を有し、適切な活用方針を決定して、事業活用する能力や知識を有しているかを検定する、ジェネラリスト※向けの検定試験です。

プロジェクトの検討・企画・推進のために必要な、実践要素を含むリテラシー習得のための学習にも最適です。

※AI・ディープラーニングを活用したプロジェクトに関わるすべての人

G検定とは？

ジェネラリスト
ディープラーニングの基礎知識を有し適切な活用方針を決定して事業応用する能力を持つ人材



受験資格
制限なし

実施概要
試験時間：120分
知識問題（多肢選択式・220問程度）
オンライン実施（自宅受験）

受験費用
一般：12,000円（税別）
学生：5,000円（税別）
※2020年2月～一般：6,000円（税別）学生：2,500円（税別）

G検定の試験範囲

人工知能（AI）とは（人工知能の定義） 人工知能をめぐる動向 探索・推論、知識表現、機械学習、深層学習 人工知能分野の問題 トピック別・フレームワーク別、弱いAI、強いAI、身体性、シンボリック・ニューラルネットワーク、特許・知的財産、セキュリティ 機械学習の具体的な手法 代表的な手法（教師あり学習、教師なし学習、強化学習） データの扱い、評価指標 ディープラーニングの概要 ニューラルネットワークとディープラーニング、既存のニューラルネットワークにおける問題、ディープラーニングのアーキテクチャ、GPUとCPU、ディープラーニングのデータ量、活性化関数、学習率の最適化、変化するネットワーク	ディープラーニングの手法 CNN、深層生成モデル、画像認識分野での応用、音声処理と自然言語処理分野、RNN、深層強化学習、ロボティクス、マルチモーダル、モデルの解釈性とその対応 ディープラーニングの社会実装に向けて AIプロジェクトの計画、データ収集、加工・分析・学習、実装・運用・評価・法律（個人情報保護法・著作権法・不正競争防止法・特許法）、契約倫理、現行の議論（プライバシー、バイアス、透明性、アカウンタビリティ、ELSI、XAI、ディープフェイク、ディバーシティ）
---	--

<E 資格>

ディープラーニングの理論を理解し、適切な手法を選択して実装するための知識・能力を有しているかを認定する、エンジニア向けの資格試験です。

アルゴリズムの構築・実装など、開発や実装に関わる理論的な知識やスキルの習得に役立つ資格です。エンジニアだけではなく、ディープラーニングプロジェクトの開発・推進に関わる人の学習にも活用できます。

E資格とは？

エンジニア

ディープラーニングの理論を理解し
適切な手法を選択して
実装する能力を持つ人材

受験資格
JDOLA認定プログラムの修了

実施概要
試験時間：120分
知識問題（多肢選択式・100問程度）
会場試験（各地の指定試験会場）

受験費用
一般：30,000円（税別）
学生：20,000円（税別）

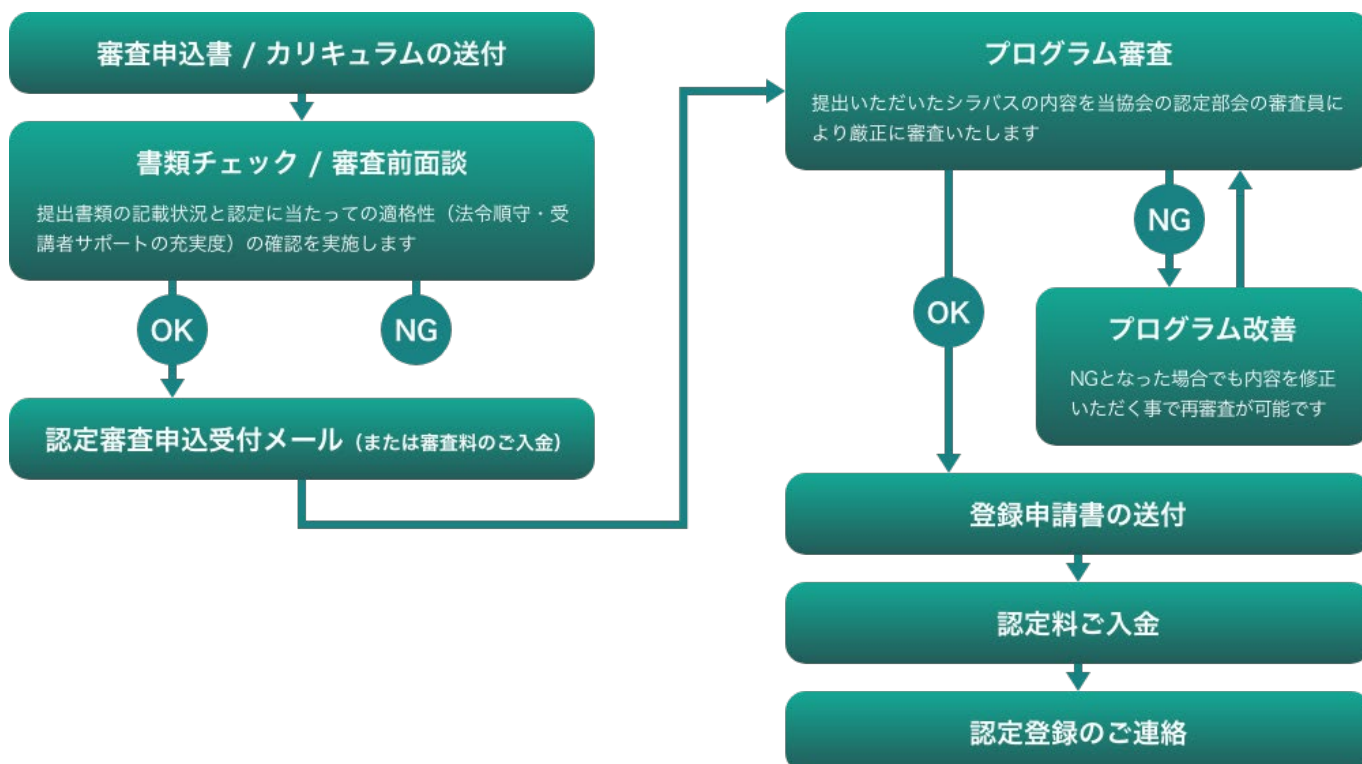
E資格の試験範囲

応用数学 線形代数 確率・統計 情報理論 機械学習 機械学習の基礎 実用的な方法論 深層学習 順伝播型ネットワーク 深層モデルのための正則化 深層モデルのための最適化 畳み込みネットワーク	深層学習 回帰結合型ニューラルネットワークと再帰的ネットワーク 生成モデル 強化学習 深層学習の適応方法 開発・運用環境 ミドルウェア 軽量化・高速化技術
--	---

<認定プログラム>

JDOLA 認定プログラムとは、E 資格の受験資格を得るために修了が必要となる教育プログラムです。プログラム認定を目指す教育事業者からの申請に基づき、その教育プログラムが、JDOLA が定める最新のシラバスの内容を網羅しているか等を審査し、認定推奨しています。

認定プログラム申請の流れ



学校教育法第一条が定める「学校」は認定にかかる費用が無償となります。また、JDOLA 認定プログラムを提供しているプログラム実施者から、同一内容の教育プログラムを受託し提供できる受託制度もあります。

3. 他団体との連携

JDLA の活動に賛同し、互いの発展のために協力する公共団体、非営利団体との連携も積極的に行っており、人材育成領域においては独立行政法人情報処理推進機構（IPA）、一般社団法人データサイエンティスト協会（DSS）と共同で 2021 年 4 月にデジタルリテラシー協議会（Di-Lite）を発足しました。「Di-Lite(ディーライト)」とは、デジタル時代の全ビジネスパーソンが持つべきデジタルリテラシーのことであり、IT・ソフトウェア領域に、デジタル時代、産業界において重要さが高まる「データ×AI」活用に関連した数理・データサイエンス、AI・ディープラーニング領域を加えた領域の基礎領域から、共通リテラシー領域を定義します。その中で、AI・ディープラーニング領域におけるリテラシー範囲として「G 検定」のシラバス範囲が設定、取得推奨されています。

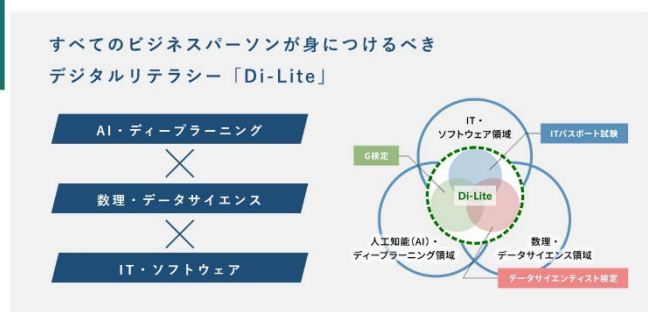


デジタルリテラシー協議会 協議委員



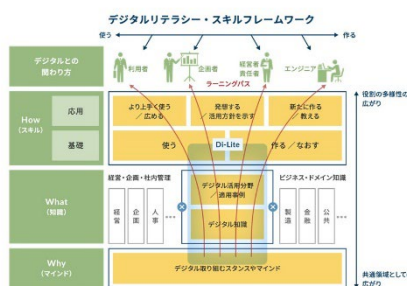
「Di-Lite」とその領域・範囲について

Di-Lite



「Di-Lite」をベースに、今後デジタル領域の全体像を整理

Di-Lite



デジタルリテラシー協議会は、この「Di-Lite」をベースに、増え続ける IT スキルや知識をビジネスとの関連性から体系化することで各ビジネスパーソンが取るべきラーニングパスの見える化をしていきます。変化のスピードが早いデジタル社会に対応していくため、この「Di-Lite」は、各参加団体を中心に開催する協議会において毎年その領域定義をアップデートするものとし、産業界の声を取り入れながら社会標準実装に向けて努めてまいります。

4. AI 人材コミュニティ活動

JDLA では、G 検定・E 資格の合格者のみが参加できる AI 人材コミュニティ「CDLE (Community of Deep Learning Evangelist)」を主宰し、2022 年 3 月現在で 4 万人を超える日本最大の AI コミュニティとなっております。CDLE ではディープラーニングの社会実装の日本代表として、社会を発展させるエバンジェリストたちが集まり、学び合い・アウトプットする場を提供しています。



CDLEについて
コンテンツ
ブログ
イベント
利用規約
プライバシーポリシー
特定商取引法に基づく表記
反社会勢力に対する基本方針
お問い合わせ
よくある質問



CDLE VOICE - AIを学べるラジオ

- 概要：AI学習や情報発信に役立てるためのCDLEメンバーによる参加型・対談式・収録式ラジオ番組
- 配信内容：
 - AI(ML, DL)に関する日々の学び、研究の成果、実務からの経験談、CDLE内での活動
- 収録・配信方法：
 - 番組MC Kenとのオンライン(zoom)収録
 - 収録された音声データを右記のメディアで配信



DM、その後収録日程調整を行います

対談式・収録式のラジオ番組です！



テーマ
スポンサー

Tokyo Marunouchi Innovation Platform (TMIP)

テーマ

都市課題をデジタル・データを活用して解決する
サービス・ソリューションの開発

- 目指す街の姿例
- ✓ 個人や企業が交流しイノベーションが生まれる街
 - ✓ 安全・安心かつ快適に移動できる街
 - ✓ 環境負荷が低くサステナブルな街
 - ✓ Well-Being (健康・快適・便利) を実現する街
 - ✓ 多様なワークスタイル・ライフスタイルに対応する街

データ提供

三菱地所

丸の内カード購買データ

デンソー

スマホで安全運転意識を高めるアプリ
yuriCargo*のデータ
*<https://www.yuricargo.com/>

2022/3/4

情報処理学会 第84回全国大会 一般情報教育と数理・データサイエンス・AI

1

情報処理学会における
データサイエンス教育の取り組み

萩谷 昌己 (情報処理学会データサイエンス戦略推進委員会委員長、東京大学)

以下のスライドは、佐賀大学の樹下 哲郎先生 (情報処理学会データサイエンス教育委員会委員長) からご提供いただいたものです。

2022/3/4

情報処理学会 第84回全国大会 一般情報教育と数理・データサイエンス・AI

2

データサイエンス分野の教育と
資格制度に関する
情報処理学会の取り組み

2022/3/4

情報処理学会 第84回全国大会 一般情報教育と数理・データサイエンス・AI

3

データサイエンス・カリキュラム標準 (専門教育レベル) の公開

情報処理学会
プレスリリース

2021年4月15日

一般社団法人 情報処理学会

データサイエンス・カリキュラム標準 (専門教育レベル) の公開

一般社団法人情報処理学会 (会長：江村克己) は、データサイエンス分野における大学レベルの専門教育を対象としたカリキュラム標準の策定を進めています。このたび、同カリキュラム標準がまとまりましたので、公開いたします。データサイエンス教育やデータサイエンス分野の高度化は社会に必要とされているため、本学会では、データサイエンス教育やデータサイエンス・カリキュラム標準 (専門教育レベル) は、関連する様々な取組を参照して策定されており、以下に挙げる様々な取組を持っています。

- ・ACM Data Scienceのカリキュラムおよび欧州EDISON Data Science Frameworkの参照を通じて、国際的通用性を確保する。
- ・データサイエンス・DSスキルチェックリスト (★レベル) の参照を通じて、知識・スキルの具体化を図る。
- ・データサイエンスを専門とする学生に求められる知識・スキルの具体化を図る。
- ・60単位程度 (標準的な授業時間：約675時間) の学修を前提とする。
- ・数理・データサイエンス・AIモデルカリキュラム (リテラシーレベル) の学修を前提とする。

大学側関係者の改正 (令和3年2月26日) に伴って送られた、連携関係科目や、新型コロナ禍に伴って普及したオンライン教育を機動的に活用して、本カリキュラム標準を参照したデータサイエンス専門教育の推進されることを期待しています。

記

1. 情報処理学会 データサイエンス・カリキュラム標準 (専門教育レベル)

2. 「データサイエンス・カリキュラム標準 (専門教育レベル)」に関する最速最善の結果について

2022/3/4

情報処理学会 第84回全国大会 一般情報教育と数理・データサイエンス・AI

4

カリキュラム策定方針

ACM Data Scienceカリキュラムおよび欧州EDISON Data Science Frameworkの参照を通じて、国際的通用性を確保する。

データサイエンス・DSスキルチェックリスト (★レベル) の参照を通じて、情報処理学会で策定中のデータサイエンス資格との連携を図る。

データサイエンスを専門とする学生に求められる知識・スキルの具体化を図る。

60単位程度 (標準的な授業時間：約675時間) の規模とする。

数理・データサイエンス・AIモデルカリキュラム (リテラシーレベル) の学修を前提とする。

参照王子的特徴

5

ACM DS Curriculum

- 教育項目を詳細に定義
- トピックの優先度を定義: T1, T2, E
- 不足項目

Edison DS
Framework

- 最も広範囲をカバー
- ECTS(単位数)の割り当て
- 詳細な教育項目は定義されていない

DSスキルチェックリスト

- 詳細なスキル定義に特化
- 実務家に求められるスキルを定義
- ビジネススキルを含む

T1: 必須項目 (100%)
T2: 準必須項目 (80%)
E: 選択項目 (50%)

3つの参照モデルを対応付け

3つの参照モデルの長所を活かす

中分類	T1	T2	E	統計	有効教育 時間数(h)
A: 基礎数学と数理統計学	18.25		19.00	37.25	27.75
B1: データマイニング	28.00	36.00	78.00	142.00	95.80
B2: 人工知能	3.25	17.00	12.50	32.75	23.10
B3: 機械学習	13.75	19.50	13.50	46.75	36.10
C: モデリングとシミュレーション	11.80	18.38	47.00	77.18	50.00
D1: コンピュータサイエンス基礎	81.16	22.50	26.25	129.91	112.28
D2: ビッグデータ・システム	7.00	59.50	47.00	113.50	78.10
E1: セキュリティとプライバシー	41.93		4.00	45.93	43.93
E2: ヒューマン・コンピュータ・インタラクション	16.83			16.83	16.83
F: ソフトウェア工学	53.50	24.50	38.00	116.00	92.10
G1: データ収集・管理・ガバナンス	27.01	25.67	17.00	69.68	56.04
G2: プロフェッショナルイズム	16.14	12.01	7.99	36.14	29.74
G3: ビジネス基礎	10.50		8.33	18.83	14.67
総計	326.12	235.55	318.57	880.24	673.85

DSカリキュラム標準(専門教育レベル)の構成



T1	• 必須項目
T2	• 進必須項目 • 80%履修
E	• 選択項目 • 50%履修

7

BC: 人工知能

[illegible]

ACH-39-08	T1	0.25	KAD 02
1. 処理に異常はないわけではなく、むしろ正常な結果をもたらす可能性を十分に把握している。			
ACH-39-09	T2	0.50	KAD 02
2. 知識表現と推論 (論理ベースモデル)			
ACH-40-01	T2	0.50	KAD 02
ACH-40-02	T2	0.50	KAD 02
ACH-40-03	T2	0.50	KAD 02
ACH-40-04	T2	0.50	KAD 02
ACH-40-05	E	0.50	KAD 02
ACH-40-06	E	0.50	KAD 02
ACH-40-07	T2	0.50	KAD 02

IPSJデータサイエンティスト育成戦略

2021年4月13日

合

- ・IPSIは、DS人材の系統的な育成および評価の枠組みを構築し、育成または資格を付与したDS人材による価値創造活動に対する支援を通じて、デジタル社会の発展に貢献してまいります。

活動方針

- ・小学生から専門人材に至る系統的なDS人材の育成および育成／評価の枠組みを構築します。
- ・データサイエンスカリキュラム標準（専門教育レベル）に関する大学間連携や教育コンテンツの共同利用を推進します。
- ・関連する既存の取り組みの活用と相互連携を推進します。
- ・DS人材による相互研鑽及びデータ利活用の倫理を含めた技術者倫理を規範とする社会貢献を推進します。

海外の取り組みとの連携

27 September
15:00-16:15 CET



Strategic Collaboration of Certification and Education on Data Science



Tetsuro Kakushita
Information Processing
Society of Japan



Yuri Demchenko
Edison Data Science Project
Society of Japan



Boote Cassel
ACM Data Science
Task Force



Kewal Dhariwal
Institute of the Certification
of Computing Professionals



Moira de Roche
IP3 / IITPSA

IFIP 60th Anniversary *Future of Information Processing* Event Series

<https://www.ipthreer.org/knowledge-portal/ip3-presentations/ifip60-webinar-strategic-collaboration-of-certification-and-education-on-data-science/>

求められる問題解決力のための知識を測る

平岡 靖敏

日本規格協会 品質管理検定センター
東京都港区三田 3-13-12

Tel : 050-1742-6445

hiraoka@jsa.or.jp

1. QC(品質管理)検定とは

1.1 QC 検定の概要

- * 日本規格協会と日本科学技術連盟が共催の品質管理の知識を問う民間検定試験である。
- * 2005 年開始。年 2 回、累計 32 回(2 回は中止)実施してきている。
- * 累計受検者(~32 回) : 1,336,143 人
累計合格者(~32 回) : 643,582 人
- * 働く人、これから働く人を対象として、4 つの級で構成
 - 1 級: 品質管理専門家として企業内のあらゆる品質や品質管理に関する仕事を主導する人々
 - 2 級: 技術部門や生産部門において、品質問題解決を主導する人々
 - 3 級: 業種・業態にかかわらず、自分達の職場の問題解決を行う人々
 - 4 級: 派遣社員、高校生、初めて品質管理を学ぶ人々

1.2 品質管理と QC 検定

- * 品質管理の目的
品質管理とは、お客様の要求を満足する製品・サービス(Q)を、安く(C)、欲しい時(D)に提供するための活動である。
それらを提供するのは自身だけでなく、同業他社が存在し、市場競争の中にあるので、製品・サービスの QCD の向上を常に図り、市場競争力を強化させて収益を向上させることにある。
- * 品質管理の基本
品質管理は「管理」と「改善」から成り立っている。製品・サービスの QCD について、企画・設計された通りに提供するために行われる活動が「管理」である。しかし実際には、決められた通りに行っても、狙った QCD が得られないことが多い。そのため、「管理」と「改善」を継続して行われる。このサイクルがいわゆる PDCA

(Plan-Do-Check-Act) である。その中で、“CA”にあたる部分の活動が「改善」であり、「Kaizen」として海外でも通じるほど日本的品質管理の特徴ともなっている。

* 改善と問題解決

この改善の進め方、プロセスが科学的に整理されたものが、QC ストーリーである。

QC ストーリーの流れは、現状把握⇒要因分析⇒対策立案・実施⇒効果確認となっており、この流れに沿って改善を進めることが即ち、問題解決である。

* 問題解決と QC 検定

この問題解決の進め方／考え方についての知識に加え、現状把握や要因分析、効果確認などで使われる手法についての知識を具備することで、改善が効果的に実施することが可能となる。一方、改善しなければならない問題はさまざまであり、それぞれ適切な手法を、適切に使用するための知識が求められる。

そのため QC 検定では、問題解決に必要な手法や知識を、先に述べた働く人々の役割による 4 つの級に分けて試験を行っている。

2. QC 検定の普及

2.1 QC 検定制度スタートの背景

* 品質管理教育の広がり実践

1960 年代以前の品質管理教育は、企業の従業員など社会人を対象として、品質管理普及団体で行われてきていた。さらに、品質管理が多くの企業に広まるにしたがって、1970 年代以降は企業内でも盛んに行われるようになった。

これは、問題解決についての教育が学校教育では行われていなかったことに加え、問題解決による製品・サービスの QCD 改善が自社の競争力向上に直接的に効果があることを理解していたからである。

現場での改善は、職場単位で組織される QC サークルによる改善活動が有名であるが、QC サークルは同時に改善の実践の場としての教育機能も担っており、現在も活動が行われている。

＊ 派遣労働者の増加

バブル崩壊後の 1990 年代中頃より、工場での派遣労働者の活用が急速に増加した。

派遣労働者に対する教育は企業側では行わず、人材派遣会社でも品質管理教育は行っていない。また、多くの場合、派遣労働者は QC サークルにも組み入れられないため、工場において多くの問題が発生していた。

ただし、工場で働く派遣労働者に対する期待は、改善や問題解決までではなく、生産活動の常識や品質管理の基礎を理解しているレベルの期待であった。

＊ 品質管理能力の評価

1990 年代中盤以降、経済成長の変化に伴って企業の人事制度も影響を受け、能力主義・評価が広がると、品質管理能力についてもどのような問題を解決できるかといった評価に対するニーズが徐々に高まった。

2.2 ニーズに応えるための制度設計

＊ 能力評価の方法

品質管理の能力、特に改善・問題解決に関する能力と、その評価に対する企業ニーズは高いが、問題解決を実践するためには、推進力やリーダーシップに代表される個人の資質や経験知に加え、モチベーションを持たせる組織体制や取り組みも大きく影響する。外的要因を取除き、個人の能力を総合的かつ客観的に多くの人数を評価することは極めて困難である。

一方、問題解決は現状把握や原因分析などを行う手法や考え方の知識がなければできないことから、それらの知識を測ることに限定する試験の提案を 100 社以上に尋ねた結果、多くの前向きな回答を得た。

＊ レベルの設定

企業等において働く人々について、業務や品質管理活動でどのような役割を担うのかを基本に、その役割を果たすためにはどのような能力や知識が必要であるかを検討・整理し、4 つのレベルの「対象となる人材像」、「認定する知識と

能力のレベル」に表した。

＊ 品質管理検定レベル表

レベル表では、「対象となる人材像」、「認定する知識と能力のレベル」を提示し、知識についてさらに詳細にブレイクダウンを行った。

これにより、個人若しくは組織が品質管理・改善、問題解決についての能力とそれに必要な知識が明確になった。

また、レベル表の構造は、上位級は下位級の知識をすべて包含した形となっている。これは品質管理での問題解決では、難解な問題は高度な手法を用いれば解決できるものではなく、複数の部署や職位の人々が関わることも多く、広い範囲の手法に精通していることが求められるからである。

＊ 4 級とテキスト提供

一番基本のレベルとして設定した 4 級は、初めて品質管理を学ぶ人、具体的には新入社員や派遣労働者、高校生などを対象としたが、企業関係者からの強い要望を受けて、品質管理の基本だけでなく、企業活動の基本常識に関する知識についても範囲とした。

また、これらの 4 級範囲をカバーする教材として、自ら学ぶことができる「QC 検定 4 級テキスト」を新たに作成し、無償で提供することとした。これにより、品質管理を学ぶ機会が得られない人々でも独学を可能とした。

2.3 受検者（申込ベース）の推移

	第 1 回	第 5 回	第 10 回	第 15 回	第 20 回	第 25 回	第 30 回
1 級	192	987	1300	1653	1462	1738	1171
2 級	916	5324	8431	12089	13758	15198	13312
3 級	1606	9650	17619	26515	34530	38038	32530
4 級	1245	3712	6786	8795	10336	12205	9304
合計	3859	19673	34146	49052	60086	67179	56317

2.4 各級の累計受検者と割合（申込ベース）

	累計受検者数	割合
1 級	39,416	2.8%
2 級	331,906	24.0%
3 級	762,907	55.1%
4 級	250,579	18.1%

あなたは、データを分析・活用して 新たな価値を創造できますか？

近年、顧客行動の複雑化により、企業においては、データに基づいた確かな戦略・施策をすることが、どんな業種・職種・プロジェクトにおいても求められるようになりました。
いまや、「文系だから」「専門ではないから」という理由で、データから目をそむけることはできない時代です。
今後、「データ」の把握・分析・利活用は、ますます注目されることになるでしょう。
そして、その基盤となる教科は、間違いなく「数学」です。



「データサイエンス数学ストラテジスト」は
データサイエンスの基盤となる数学スキルを
測定して認定する**新しい資格制度**です。

国が目標に掲げる 数理・データサイエンス・AI 人材の育成方策

AI 戦略、データサイエンス、IoT (Internet of Things) などに対応できる人材の需要は高まり、政府の「AI 戦略 2019」(2019 年 6 月策定)では、文理を問わず年間約 25 万人の数理・データサイエンス・AI 人材育成(応用基礎レベル。さらにリテラシーレベルで年間約 50 万人)を 2025 年までに実現するよう目標に掲げられています。



(参考)

- ・首相官邸 統合イノベーション戦略推進会議「AI 戦略 2019」
- ・数理・データサイエンス教育強化拠点コンソーシアム「数理・データサイエンス・AI(応用基礎レベル)モデルカリキュラム」



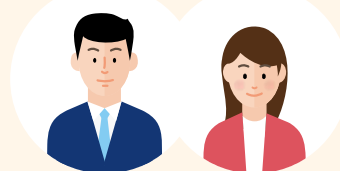
■ 資格のメリット

社会人の方

仕事に生かせる！キャリアアップや転職にも活用できる

- 業務データを数値的に解釈して、ビジネス課題の解決に貢献できます。
- データサイエンスに必要な数学を効率よく身につけられます。
- データサイエンスを主とした事業戦略・施策に関わるデータサイエンス数学の一定の知識を習得できます。

\ GOOD! /



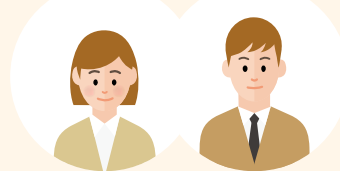
大学生・高校生・高専生の方(文系・理系問わず)

進学・就職に役立つ！

データサイエンスの基盤となる数学スキルを証明できる

- 数理・データサイエンス・AI人材としての素養があることをアピールできます。
- 就職後にデータサイエンス戦略(データの把握や分析など)の即戦力として活躍できます。

\ GOOD! /



■ 有識者のメッセージ

村上 雅人氏／芝浦工業大学 学事特別顧問・前学長

現代社会では、DXやSociety5.0など、IT技術を活用した変革が求められています。その基本は、データ駆動型社会であり、理系だけでなく、文系においてもデータサイエンスの素養が必要とされています。企業や組織のプロジェクトにおいて何かを決断する際にも、思い込みではなく、数値データを基礎に置くこと、そして、数値データの正当性を確かめ吟味する力も必要です。それには四則計算で十分です。あなたの新しい未来を開くためにも、ぜひ「データサイエンス数学ストラテジスト」の資格にチャレンジしてみてください。



三井住友海上火災保険株式会社 デジタル戦略部 データサイエンティストのみなさま

ビジネスパーソンにとって、データサイエンスの基礎理論にあたる「数学」を学ぶ意義は何でしょうか。市販されているデータ分析ツールはその中身のしくみを知らなくても使うことは可能ですが、基礎理論を身につけていれば、データ分析／AIツールによる結果を的確に人に説明できるようになります。ツールの結果を業務に生かすには理論的背景が欠かせないのです。ツールを使う側がデータサイエンスの理論的背景を加えることで、より効果的に使うことができるようになるのです。

立ちどまらない保険。

三井住友海上

MS&AD INSURANCE GROUP

試験内容

資格試験の概要

「データサイエンス数学ストラテジスト」には、中級と上級の2つの階級があり、それぞれデータサイエンスの基盤となる基礎的な数学(確率統計・線形代数・微分積分など)と実践的な数学(機械学習系・アルゴリズム系・ビジネス系数学など)の習熟度・理解度を測定する問題を出題します。資格試験は、5肢択一のIBT(Internet Based Testing)形式で行います。

	中級	上級
受験資格	なし	
対象の目安	社会人・大学生・高校生	
数学のレベル(目安)	数検準2級程度 (数学Ⅰ・Aまで)	数検2級・準1級および 大学初学年程度まで
問題数(5肢択一問題)	30問	40問
試験時間	90分	120分
合格基準	60%(18問)以上	70%(28問)以上
受験料	7,000円	9,000円

受験方法

資格試験は、個人受験または団体受験で受験できます。個人受験は、「CBT-Solutions」「アガルートアカデミー」のいずれかの外部サービスで、申し込み、受験が可能です。団体受験は、企業や学校などで一括申し込みを行って受験するものです。団体ごとに合わせたプランをご提案いたしますので、当協会までお問い合わせください。

個人受験の流れ

① 申し込み手続き

「CBT-Solutions」または「アガルートアカデミー」の受験申込サイトにアクセスして、「データサイエンス数学ストラテジスト」資格試験の申し込み手続きをします。



▶ 資格試験のお申し込みはこちら

② 受験

試験有効期限までに、資格試験を受験します。

用意するもの

- ・筆記用具
- ・電卓または関数電卓
- ・計算用紙
- ・表計算ソフト(必要に応じて)

③ 結果

試験終了後、合否結果が画面上に表示されます。「合格」「不合格」のほか、総得点、評価コメントなどが試験結果として表示されます。

④ オープンバッジの発行

各階級の合格基準点に到達された方には、オープンバッジを発行いたします。

「データサイエンス数学ストラテジスト」オープンバッジ一覧(全6種) ▶



オープンバッジとは

オープンバッジとは、世界的な技術標準規格「IMS Global Learning Consortium」に準拠して発行されるデジタル証明・認証のことです。世界では、すでに年間2400万個が発行され、資格のほかに免許証や卒業証明書など用途はさまざまに広がっています。オンライン上で公開したり、SNSなどで共有できたりするなど、利便性の高さが特徴です。偽造・改ざんが困難なブロックチェーン技術に応用しており、学習・スキルの証明書として信頼性の高いものとなっています。オープンバッジの詳しい情報は、下記のWEBサイトでご確認ください。

一般財団法人オープンバッジ・ネットワーク <https://www.openbadge.or.jp/>

試験内容

「データサイエンス数学ストラテジスト」の資格試験は、以下の4つの学習分野で構成しています。

【学習分野①】 AI・データサイエンスを支える 計算能力と数学的理論の理解	問題配分 中級：50% 上級：50%	【学習分野②】 機械学習・深層学習の 数学的理論の理解	問題配分 中級：16.7% 上級：25%
- 確率統計系分野(統計・確率・場合の数 など) - 線形代数系分野(行列・ベクトル など) - 微分積分系分野(微分積分・関数・写像 など)		- 基礎理論(活性化関数・類似度・最小二乗法) - 機械学習(回帰・分類・クラスターリング など) - 深層学習(ニューラルネットワーク など)	
【学習分野③】 アルゴリズム・プログラミングに必要な 数学リテラシー	問題配分 中級：16.7% 上級：12.5%	【学習分野④】 ビジネスにおいて 数学技能を活用する能力	問題配分 中級：16.7% 上級：12.5%
- アルゴリズム(探索・ソート・暗号、計算量) - プログラミング言語に依存しない手続き型思考 - 数学的課題解決(論理的思考＋数学的発想)		- 把握力(データ・グラフの特徴の把握 など) - 分析力(売上・損益等財務的な分析 など) - 予測力(データに基づいた業績予測 など)	

資格到達目標 「データサイエンス数学ストラテジスト」資格では、以下の内容を到達目標としています。

各領域	到達目標
領域 1 データ 集計・分析	データサイエンスに必要なデータの集計・分析手法の理解・習熟 ・データ分析目的の設定、データの収集・加工・集計、比較対象の選定 ・データのばらつき度合、傾向・関連・特異点の把握 ・時系列データ、クロスセクションデータ、パネルデータの理解 ・目的に応じた図表化・可視化(棒グラフ、折線グラフ、散布図) など
領域 2 数学基礎	データサイエンス戦略・施策に必要な数学の基礎 ■算数・中学校数学分野 ・四則計算、グラフ、比例と反比例、単位あたりの大きさ、文字式の計算、方程式、1次関数、三平方の定理、思考力を測る問題 ■確率統計系分野 ・平均値・中央値・最頻値、分散、標準偏差、統計基礎 ・割合、順列・組み合わせ、二項定理、確率、確率分布 ・データの分析、資料の整理・活用、標本調査 ■線形代数系分野 ・ベクトルの演算(和とスカラー倍、内積) ・行列の演算(和とスカラー倍、積)、行列式 ・固有値と固有ベクトル ■微分積分系分野 ・指数関数、対数関数、三角関数、2次・多項式関数、写像 ・数列、関数と極限、微分・積分 ・偏微分、重積分、微分方程式の基礎 など
領域 3 機械学習基礎	データサイエンス戦略・施策に必要な機械学習の基礎 ・基礎的な理論(活性化関数、距離による類似度、最小二乗法) ・教師あり学習(回帰(回帰直線)、分類(線形識別・混同行列)) ・教師なし学習(クラスターリング、次元削減) ・関連研究分野(自然言語処理、データマイニング) など
領域 4 深層学習基礎	データサイエンス戦略・施策に必要な深層学習の基礎 ・ニューラルネットワークの原理、勾配降下法 ・ディープニューラルネットワーク(DNN) ・畳み込みニューラルネットワーク(CNN) など
領域 5 アルゴリズム・ プログラミング的思考	データサイエンス戦略・施策に必要なアルゴリズム、プログラミング的思考 ・アルゴリズム(探索・ソート・暗号)、計算量理論 ・特定のプログラミング言語に依存しない手続き型思考、情報理論 など
領域 6 数学的課題解決	論理的思考と数学的発想を用いた課題解決 ・課題から解答まで矛盾なく導く論理性、一貫性 ・課題の読み取り、規則性・法則性の発見
領域 7 コンサルティング	ビジネスシーンでのデータサイエンス戦略・施策の実現方法の検討、提案 ・顧客、ステークホルダーの要望・意見を聞くコミュニケーション力 ・戦略・施策の実現方法を検討し、提案するプレゼンテーション力

数理・データサイエンス・AI 教育プログラムの導入をお考えの方へ

当協会は、内閣府・文部科学省・経済産業省の3府省が進める「数理・データサイエンス・AI 教育プログラム」の趣旨に賛同し、数理・データサイエンス・AIの基礎的素養を兼ね備えたDX(デジタルトランスフォーメーション)を担う人材の育成と活用を推進しています。

■ 数理・データサイエンス・AI 教育プログラム認定制度とは

数理・データサイエンス・AI 教育プログラム認定制度とは、内閣府・文部科学省・経済産業省の3府省が連携し、各大学・高等専門学校における数理・データサイエンス・AI 教育の取り組みを奨励するための認定制度です。

政府の「AI 戦略 2019」においては、データとデジタル技術を活用したビジネスモデルの抜本的な変革(DX)をしていくことが重要で、このDXを担う人材として、「数理・データサイエンス・AI」を理解し、活用できる人材が必要であるとしています。これらを身につけた人材を育成するしくみとして、大学(大学院を除く)、短期大学、高等専門学校が実施する教育プログラムを認定する「数理・データサイエンス・AI 教育プログラム認定制度」が創設されました。

▶ 内閣府「AI 戦略 2019」

www8.cao.go.jp/cstp/ai/index.html

▶ 文部科学省「数理・データサイエンス・AI 教育プログラム認定制度(リテラシーレベル)」

www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/00002.htm

■ 「MDASH SUPPORTERS」とは

当協会は、「数理・データサイエンス・AI 教育プログラム」の趣旨に賛同し、支援する団体として、経済産業省の「MDASH SUPPORTER*」に認定されています。

※Approved Program for **M**athematics, **D**ata science and **A**I **S**mart **H**igher Education

「MDASH SUPPORTER」は、数理・データサイエンス・AIに関する基礎的な能力を修得した人材がより多く輩出されることを期待する産業界の声を集めることを通じて、大学・高等専門学校での教育プログラムを整備する取り組みを後押しするとともに、人材の活躍の場が広がることを目的としています。

▶ 経済産業省「数理・データサイエンス・AI 教育プログラム支援サイト」

www.meti.go.jp/policy/it_policy/jinzai/MDASH/mdashsupport.html

日本数学検定協会から、大学生や受講生などに向けたメッセージ

学校教育で育んだ数学リテラシーを実社会で生かす産業人材育成のために、
私たちは「数理・データサイエンス・AI 教育プログラム」で支援いたします。



■ 当協会における数理・データサイエンス・AI 教育の取り組みについて

当協会は、「数理・データサイエンス・AI」の基盤となる教科「数学」に関する検定事業や資格制度などの取り組みをとおして、数理・データサイエンス・AIの基礎的素養を兼ね備えたDXを担う人材の育成と活用を推進しています。

「数検」の実施

実用数学技能検定「数検」(後援＝文部科学省。対象：1～11級)を通じて、数学・算数の学習成果を評価・顕彰する生涯学習の場を提供しています。累計志願者数はのべ600万人を突破しており、数学・算数に関する検定のスタンダードとして、児童・生徒の学力定着の確認や学習意欲の向上、進学・就職をはじめ、社会人の学び直しなどに活用されています。



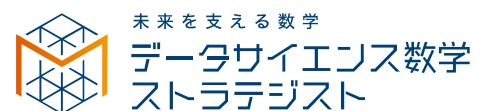
ビジネス数学の普及

社会・企業と数学との接点を明らかにし、実社会における数学的リテラシーの向上や、企業の効率的な人材育成につなげるため、ビジネスにおいて必要とされる数学の検定および研修などを実施しています。



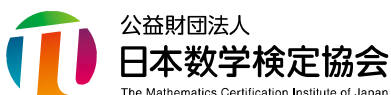
「データサイエンス数学ストラテジスト」の資格認定

データサイエンス戦略(データの把握や分析など)において必要とされる数学的スキル・リテラシーの学習機会を提供しています。また、理解度・習熟度を測定し、データサイエンス数学ストラテジストとして資格認定しています。



- ▶ 当協会の「数理・データサイエンス・AI教育プログラム」に関する最新の情報は、右記の二次元バーコードを読み取るか、下記をご覧ください。

www.su-gaku.net/business/mdash/



〒110-0005 東京都台東区上野 5-1-1 文昌堂ビル 6 階

TEL : 03-5812-8340 FAX : 03-5812-8346

電話お問い合わせ時間：月～金 10:00～16:00（祝日、年末年始、当協会の休業日を除く）

E-mail : mdssinfo@su-gaku.net

データサイエンス数学ストラテジスト 特設サイト : ds.su-gaku.biz



SP01-2110

データサイエンスの能力とその測定方法
～統計検定 CBT データサイエンス発展級・エキスパート級の設計を通じて～

佐藤彰洋・横浜市立大学データサイエンス学部

竹村彰通・滋賀大学データサイエンス学部

横浜市金沢区瀬戸 22-2・045-787-2208・ahsato@yokohama-cu.ac.jp

1. はじめに

統計検定データサイエンス発展(DS 発展)[1]ではデータサイエンス要素を4つの領域(数理、情報、統計・可視化、倫理・AI)の分類し、これらの領域間の重複を考慮した7種類の問題型の正答率により受講者のデータサイエンス能力を多角的に測定することを目指し、2019年5月から基本設計を開始し、2021年9月28日に試験配信を開始した。

本稿では、その問題傾向を示すとともに、統計検定データサイエンス発展を設計するのに際し、検討したデータサイエンス能力の定義とその計測の方法について説明する。更に、現在準備中の統計検定データサイエンスエキスパート(DS エキスパート)[2]の設計指針について、その方向性をDS発展のより高度化したものとしてDS エキスパートを位置付けて述べる。

今回、データサイエンス能力測定の検定試験設計の基本的戦略として、我々は、椿(2019)が言及するプログラム科学の文法[3]を採用した。

椿は、プログラム科学へのアプローチ法として、QCストーリーとPDCAとを結合した次のような方法論を提案している。

1) あるべきプログラム・パフォーマンスの明確な定義

2) プログラム・パフォーマンスにおける看過できない違反や例外の発見

3) パフォーマンスに影響をあたえる可能性のある記号列の論理的網羅

4) 記号列の必要十分性の観点からの比較・評価

5) 最適な記号列プログラムの実装

6) プログラムの実施とそのパフォーマンスのモニタリング

本稿では、この椿の方法論に沿って、DS発展、DS エキスパートの検定試験策定の過程を追って説明していく。我々は、プログラム科学の手法を用いながら、データサイエンスの能力を最大限に活用し、データサイエンス検定を設計することを志向してこの作業に取り組んだ。

本稿は以下の構成となっている。2章において、データサイエンス検定試験のあるべき姿について椿の方法1)～4)の分析結果について説明する。3章では、DS発展、DS エキスパートの問題検討に用いた試作問題の例について述べる。4章ではDS発展、DS エキスパートの作問プロセスについて概要を述べる。5章は今後の課題と展望である。

2. データサイエンス検定の分析戦略

椿の方法では、5)の最適な記号列プログラムの実装の前に、1)～4)の観点から取り組みたいプログラムの詳細な分析と検討を行うべきであることを提案している。そのため、データサイエンス検定を設計するに当たり、「データサイエンス能力を計量する」というデータサイエンス検定プログラムを1)～4)の観点から分析を行うこととした。

2. 1 パフォーマンスの明確な定義

椿の方法では、1)「あるべきプログラム・パフォ

パフォーマンスの明確な定義」を行うことが最初に重要とされている。我々は、データサイエンス検定のプログラム・パフォーマンスを次の3つの部分に分解して検討を行うこととした。

- ・データサイエンスの定義
- ・出題範囲の特定
- ・受講者の特定と受講メリットの醸成方法

データサイエンティストの基本的なスキルについて議論がいくつかあるが、Shelly Palmer[4]が提案したとされる図1に示すドメイン専門性、数学、コンピュータサイエンスの3つの領域とその合成能力がデータサイエンス能力[5]であるという定義に着目した。この定義に従うと、ドメイン専門性とコンピュータサイエンスの合成能力はデータ処理であり、ドメイン専門性と数学の合成能力は統計的研究、コンピュータサイエンスと数学との合成的能力は機械学習、ドメイン専門性とコンピュータサイエンスと数学の合成能力が狭義のデータサイエンスである。この定義から、データサイエンス力とはいくつかの分野にまたがる能力を同時に利用して問題解決をはかる能力と考えてみる。

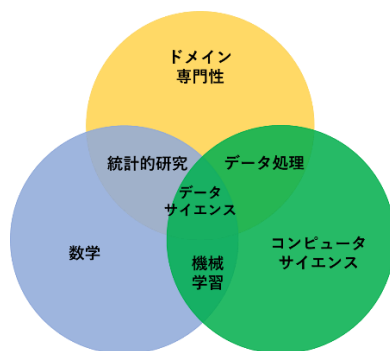


図1：データサイエンティストの基本的スキル[4]

次に、2)「プログラム・パフォーマンスにおける看過できない違反や例外の発見」を目指して、データサイエンス的な作業を実務的に行うときに起こりうる事例について事例分析を行った。この作業により次のようなストーリーがデータサイエンス的な作業を行う場合に看過できないパフォーマンス上の劣化をもたらすことに着目した。

ストーリー（1）「建築資材会社において建築資材の受注効率化を行うための機械学習アルゴリズムを

含む情報データシステムの構築を試みる。建築資材発注部門のマネージャーと情報システム設計の専門家、データベースエンジニア、プログラマ、数理統計の専門家からなるチームを社内に構築した。数理統計の専門家は建築資材発注のプロセスも知らないし、情報データシステムの設計やコーディングについて専門知識を有していない。情報システム設計の専門家、データベースエンジニア、プログラマは機械学習アルゴリズムも分からないし、建築資材会社のビジネスアーキテクチャも知らない。建築資材会社の資材発注部門のマネージャーはコンピュータも数理統計アルゴリズムもわからない。」

このプロジェクトを開始すると何が起こりそうだろうか。まず、

（1）情報システム設計の専門家、データベースエンジニア、プログラマはドメイン知識の定義のために建築資材に関するレクチャーを受ける。

（2）データベースエンジニアは建築資材発注にかかるデータベースからデータを取り出して定量分析を行ってみる。

（3）情報システム設計の専門家、プログラマはシステムアーキテクチャを判明させるために作業をする。

（4）統計数理の専門家はアルゴリズムの指示をするがプログラマには数理的背景は分からないためコーディングが滞る。

異なる分野の間でのコミュニケーションができないことにより、プロジェクト実施期間の多くの時間が、相互理解のために消費されていく。

ストーリー（2）「英語しか話せない人と日本語しか話せない人からなる日英両方に通用するプロジェクトを構成する。英語しか話せない人は英語の作業に集中するだろうし、日本語しか話せない人は日本語の作業に集中するだろう。」

このプロジェクトは日英両方に通用するプロジェクトにはおそくならないだろう。なぜなら、英語と日本語で共通の認識を持つことができないからである。英語と日本語で共通認識を持とうとして辞書を作り始めるかもしれないが、辞書から作っているとプロジェクトは長期化してしまう。最終的に、プロジェクトは予定期間を終えて、目標を達せず終わってしまうだろう。

これらの2つの事例から導き出されることは、個

別の単一専門能力を有する人を集めても複合問題のプロジェクトを貫徹できないということである。

これは、数学（数理・統計）、コンピュータサイエンス、ドメインでは表現形式が大きく異なることに起因する。すなわち、数学（数理・統計）では数式による表現を多用し、数量的データの理解と数理モデルの記述方法を用いる。コンピュータサイエンスでは、コンピュータコードによる表現を多用し、システムアーキテクチャとデータモデルの記述方法を用いる。ドメインでは、用語、物語による表現を多用し、ビジネスプロセスモデルや参照アーキテクチャーといった概念的モデルを用いる。

これら、分野固有の表現形式を理解できるようになるまでには多くの学習を必要とし、その習熟に長い時間を要する。そのため、単一専門性しか有しない場合、課題を解決するために必要となる他分野の知識や技術を利用することができず、また、分野の異なる専門家では、相互にコミュニケーションができない。このため、データサイエンス分野における分野横断的なプロジェクトが貫徹できないことがデータサイエンスプロジェクトを実行する場合において、パフォーマンスの劣化をもたらす看過できない大問題と考えた。

これらから、1)「あるべきプログラム・パフォーマンスの明確な定義」については、次のように検討をした。

データサイエンス型プロジェクトを、数学（数理・統計）、コンピュータサイエンス、ドメインの3領域の知識と技能を複合的に用いないと解決できないデータ駆動型のコンバージェンス型（単独専門性では解くことが出来ない社会的課題を複数の専門性を用いて対応するべき）プロジェクトと定義した。このような、データサイエンス型プロジェクトを志した者たちが、実社会においてプロジェクト実施上で、最大の課題となると予想される事柄を回避できるような人材として登用・育成されることがデータサイエンス教育におけるあるべきプログラムの目的であるとした。そして、データサイエンス型プロジェクトを社会に実装する上で必要となる異分野間のコミュニケーションラインの基礎を与えるものであるべきと考え、このような知識と技能的に異なる分野間でのインターフェースの基礎となるような、**基礎力**と、**分野**

間での通訳・翻訳ができる能力、**分野複合的な課題の記述、解決、運用力**を内包する課題解決能力を鋭敏に計測することができることをデータサイエンス検定の識別能力と定義した。

2. 2 パフォーマンスに影響をあたえる可能性のある記号列の論理的網羅

次に3)「パフォーマンスに影響をあたえる可能性のある記号列の論理的網羅」について検討した。データサイエンス検定では、直交性の高い分野として数学を数理と統計に分解、コンピュータサイエンスを計算、ドメイン知識をAI・倫理と対応させて、4つの分野とすることにした。

そして、数理・データサイエンス教育強化拠点コンソーシアムカリキュラム分科会データサイエンス教育に関するスキルセット及び学修項目（第1次報告2019年11月発表）[4]に準拠して、大分類をこの4つの要素（数理、計算、統計、AI・倫理）に当てはめてみた。ここで、これら4つの要素の全ての組合せを

数えてみる。組合せを $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ とすると、

$$N = \binom{4}{1} + \binom{4}{2} + \binom{4}{3} + \binom{4}{4} = 15 \quad (1)$$

から $N=15$ 通りであることが分かる。

2. 2. 1 単一要素能力

単一要素能力として、数理、計算、統計、倫理・AIを次のように能力定義を行った。

数理・・・物事を定量的に表現するための関数、変数、方程式の概念、ベクトルと行列

計算・・・計算を自動的に行うための原理、質的、量的にデータを取り扱うための方法と形式、計算機システムの基本原理と構成、データベースとデータモデル

統計・・・不確実性のある現象（観測できていない不可知）を量的に記述し、パターンを発見し、分類、推定と予測を行う方法

倫理・AI・・・ドメインでの振る舞い方。ここにアプローチの方法が複数あっても良い。フレームワーク、方法論を計算に含めるかドメインに含めるかは場合による。社会的要請から検討する場合は、倫理の問題

となるし、技術的要請から検討する場合はAIの問題となると考えた。

2. 2. 2 二要素複合能力

二要素複合能力は4分野から2分野を取り出した組合せであり、全てで6通り存在している。

数理×計算＝数値計算（数理的に表現されるものの計算を数值的に実行する場合の問題）・・・行列やベクトルを計算機で計算する場合に実行されるアルゴリズムや計算の手順、計算量の見積もり。

数理×統計＝数理統計（数理統計の問題）・・・ベクトルや関数、行列を用いて統計モデル確率モデルを記述し、または、演算する問題。

数理×倫理・AI＝数理的研究倫理（数理的に社会を研究するときにある課題）・・・社会的な課題を数量でとらえてこれを数理的なベクトルや関数を用いて表現する問題または、その時の方法論と社会通念上や倫理上気を付けるべきこと。

計算×統計＝機械学習理論（機械学習のアルゴリズム）・・・計算機を用いて統計の演算を行うときに現れるアルゴリズムやその計算量的見積もり。

計算×倫理・AI＝情報的研究倫理（情報システムで社会を研究するときにある課題）・・・計算機を用いた社会課題へのアプローチ、データベースによる集計を実社会問題で行う場合の手順や社会通念上及び倫理上気を付けるべきこと。

統計×倫理・AI＝統計的研究倫理（統計的方法により社会を研究する場合に起こり得る課題）・・・統計的研究方法により社会的課題を取り扱う場合に現れる理論上の課題や表現方法。また、その時に気を付けるべき社会通念上の課題と問題。

2. 2. 3 三要素複合能力

三要素複合能力は4分野中3つを取り出す組合せであり、全部で4通りが存在する。

数理×計算×統計＝機械学習アルゴリズムの実装（機械学習アルゴリズムの問題）・・・機械学習アルゴリズムを用いて計算機でそれを実装する課題。また計算機システムとしての構成やそのコスト分析。

数理×計算×倫理・AI＝計算機科学倫理（社会的課題を数式により表現し、その数式を実際に計算機により社会的課題を解くときにあり得る課題）・・・社会的課題の数理モデルを作成して、その代表変量をデータベース上に集約後、データベースを用いて計算機による集計を行い、社会的な課題理解や問題解決等を実行する問題。また、その時に課題となる社会通念上や倫理上の問題。

数理×統計×倫理・AI＝数理統計科学倫理（数理統計の理論的な方法を社会に適用するときにある課題）・・・数理統計的な理論を社会に適用するときに起こり得る課題。数理統計的な量的な記述と分類、推定予測を社会的課題に対して行う場合の記述能力を問う問題。また、社会通念や倫理上問題となる課題。

計算×統計×倫理・AI＝機械学習科学倫理（統計の検定や機械学習のアルゴリズムを社会に適用しようとする場合に起こり得る課題）・・・機械学習アルゴリズムを用いて実際に社会に存在するデータを計算して、分類、予測、推定を行う課題。また、そのようなときに社会通念や倫理上問題となる課題。

2. 2. 4 四要素複合能力

四要素複合能力は、**数理×計算×統計×倫理・AI**全ての能力を統合して実施する能力であり、データサイエンスの能力である。計算機を用いて社会の問題を数理的な定量的な捉え方により記述し、必要に応じてデータベース化または計算機上で、コンピュータプログラムを用いて集計、計量しながら、パターン発見、分類、推測や予測を実行し社会的な課題を解決する能力である。

計算機を用いて社会的に存在するデータを集め、それらのデータを数学的表現として数的または量的に記述して、分類、予測、推定を実データで行うことで意思決定を行う課題対応能力である。または、そのようなことを行う場合に認識される社会通念上または倫理上の問題に対処する能力となる。更に、そのような実装を行う場合に必要となる計算資源やシステム実現のための費用便益の問題に対処する能力となる。結果、データサイエンス型プロジェクトの課題対処を統合的に行える能力となっている。検討の結果、

数理、計算、統計、倫理・AI の知識と技術を複合的に用いることにより、現実社会におけるデータを起点とした問題解決のかなりの種類を取りあつかうことができることが分かった。

2. 4 記号列の必要十分性の観点からの比較・評価

次に、ここまでに網羅的に列挙した 15 通りの能力に対して、4)「記号列の必要十分性の観点からの比較・評価」を行う。能力の困難さは、そこに含まれる要素数合計であろう。この検討から、DS 発展では単一要素と二要素複合能力までを、DS エキスパートでは、DS 発展を超える高度な要素からなる単一要素と二要素複合能力及び、三要素複合能力、四要素複合能力を問うことにより、異なるレベルでのデータサイエンス能力の測定を可能とすることとした。このように 2 つの検定に対して、その能力測定レベルを分割した。

更に、4 つの要素を図 2 に示すように分解してその連結性について検討を行った。

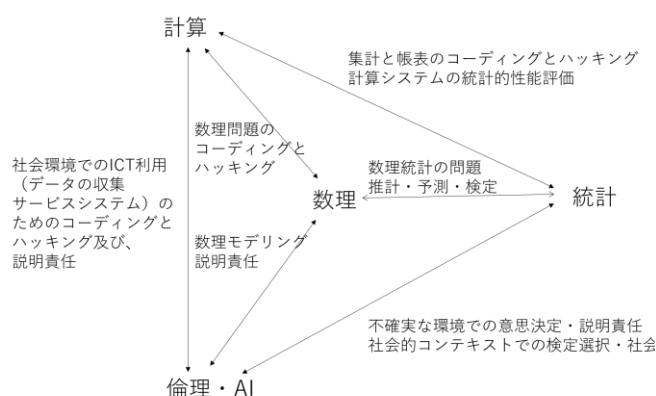


図2 分野間の連結性の検討

表1 数理、計算、統計、倫理・AI の基本要素

分野	基本要素
数理	微積分、行列、ベクトル、数列、集合
計算	離散数学の基礎、デジタル情報、アルゴリズム、プログラミング
統計	データ、グラフ、相関、散布図、分布、回帰、予想、残渣、母集団、分布、推定（最尤法、モーメント法）
倫理・AI	倫理・AI をデータ取り扱いの注意やデータサイエンスの役割

図2の頂点は数理、計算、統計、倫理・AIの4つの分野を示しており、それぞれ表1に示される基本要素を含むとする。

このとき、各頂点を結ぶ辺は、それぞれの分野の複合性を意味する。すなわち、計算と統計の複合性は集計と帳表のコーディング（プログラミング）とハッキング（読み取り解釈）、計算システムの統計的性能評価に対する問題となる。数理と統計の複合性は、数理統計の問題（推計、予想、検定）である。計算と数理の複合性は数理問題のコーディング（プログラミング）とハッキング（読み取り解釈）の問題である。また、倫理・AIをデータ取り扱いの注意やデータサイエンスの役割を基礎要素として、倫理・AIと他3分野との連結分野はそれぞれの要素を利用した説明責任、社会的コンテキストでの手法選択とした（実際には、作問が困難であったことからこの種の複合問題については分類を見送った）。

この検討を踏まえて、DS 発展は、[6]に準拠したスキル要素を含みつつ、数理、計算、統計、AI・倫理および、数理×計算、数理×統計、統計×計算の単一分野問題と二分野複合問題の7つの分野についての問題分類に従い問題を作成し、それぞれの分野問題の回答率から単一能力と二分野複合能力を計測することにより、データサイエンス能力の発展的能力を計量することを基本設計とした。

DS エキスパートについては、高度なデータサイエンス業務の実行能力を評価することを目指し、上位2000名/年程度を目指す先鋭的な人材（課題解決型AI人材）を発掘・伸ばす環境を整備・構築することを目的と設定した。そのため、それぞれの分野ごとの高度な知識を要求し、コンバージェンス型データサイエンスプロジェクトを達成できる分野横断的なデータサイエンス能力を複合的に測定することを目指すこととした。

数理・データサイエンス教育強化拠点コンソーシアムカリキュラム分科会データサイエンス教育に関するスキルセット及び学修項目（第2次報告2021年6月発表）[7]に準拠することにより、データサイエンスの基礎能力を4分野（数学基礎、統計基礎、計算基礎、モデリングと評価）として、その複合問題を小問4題から構成される大問として問題を構成することにより、複合的データサイエンス能力を計量する

ことを基本設計とした。

3 分野および 4 分野複合問題は厳密には分野の連結性が 2 分野問題の連鎖に帰着できる線の問題と、相互に依存性を有する環の問題とに分類される。複数知識を循環的に用いる問題(環の問題)は相互依存性が存在するため線の問題を中心として作問検討を行うことにした。

多くの問題はある局所的な同一分野内知識あるいは 2 分野横断的知識の連鎖として記述できる。このようなことがよく起こる仮説として、多くの問題はプロセス(過程)として記述がなされるというものがある。

プロセスの考え方はフローチャートなどアルゴリズムの考え方の基本となっている。プロセスとは入力と出力と処理の関係である。プロセスは前段のプロセスの出力を後段のプロセスの入力として接続することで、何段でも連鎖を形成することができる。また、プロセスは関数として記述が可能である。

問題を K 個のプロセスに分解できるとする。分解されるプロセスの数 K が多いほど困難な問題と考える。分解された単一プロセスに含まれる内容が高度であるほど困難な問題であると考えている。

高度な内容も、いくつかのプロセスに一般的には分解ができるはずである。そのため、高度な内容と認識される概念も、それをプロセスに分解したときのプロセスの数で測定できる可能性がある。

3. 最適な問題の検討

次に 5)「最適な記号列プログラムの実装」の実装を行うために、2 章で検討した分野ごとの具体的な問題について検討をおこなった。DS 発展と DS エキスパートは CBT(Computer Based Test)で実施されることになったため、基本的に、一択選択問題、複数選択問題、数値記入式問題のみが許される。

このような制限の中で、主として 5 択選択問題によるサンプル問題について例示する。3.1 節で述べる問題例は検討した問題のひとつである。実際には、図 2 で示した 6 種類全ての問題パターンに対して、どのような問題が最適な記号列プログラムの実装となるかについて同様の検討を行った。

3. 1 問題

ある関数 $f(x)$ について、マクローリン展開を考えると、 $N \rightarrow \infty$ として、

$$f(x) = 1 + \sum_{n=1}^N \frac{x^n}{n!}$$

と表現できることがわかり、上式を用いて、 $f(1)$ を計算することを考えた。計算上は、 N を無限大にはできないので、 $N = 10000$ として、 $n > N$ の項は無視することにし、次の Python コードで算出することにした。

```
v = 1
factn = 1
for n in range(10000):
    factn = 【ア】
    v += 1/factn
print('{:.4f}'.format(v))
```

【ア】の選択肢：

- ① factn + n
- ② factn * n
- ③ factn * (n+1)
- ④ factn ** n
- ⑤ factn / n

【ア】に当てはまる適切なものを①～⑤のうちから一つ選べ。また、出力値として適切なものは次のどれか。

解の選択肢：

- ① 4.9348
- ② 3.2899
- ③ 3.1416
- ④ 2.7183
- ⑤ 2.2361

3. 2 問題の解釈

この問題は数学分野におけるマクローリン展開を Python コードとして計算機プログラムとして表現する計算分野の記述能力を用いて、計算結果の推測を数学の能力により問う問題となっている。すなわち、数学×計算の能力を問う 2 分野複合問題となっている。

3. 3 より複雑な複合問題についての発展形

分野の異なる複合知識を用いる線の問題（2分野能力を組み合わせるプロセスの列として適用する問題）で、データに関連する問題は、次の1～4の4つのステップと異なる分野知識にまたがる表現形態を利用することがひとつの典型的な類型である。

1. データの生成(倫理・AI)
2. モデルの選択(統計)
3. 計算の方法(アルゴリズム)
4. 計算の実行(計算・プログラミング)

例えば、2値変数と連続変数を組とする観測データが得られた場合におけるロジスティック回帰の問題を考える。2値変数と連続変数の組みが社会的な分類問題の課題設定(1. データの生成)が存在する。このような変数の組に対する分類問題では、ロジスティック回帰がよく用いられるモデル(2. モデル選択)となる。ロジスティック回帰において回帰係数を求めるアルゴリズムには、対数尤度の算出と2変量ニュートン・ラフソン法による計算手続き(3. 計算の方法)が存在する。更に2変量ニュートン・ラフソン法を、実際に計算するには、コンピュータプログラムへの実装(4. 計算の実行)が必要となる。

4. 作問プロセス

6)「プログラムの実施とそのパフォーマンスのモニタリング」を実施するための作問プロセス設計について述べる。

DS発展の問題策定を行う上で、まずは7種類の問題分類に属する問題アイデアプールを構成し、問題と解答、解説を記述することにより問題試作を作成した。更に、解答の検算と誤植確認、問題のラベリングを行った。これにより、検査済みサンプル問題集を作成した。

次に検査済みサンプル問題を問題種類ごとに問題策定委員の専門性を考慮して配布し、サンプル問題を参考にして問題作成を依頼した。問題策定委員の専門性については、アンケート調査を行った。

問題策定委員の専門領域測定は、図3に示す問題

策定委員と分野との間の二部グラフ構造によるモデル化を行った。このモデル化により二部グラフの連結行列 A をアンケート調査より特定する。すなわち、問題策定委員 j の分野 i に対する担当可否に関するアンケート回答から(2)式で定義するように、担当可能を1、担当しないを0とした連結行列を構成する。

$$A_{ij} = \begin{cases} 1 & (\text{担当可能}) \\ 0 & (\text{担当しない}) \end{cases} \quad (2)$$

連結行列 A は分野数を K 、担当者を M とすると、 $K \times M$ 行列として表現される。

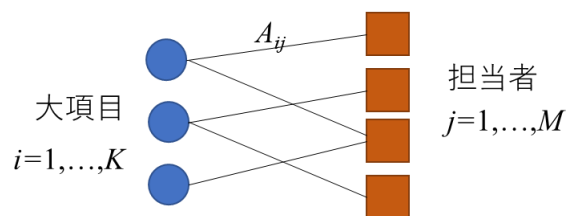


図3 専門領域と問題策定委員との二部グラフ表現

そして、

$$W = AA^T \quad (3)$$

を計算することにより、 W の値は該当単独問題、連結問題の担当可能人数の総数の数え上げとなる。更に、 V の値

$$V = A^T A \quad (4)$$

は問題策定委員間の分野を通じた接続のべ数の数え上げとなる。

アンケートからこの行列の値を確認したところ、全ての W の要素について非零であることが確認できたことにより、問題策定委員会の作問能力は、基礎設計を行った単独問題、複合問題に対する問題策定を行う能力を有することが確認できた。

さらに、データサイエンス能力測定は4種類の単独能力(数理、計算、統計、倫理・AI)と3種類の2分野複合能力(数理×計算、数理×統計、計算×統計)の正解率により、受験者のデータサイエンス能力のバランスがどのように分布するかで判断することとした。7種類の能力強度を「できる」と「できない」の2値分類にした場合に実現する組み合わせは $2^7=128$ 通りである。すなわち、受験者の7種類の回答率が50%を超えているものを「できる」、50%を下回っている場合を「できない」とし、7種類の能力強度の組み合わせから、受験者の7種類能力の強度分

布をみることで、検定結果から受験者のデータサイエンス能力を 128 通りに分類し、評価ができることが分かった。

更に、このようなプロセスにより作成される DS 発展、DS エキスパートの問題はその後の実試験を経て回答率や正解率の観点から問題ごとにそのパフォーマンスの測定が継続して行われるべきである。

そして、これらの問題は定期的に問題パフォーマンスを基に連続的改善を行うことになっており、この連続的改善を行うための問題検討、問題の入れ替えを担当するための問題策定委員会については、引き続き担当可能問題のカバーが継続しているかを継続的にモニターすることが求められる。これが、樁フレームワークの方法の 6 番目のポイント「プログラムの実施とそのパフォーマンスのモニタリング」となる。

5. まとめと今後の展望

本稿では、統計検定データサイエンス発展級、エキスパート級の問題設計の考え方として採用した、樁(2019)が言及するプログラム科学の文法[3]に基づき、どのような手順により検定の設計作業をおこなったかについて報告した。データサイエンス能力を、独立性を有する数学、統計、計算、ドメイン（倫理・AI）の単独能力と複合能力の総体であると定義し、分野間での複合能力を測定するための問題分類と、その達成度によるデータサイエンス能力の測定方法を開発した。

樁は、プログラム科学へのアプローチ法として、QC ストーリーと PDCA とを結合した方法を提案しているが、統計検定データサイエンス発展級とエキスパート級は QC ストーリーと PDCA とを結合したプログラム科学の方法に基づき設計、試作、作成、実施、運用、改善がなされている。

データサイエンスの方法によりデータサイエンス検定を設計できるかという問については、データサイエンス的方法を用いて問題の作成中から問題策定委員の作業のモニタリングを行うとともに、実試験を通じた問題のモニタリングを継続して行っている。

今回多くの問題策定委員の先生方の協力と問題作成への貢献により、統計検定データサイエンス発展級が実施でき、エキスパ級が検討できるに至って

いる。更に、この検定試験が多くの受験者を迎え、わが国のみならず世界中のデータサイエンスの発展に貢献していくことを切望する。

また、この統計検定データサイエンス発展級とエキスパート級がプログラムとして安定的に継続できることを実証的に示すことにより、樁が提唱するプログラム科学の方法論の正しさをも証明することができると考えている。

この検定試験が、長く継続性を有し、多くの方々に利用して頂き、データサイエンス分野の人材育成、人材登用、データサイエンス型の課題解決グループのバランス計測などに産官学でともに利用して頂けるようになればこれ以上の喜びはない。

参考文献

- [1] 統計検定データサイエンス発展(DS 発展), <https://www.toukei-kentei.jp/about/grade12/>, 最終アクセス日 2022 年 2 月 28 日
- [2] 統計検定データサイエンスエキスパート, <https://www.toukei-kentei.jp/about/grade13/>, 最終アクセス日 2022 年 2 月 28 日
- [3] 樁広計, 統計科学の文法と論理的推論, 統計数理, Vol. 67, No. 2, pp. 165—180 (2019).
- [4] Shelley Palmer, Data Science for the C-Suite. Digital Living Press, New York (2015).
- [5] Brian Keng, Data Science Needs and Stress Points, Canadian Data Science Workshop 2018, https://dsg.uwaterloo.ca/CDSW/slides/Keng-cdn_ds_workshop-v3.pdf, 最終アクセス 2022 年 3 月 8 日
- [6] 数理・データサイエンス教育強化拠点コンソーシアムカリキュラム分科会データサイエンス教育に関するスキルセット及び学修項目第 1 次報告, http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_curriculum_2.pdf, 最終アクセス日 2022 年 2 月 28 日
- [7] 数理・データサイエンス教育強化拠点コンソーシアムカリキュラム分科会データサイエンス教育に関するスキルセット及び学修項目第 2 次報告, http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_curriculum_2021_2nd.pdf, 最終アクセス日 2022 年 2 月 28 日

CBT によるデータ対話型アナリティクススキルの測定と評価

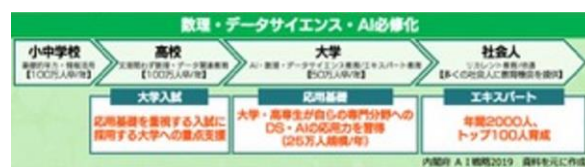
～統計検定データサイエンス基礎試験における実装～

渡 辺 美 智 子(立正大学データサイエンス学部)

1. AI 戦略 2019 でのデータサイエンス教育

社会の AI・デジタル化は急激に進行している。センサー、高性能カメラ、モーションキャプチャー、モバイル IOT などのデータ取得・通信機器の爆発的ともいえる普及で、企業経営、医療・健康、スポーツマネジメント等、社会のほぼすべての領域でデータの利活用と可能性への期待がかつてない高まりを見せている。その中で、デジタル時代の新しい職能として、社会と組織の課題をとらえ、定量的データに基づいて問題を理解・分析・解釈し、そのエビデンスに基づいて問題解決と意思決定の高度化を先導するデータサイエンティストをどう育成していくのか、人材育成と教育システムの構築は、どの国においても喫緊の課題である。

日本政府は、2019 年 6 月に閣議決定された「AI 戦略 2019」において、数理・データサイエンス・AI 必修化として、小中高等学校から大学・大学院、社会人のリカレント教育に至る教育の体系化を示し(図 1)、全ての生徒が高等学校卒業時までに、数理・DS・AI の基礎的なリテラシーを習得することを掲げている。かつて日本人が苦手とする英語教育の重要性から、「100 万人の英語」が謳い文句の時代があったが、まさに「100 万人のデータサイエンス」時代が到来したことになる。



2. 指導要領改訂と高校でのデータサイエンス探究力強化

今回の新学習指導要領で最も大きく変わるのは 2022 年 4 月から実施予定の高校教育である。既にその重要性から共通必修科目として新設された「総合的な探究の時間」が 2019 年から前倒しで先行実施されている。新課程ではさらに、地理探究、数学探究、理科探究、理数探究とほぼすべての教科に探究が位置づけられ、実社会・実生活から自ら見出した課題の探究活動（アクティブラーニング）を通して、生徒の探究能力を育むことが意図されている。

同時に、科学的探究の方法論である、リアルワールドデータに基づく統計・データサイエンスの分析手法の理解と実践スキル習得は、数学Ⅰ「データの分析」、数学Ⅱ「統計的推測」と情報Ⅰ「情報通信ネットワークとデータの活用」、情報Ⅱ「データサイエンス」が連携して担う。具体的に取り扱われる内容は、相関、単回帰、重回帰、クロス集計、仮説検定、信頼区間に加え、文科省が作成し今年度中に公開される情報Ⅱ「データサイエンス」単元の指導書には、ビッグデータ整形処理、

主成分分析、決定木、サポートベクターマシーン、クラスタリング、ニューラルネットワーク、テキストマイニングなど機械学習系手法が Python や R のコードと共に扱われている。

一見すると驚くほどの高度な手法ではあるが、インターフェースに優れたツールの普及で、既にスーパーサイエンスハイスクールなど全国大会では、AI、機械学習、多変量解析を応用した高校生ならではの創意あふれる分析実践例が多数みられる。

3. データサイエンスの国際的枠組みと評価

(1) AP Statistics

1990 年代初頭より、OECD など国際社会は、予測困難な変化に向かう 21 世紀の人材スキルとして、従来の専門知識・技能(ハードスキル)から問題解決力・科学的思考力・コミュニケーション力・連携力などの業種を超えて転用可能なパフォーマンス(行動特性)基盤型ソフトスキルを 21 世型スキルとしてより重視するという方向を示し、初中等から高等教育、社会人教育の枠組みと評価システムの確立を行ってきた。統計教育に関しても問題解決型を指向し、Real Data, Real Problem, Real Learning を軸に、問題解決のサイクル PPDAC: Problem-Plan-Data- Analysis- Conclusion を広く普及させ、1992 年には高校生を対象とした Advanced Placement Statistics の高校 1 年間の教育プログラムと統一試験を始め、現在、毎年 20 万人を超える高校生が現実の文脈の中で統計分析の設定から解釈までを行う問題の試験を受験している。AP Statistics では、計算の公式はすべて与えられ、単純な計算問題や公式を覚えているかどうかなどの問題は出題されず、ブートストラップなどのリサンプリングや仮説検定、

ベイズ統計の考え方を理解し応用する問題が統計研究者によって開発・出題されている。

AP 統計が新しい統計能力の評価として登場してから 30 年近くが経過し、実際に 21 世紀を迎え冒頭で述べたデジタル化社会への変革が迫られている現在、AI とアナリティクスの普及は、専門職能とその基礎として、分析結果の解釈方法とともにビッグデータ・スモールデータの収集方法と解析方法、データハンドリング・解析アルゴリズムの機能を理解し問題解決を通して社会の価値に結び付けるデータサイエンスのパフォーマンス評価の段階に至っている。

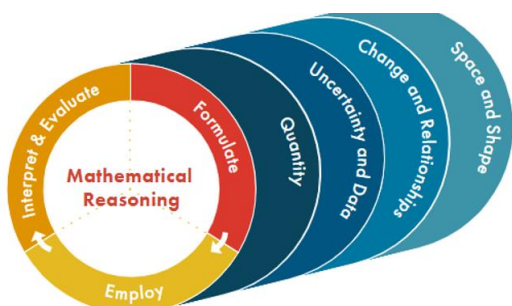
(2) PISA2021 数理リテラシーの新フレーム

OECD が実施する生徒の学習到達度調査 PISA は、義務教育修了時点(15 歳)で身に付けてきた知識や技能を実生活の課題解決に活用できるかを評価する試験として 2000 年から 3 年ごとに、読解力、数学リテラシー、科学リテラシーの主要 3 分野で実施されている。この時点から世界的に、「学力」論は知識・技能を使いこなす「資質・能力」の育成へと変化した。PISA2015 では、先述の 21 世型スキルも、21 世型学力(世界共通の力)として導入され測定が行われた。昨年末に、PISA 2018 の結果が日本でも公表され、「数学的応用力」で 6 位、「科学的応用力」では 5 位と数学、科学はトップレベルを維持したと報道されたが、果たしてこれが次回も達成できるのか。

OECD は既に、劇的にデジタル化した現代社会の変容に合わせ、次回 2021 年の PISA 数学の評価フレームを数学リテラシーの定義、内容、測定方法とも大幅に改訂したことを発表している。主な変更点は、データサイエンスを中心に据えたこと。測定方法も表計算機能を使って段階的にデータ分析的思考で問題を解いていく Computer

based assessment (CBA) 方式に本格的に変更した。内容領域は「数量」「不確実性とデータ」「変化と関係」「空間と形」で、これまでトップレベルと言われながらも領域的に日本が弱かった「不確実性とデータ」の部分が強調されている。AI時代の予測モデルの基礎である散布図や相関、関数の適合、集計から得られる条件付確率を基に意思決定するなどの問題がコンピュータベースでインタラクティブに課される。数式を中心とした旧来の紙ベースのセンター試験の形式では測れない、データ対話型のアナリティカルな思考とスキルである。表計算ソフトは、米国の調査では6割強の大人がデータ処理に日常で活用している。子供の未来のために、入試改革も内容だけではなく評価方式もアナリティクスを組み込むなどのデジタル化も急務となっている。日本では、PISAは高校1年の春に実施されているが、実はPISA2021の散布図・相関・クロス集計表から条件付確率を推測して意思決定する等は、次期の指導要領でも高校生の内容である。つまり、義務教育段階の数学(統計)内容が改訂しても既に世界とずれているのである。

更に、数学(統計)内容は、具体的な文脈(身の回り、職業、社会、科学)において問題解決サイクルのプロセス:「定式化」→「活用」→「解釈と評価」の中に組み込まれる形で問題が開発されており、プロセスごとの評価もできるようになっている。



(3) IDSSP (International DS in School Project)

2018年に、国際統計協会や米国統計学会、王立統計学会、国際コンピュータ科学学会(ACM)、オーストラリア統計局、Googleなどの支援で、AP統計受講後の高校の2年間分のデータサイエンス入門コースのカリキュラムチームIDSSPが国際的に組織されているが、2019年9月に公開されたガイドラインでも、先ず、データに基づく学習(知識獲得)のサイクルを理解し、具体的な課題に応じてこのデータサイエンスサイクルをまわす



ために、5つの各ステップでどのようなスキルが必要とされるのかを実践的に学ぶ構成が示されている。

4. 統計検定 CBT-データサイエンス基礎

これまで述べたように、社会のデジタル革新が急速に進行する中で、データサイエンス人材の育成は喫緊の課題であり、その際、日本においても規模の拡大だけではなく、国内外のガイドラインに沿ったデータサイエンス人材の質評価のシステムを確立する必要がある。このため、一般財団法人統計質保証推進協会は、統計検定に新たに CBT『データサイエンス』を創設し、来年度から順次、「基礎」、「発展」、「応用」の3水準でその能力評

データサイエンス基礎

表示サイズ 100%

1問目 / 全2問

□あとで復習す

エクセルデータシート「アンケートリゼ11」は、ある不動産会社で取り扱っている中古住宅17戸の価格と属性のデータである。1行目は表頭で、()内は、それぞれの変数の単位を表している。2行目以降がデータの値である。データを分析して、下記の問いの空欄に適切な数や文字を入力せよ。

[1] 木造住宅と鉄骨住宅、それぞれの住宅の価格の平均を求めた、木造住宅の一戸当たりの平均価格は、 百万円となる。※小数第2位までの値を半角で入力すること。

[2] 価格を床面積の価格に変換し、木造住宅について、築年数によって一坪 (3.3m²) 当たりの価格がどのように変化するのを見てみることにした。そのため、先ず、築年数と床面積一坪当たりの価格の関係を散布グラフを作成する。次に、そのグラフに、回帰直線をあてはめることにした。このようなグラフを作成できずか、一般的な統計グラフの名前を誤って記入し、

[3] 当てはめた回帰直線の式から、築年数が1年減ることにより、床面積一坪当たりの価格は、 万円ずつ、変化する場合にあることがわかった。※小数第2位までの値を半角で入力すること。

[4] この単回帰直線上の決定係数と残差の総和の値 (残差) が最も大きな木造住宅の番号は、 である。※番号の番号を半角で入力すること。

[5] この単回帰直線が築年数の変数に十分に表しているか否かを調べる。下記の①～④から、先の問4の住宅の価格が大きくなった原因として最も適切なものを一つ選ぶ。

※適切な理由を半角で入力すること。

① 住宅の床面積
② 住宅の価格 (価格、木造)
③ 住宅の築年数
④ 駅からの距離など、①～③の変数以外の要因



統計検定 CBT 「データサイエンス基礎」

解答形式

- 択一問題** : 解答番号をキーボードで入力する方式の問題は、表示された選択肢①～⑤から解答番号を半角数字で入力してください。
- (例: 解答が③の場合は、半角で数字の 3 を入力)
- 計算結果等数値で解答する問題** : 計算結果等を直接、数値で入力する問題は、
- * 0 または正の整数の場合、その値を入力
 - * 小数を含む場合、小数第 3 位を四捨五入して、小数第 2 位までの値を入力
 - * 上記のいずれも、負の値の場合は、マイナスの符号を含めて入力
- (注意: 入力する数字や符号は半角の文字を使用すること。)
- 用語を解答する問題** : 漢字、ひらがな、カタカナの場合は、**全角で入力**。アルファベットの場合は、**半角で入力**。

解答・分析操作のヒントと求められる知識・能力

【問題 1】

- [1] 木造住宅に限定して、AVERAGE 関数などで計算する。

(平均の意味理解と文意に沿った計算処理能力)

- [2] 散布図

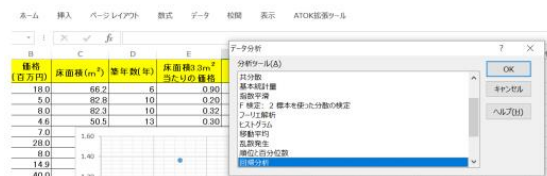
(分析目的: 単回帰予測の意味理解と関連する統計グラフを選択する能力)

- [3] 価格と床面積のデータから床面積 3.3m² 当たりの価格に変換した列を作成する。

A	B	C	D	E	F
住宅番号	価格 (百万円)	床面積 (m ²)	築年数 (年)	構造	床面積 3.3m ² 当たりの価格 (百万円)
1	18	66.2	6	木造	=B2/C2*3.3
2	5	82.8	10	木造	
3	8	82.3	10	木造	

回帰直線式 (回帰係数) を求める方法としては、

- (1) 散布図 (X 軸: 築年数、Y 軸: 床面積 3.3m² 当たりの価格) を作成し、グラフ上で「近似曲線の追加」、「線形近似」、「数式を表示する」を実行する。
- (2) 「データ」メニューの「データ分析」から、「回帰分析」を選択する。



等がある。

(変数変換式の意味理解と実際のデータに対する数的処理能力。単回帰式の意味理解および実行処理、出力結果を解釈し推論する能力)

- [4] 残差を読み取る方法としては、
- (1) 散布図から残差を読み取る。
 - (2) 「データの分析」→「回帰分析」のメニューで「残差」をチェックする。
- 等がある。

(残差の意味理解およびグラフや出力結果を読み取る能力)

- [5] 4
- (残差の解釈と活用能力)

[問題2]

ピボットテーブル等で、題意に沿ったクロス集計表を作成し考察する。

- [1] 199人 (文意に沿ったクロス集計の数的処理能力)

- [2] 71.20% (文意に沿った割合の意味理解と数的処理能力)

- [3] 7 (クロス集計の結果から変数の関連性を解釈し推論する能力)

- [4] [4-1] 1 [4-2] 6

(3変数以上のクロス集計の結果を読み取り推論する能力)

統計検定2級、3級、4級の創設時は、現行学習指導要領の告示に合わせて、問題解決型の統計思考力評価問題を開発・公表し、その後、センター試験問題、教科書・参考書・問題集、大学入試問題等に参照された。今回は、AI・デジタル社会を見据えた新学習指導要領であり、大学入試センター共通新テスト始め、多くの全国的な試験が将来

的にはCBT化されることが予想される。また、そうならなければ、データ中心のコンピュータ対話型試験は実行できず、真に測らなければいけない能力が測滴できない。

海外では、AIやデータ解析、Pythonなどの単純な知識を問う、4択や5択のクイズ形式のオンライン試験や資格は、有償・無償を含め多くネット上に存在し、簡単に試験が開始できその場で合格証が手に入る。これらは入り口に過ぎない。

OECDのPISAやAP、IDSSP、日本の大学入試改革などは、問題開発に困難は伴っても、学協会経験と知識、ICTの技術を使って、今後求められる高度に知識処理ができる人材の質保証にチャレンジしている。AI・データサイエンスは社会実装化が大学等高等教育の現行課程を追い越し、はるかに速いスピードで進んでいる。産学官の連携は、評価問題の開発においても必至である。

統計データ分析コンペティション 2021 の概要とデータ分析実践事例

堀川 顕子・独立行政法人統計センター

太佐 美結・フェリス学院高等学校（1年）

〒162-8668 東京都新宿区若松町 19-1

TEL：03-5273-1368

E-mail：statcompe@nstac.go.jp [@は半角]

本報告では、統計データ分析コンペティション及び SSDSE の概要等（1～3）について解説するとともに、2021 年優秀賞受賞者（4(2)）による分析事例の紹介を行う。

このセッションでは、本報告とは別に、総務大臣賞受賞者（4(1)）及び統計活用奨励賞受賞者（4(4)）の報告も行われる予定。

1. 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。昨年は、新たに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。

2021 年のポスター（2022 年も 4 月公開予定）

統計データ分析コンペティション 2021

高校生、大学生等の皆さんの統計分析におけるアイデアと技術を競います

論文募集

エントリー期間
令和3年5月10日(月)～8月10日(火)

最優秀作品には総務大臣賞と副賞が贈られます
このほか、奨励論文が専門誌等に掲載されます

論文締切
令和3年9月1日(水)
令和3年9月10日(金)

第4回統計データ分析コンペティション
詳細は、下記ウェブサイトをご覧ください
<http://www.nstac.go.jp/statcompe/>

共催 総務省統計局・独立行政法人統計センター
大学・研究機関関係人 情報・システム研究機構 統計数理研究所・一般財団法人日本統計協会
文部科学省・独立行政法人 科学技術振興機構 (JST) 全国高等学校長協会・一般財団法人日本統計学会・全国統計教育協議会

2. 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、2021 年版として以下の 4 種類が無償ダウンロードできる。今年も 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-A:

市区町村別、多分野データ（項目数 125）

SSDSE-B:

都道府県別、時系列、多分野データ(105)

SSDSE-C:

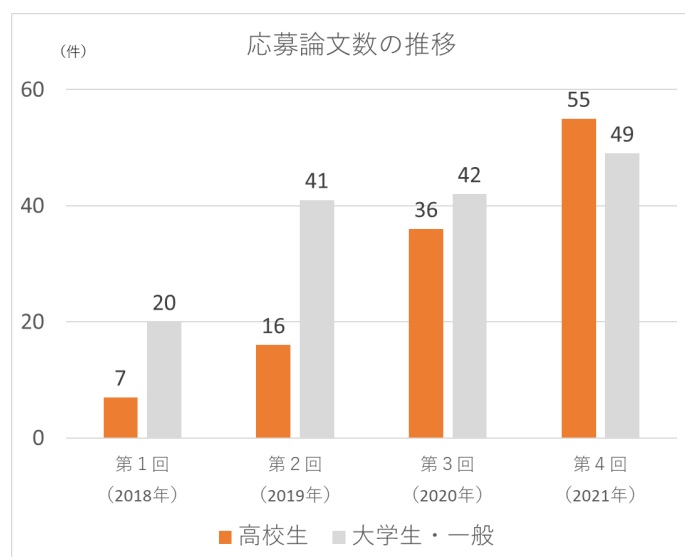
都道府県庁所在市別、家計消費データ(226)

SSDSE-D:

都道府県別、自由時間活動・生活時間データ(119)

3. 統計データ分析コンペティションの応募状況

昨年 2021 年の応募論文数は、高校生の部 55 本、大学生・一般の部 49 本であり、この応募論文数は年々増加している。



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

村澤 舞・山家 里穂（広島大学附属高等学校）

「日本におけるワークライフバランスの達成状況とその課題」

ワークライフバランスについて、生活時間を男女別に分けることで男性はワーク（仕事）とライフ（生活）が負の相関にあること、女性はワークとライフに育児時間が強く影響しており、ワークを交絡変数として偏相関分析を行うことで、育児時間とライフの時間が負の相関にあることを示した。

(2) 優秀賞

太佐 美結（フェリス女学院高等学校）

「健康寿命の延伸に向けて」

人が健康に暮らすことのできる健康寿命について、様々な変数を用いて因子分析を行うことで因子の抽出を行い、説明変数の選択に当たっては相関分析により多重共線性について考慮した上で重回帰分析を行うことで、男性はスポーツ行動率、

女性は緑茶の消費量等が健康寿命に影響を与えることを示した。

(3) 統計数理賞

森 颯太（香川県立高松商業高等学校）

「求められている住宅」

空き家が増える原因を分析するために不動産価格予測 AI を作成し、住宅の条件を学習データとして用いることで、住宅の面積や築年数、最寄駅からの徒歩時間などの重要度が高いことを示した。結論として、住宅のリノベーションやリフォームに取り組むことによって空き家を減らすことを提案している。

(4) 統計活用奨励賞

谷 優輝（慶應義塾湘南藤沢高等部）

「外国人にとっての暮らしやすさとは」

日本の少子高齢化による人口減少を解決するための足掛かりとして外国人の移住に着目し、都道府県別の外国人人口割合について相関分析を行うことで、快適な気候や日本語教育の質に加えて、工業が盛んな地域ほど外国人人口割合が高い傾向にあることを示した。

(5) 特別賞

特別賞（統計分析） 2 本

特別賞（統計活用） 3 本

特別賞（審査員奨励） 1 本

【参考URL】

・統計データ分析コンペティション 2021

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

・SSDSE（教育用標準データセット）

<<https://www.nstac.go.jp/SSDSE/>>

・総務省統計局報道資料

<<https://www.stat.go.jp/info/guide/public/houdou/pdf/ho211018.pdf>>

統計データ分析コンペティション 2021 に参加して

村澤 舞・広島大学附属高等学校（1年）
山家里穂・広島大学附属高等学校（1年）
橋本三嗣・広島大学附属高等学校（指導教員）
〒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

私たちは日本におけるワークライフバランスの達成状況とその課題について統計データを用いて研究しました。

このテーマを研究した理由は、個人的な経験が2つあり、1つは父親の仕事量が多いこと、もう1つは学校の先生方が生徒への熱心な指導のため長時間働いていることです。

分析の結果、男性は女性より仕事時間が長いこと、女性は男性よりも家事・育児の時間が長いことが分かりました。

このような男女の働き方の違いに着目して、男女に分けて分析を行いました。男性はワーク（仕事）とライフ（生活）の間の負の相関がありましたが、女性はワーク（仕事）とライフ（生活）の間に明確な相関がありませんでした。そこで女性の場合は育児の負担が大きく、ワークを交絡変数として育児時間とライフの間に負の相関があることを示しました。これは育児時間を交絡としてワークとライフの間に負の相関があると考えられることもできます。

最後に重回帰分析を用いて生産年齢人口の割合（+）、第三次産業の人口の割合（-）、販売職の割合（+）が仕事時間に影響を与えていることを確認しました。

これらの分析から男性の場合は平日の働き方を正面から改革する必要があるとあり、女性の場合は伝統的な性別役割分業観の影響から育児が過剰負担で、育児負担の分散やさらなる職場進出を進めていく必要があることが分かりました。

2. 研究を通して感じたこと

夏休みに入る前に研究テーマを決め、夏休みの間に学校や自宅で研究に取り組みました。

データを使った分析を行うのが初めてで、SSDSEの膨大なデータを扱うことが最初の壁でした。データを大きい順番に並べるだけでも一苦労でした。最終的に

はそれらの処理にも慣れて数字にとっても強くなりました。研究を進めていく中で、数字で客観的に分析できることが統計の強みだと思いました。

画像1 分析に用いたデータ

U	V	W	X	Y
男性正規 仕事	娯楽		趣味	趣味+娯楽
550	83		30	113
533	84		27	111
515	92		25	117
535	74		28	102
528	77		26	103
526	72		41	113
550	82		22	104
557	81		21	102
541	80		32	112
575	65		21	86
547	59		23	82
548	64		31	95
555	61		26	87
574	62		19	81
558	74		26	100
537	79		29	108

画像2 相関係数

A	B	C
	列 1	列 2
列 1	1	
列 2	-0.66711	1
男性正規労働時間と趣味+娯楽		
	列 1	列 2
列 1	1	
列 2	-0.49041	1
男性正規労働時間と趣味		
	列 1	列 2
列 1	1	
列 2	-0.54071	1
男性正規労働時間と娯楽		

3. 統計を扱う上で大事だったこと

統計で分析を行う上で特に大事だったことを2つ挙げます。

一つ目は分析に適切なデータを使うことです。研究したい内容に即したデータを使わなければ、正確な分析ができなくなってしまいます。今回の研究で言えば、ワークとライフの相関を分析する時に SSDSE から社

会生活基本調査に切り替えたことでより詳細に実情を把握することができました⁽¹⁾。

二つ目は私たちの論文の内容は覆されるかもしれないし、それは決して悪いことではないということです。これは統計に限らず論文全体に当てはまることだと思います。学校である先生の講演を聞いた時に、「研究内容が覆ることもある。」とおっしゃっていて、今回の研究で示したと矛盾するデータが見つかって、論文の内容が覆ったらどうしようと考えていました。それに関して先生に質問すると、「研究できる時間や費用に限度があるのだから、完璧な論文を求めるのではなく、使える時間や費用の中でできることを最大限やるのが大切だ。」とおっしゃっていました。だから今回の論文の内容も覆される可能性もありますが、悲観的な必要はないと思いました。

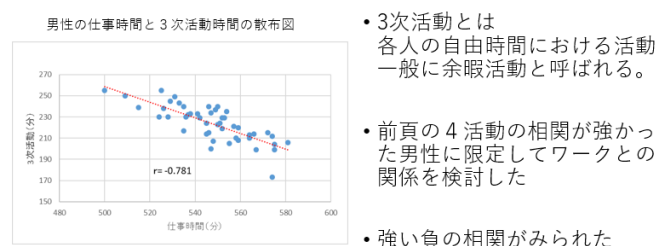
4. まとめ

今回の研究でデータの処理や論文の書き方が身につけてとても貴重な経験になりました。

今回分かったワークライフバランスの達成状況が労働環境の改善や仕事時間の短縮につながることを切に願っています。

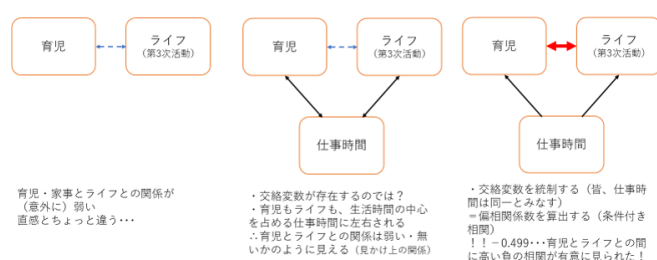
画像3 理数系教員統計・データサイエンス授業力向上研修集会（愛媛）で使用したスライド1

ワークとライフの関係



画像4 理数系教員統計・データサイエンス授業力向上研修集会（愛媛）で使用したスライド2

交絡の可能性



また理数系教員統計・データサイエンス授業力向上研修集会（愛媛）では研究内容の発表の機会を設けていただき大変勉強になりました。発表の仕方の勉強になったり、発表後の質問に答えるうちに新たな視点を持てたのはもちろんのこと、他の高校の研究内容も聞けて、私たちの手法とは異なった研究方法や着眼点、グラフや表の見せ方を知り、統計にさらなる可能性を感じました。

広島大学附属中・高等学校のホームページ⁽²⁾や機関誌「むつみ」に今回の研究が紹介されたことをとても光栄に思っています。また、ロングホームルームの時間に講堂で同学年の約200人に向けて研究内容を発表しました。たくさんの人に研究内容を詳細まで知ってもらえるのが嬉しかったのと同時に、たくさんの人が見ている中での発表は緊張しました。たくさんの人が見ている発表の難しさが分かり、そのような場に慣れたいと思いました。

さらに、中国新聞⁽³⁾にも取り上げていただき、「新聞読んだよ!」と多くの方に声をかけていただけてとても嬉しかったです。

最後にこれまで論文の作成を指導してくださった橋本三嗣先生に感謝申し上げます。

5. 参考文献・参考 URL

- (1)統計データ分析コンペティション 受賞論文
(2021) 2022年2月21日閲覧
<https://www.nstac.go.jp/statcompe/award.html>
- (2)広島大学附属中・高等学校 広島大学附属高等学校の生徒が「統計データ分析コンペティション 2021」高校生の部で総務大臣賞を受賞しました
2022年2月21日閲覧 https://www.hiroshima-u.ac.jp/fu_midori/news/68284
- (3)中国新聞 男女の仕事時間に着目、統計コンで総務大臣賞 2022年1月9日掲載

6. 資料

画像1 分析に用いたデータ

画像2 相関係数

画像3 理数系教員統計・データサイエンス授業力向上研修集会（愛媛）で使用したスライド1

画像4 理数系教員統計・データサイエンス授業力向上研修集会（愛媛）で使用したスライド2

統計活用奨励賞及び特別賞の授業実践事例 (統計データ分析コンペティション 2021)

馬場 国博・谷優輝・河又杏香・遠藤沙恵
慶應義塾湘南藤沢中・高等部
〒252-0816 藤沢市遠藤 5466
TEL (0466)49-3585 FAX (0466)47-5077
E-mail: baba@sfc-js.keio.ac.jp

1. はじめに

本校の高校3年生を対象にした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、今年度は馬場(発表者)が担当し、履修者数は20名であった。

授業の前半(4月～9月)は、次の2つのコンテストのいずれかへの参加を通して、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とした。

A.「統計データ分析コンペティション 2021」(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)：11名が応募

B.「第13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13名が応募

本発表ではAへの実践事例について報告するが、Bでは本選(高校生チーム2組、大学生チーム8組)に1組3名が進み、そのチームが最優秀賞を受賞することができた。

授業の後半(9月～11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。

2. 「統計データ分析コンペティション」に向けての取り組み

このコンペティションは、都道府県や市町村、県庁所在市別の公的統計を集めた4種類の教育用標準データセット(SSDSE)を用いた統計分析によりアイデアと解析力を競うもので、A4用紙で2～8ページに論文をまとめるものである。

(1) 授業展開

4月～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、箱ひげ図、散布図など)、重回帰分析、SSDSEのRへの取り込み方を扱った。

6月には週2時間×4回の授業時間を各自の解析、論文作成の時間に充て、適宜助言を与えた。その後、論文の中間提出を受け、7月に個別指導を行いながら返却。夏休み中に各自で論文作成に取り組んでもらい、8月末日にはほぼ完成形で再提出。助言を受けて最終的に論文を完成させた。

(2) 応募論文

11名の応募者のうち、受賞した3名(発表者)を除く応募者の論文タイトルは以下の通り

- ・都道府県別失業率と産業
- ・運転免許自主返納と高齢化社会
- ・がんと生活習慣
- ・人口移動モデルと少子化対策—「東京一極集中の是正」は適当か
- ・日本における中学校生徒の学力の解析—英語分野の成績向上の要因について
- ・学生の自殺といじめの関係
- ・学校数の増減から見る大学の存在意義
- ・人間関係による子どもの変化

身近な問題から社会的課題への問題意識など多分野のテーマが扱われた。選に漏れたものの中にもレベルの高い論文も多く、各自なりにデータと格闘したことが伝わってくるものばかりであった。

応募した生徒諸君の頑張りのおかげで、2年連続で学校表彰していただけたことは、今後の励みにつながるものである。

(3) 反省点

コンペティションへの参加について、履修者からは以下のような感想が寄せられた。

- ・統計に関する理論と実践を行えてとても良かったです。いろいろな統計の手法を学んだ今思い返すと、SSDSEでもっとこうすれば良かった、と思うようなことがたくさんあります。(同様な意見が複数あり)

家庭や地域の実態調査を取り入れた

統計的問題解決の実践について

- 総合的な学習の時間やSDGsを関連させた取組 -

令和4年3月19日
愛知県豊田市立寿恵野小学校
加藤 孝児

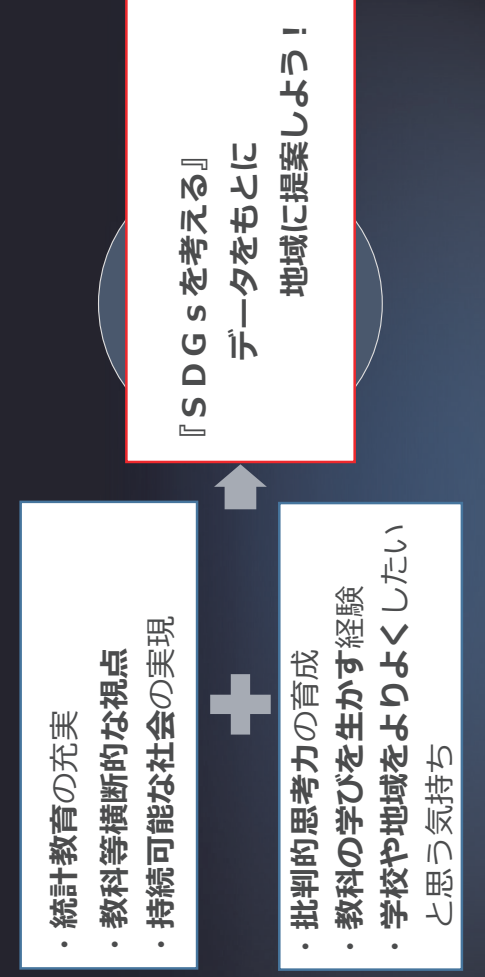
1 社会の要請



2 子どもの実態



3 単元の構想



3 単元の構想

算数科

総合的な
学習の時間

3 単元の構想

<主な手だてについて>

手だて①

「スキルBOX」の活用

手だて②

分析方法や結論を検討するための対話

3 単元の構想

<主な手だてについて>

<6年生>

<1～5年生>

3 単元の構想

<主な手だてについて>

手だて③

SDGs や地域を関連させた単元構想

手だて④

VW を取り入れた実態の数値化

3 単元の構想

<主な手だてについて>

VW（バーチャルウォーター）とは、食料を輸入している国において、もしその輸入食料を生産するとしたら、どの程度の水が必要かを推定したもの。（環境省より）



（例）牛肉100gのVWはおよそ2060 L。
牛が飲む水に加えて、牛が食べる飼料を育てる水が必要になる。

4 研究の実際と考察

（1）SDGsやVWを知り、現状を変えたいという気持ちが芽生える（手だて③④）
【総合：第1～7時】

日常生活の中から、

「何かしたい」と思える

課題を提示したい！

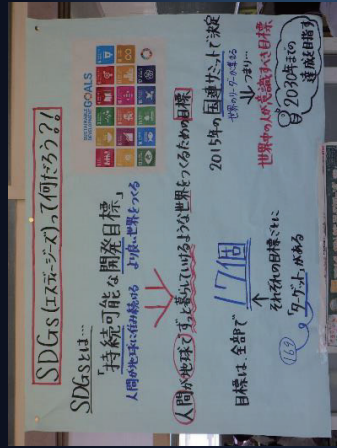
聞いたことはあるけど…

内容はよく知らないな



4 研究の実際と考察

（1）SDGsやVWを知り、現状を変えたいという気持ちが芽生える（手だて③④）
【総合：第1～7時】



4 研究の実際と考察

（1）SDGsやVWを知り、現状を変えたいという気持ちが芽生える（手だて③④）
【総合：第1～7時】

資料6 A児の振り返り（総合的な学習の時間）

① 日本が抱える課題を考えた後①
自分が気づけていない課題もあったので、取り組むことができることからやっけていき、2030年までに達成できるようにがんばりたい。

② 日本の達成状況を調べた後②
思っていた以上に緑が少なく、赤やオレンジが多かった。それに、何も問題がないと予想していたところは、安全なところもあるしやばいところもある。なので、本当に他人事じゃないし、自分も国も積極的にやっけていかないと間に合いません。

※ 達成状況：緑…◎ オレンジ…△ 赤…×

他人事ではなく

自分事として現状をとらえる

自分にできることから
やっけていきたい！

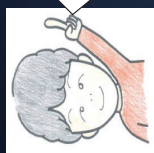
4 研究の実際と考察

(1) SDGsやVWを知り、現状を変えたいという気持ちが生える (手だて③④)
【総合：第1～7時】

個人調べ

食品ロス問題への興味

VWの導入



教師が1週間で
買った食料を
VWに変換したものを
提示すると

自分の家の
1週間のVWも
調べてみたい!!

4 研究の実際と考察

(1) SDGsやVWを知り、現状を変えたいという気持ちが生える (手だて③④)
【総合：第1～7時】

資料7 A児の振り返り (第6時)

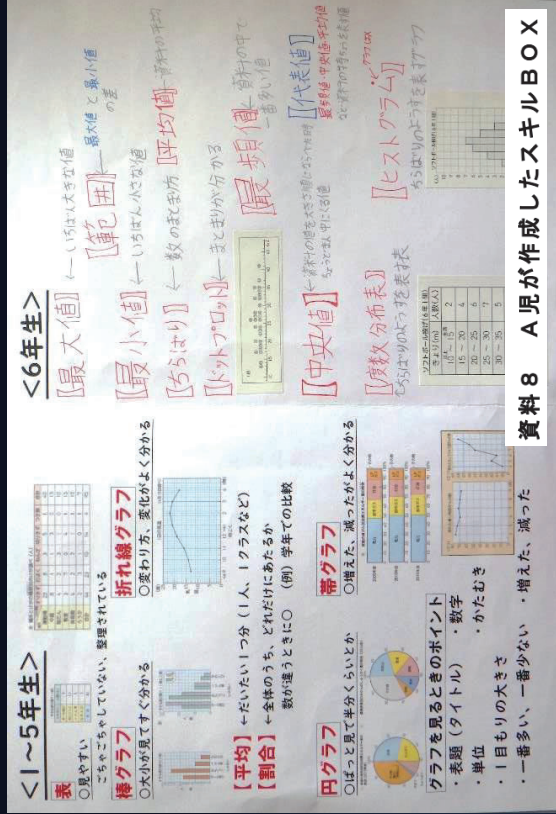
何かしないといけないという
切実感が芽生えた

日常生活の問題解決に
算数科の学習を生かしたい

いないから何とかしないかと思った。SDGsとVW、食品ロスとVWの関係を知って、算数で勉強することを使って自分にできることを考えたい。

4 研究の実際と考察

(2) 「スキルBOX」を完成させる (手だて①) 【算数科：第1～6時】



4 研究の実際と考察

(2) 「スキルBOX」を完成させる (手だて①) 【算数科：第1～6時】

資料9 A児の振り返り(第1時)

5年生までに学習したことを振り返って
て、忘れていた言葉やグラフなどを見るとき
ポイントが意外ななと思いました。これか
らの学習でスキルBOXを完成させて、いろい
ろな見方で考えられるようになります。

4 研究の実際と考察

(3) ソフトボール投げの記録を分析する (手だて②) 【算数科：第4～6時】

6年1組の記録			6年2組の記録			6年3組の記録			
番号	きより (m)	番号	きより (m)	番号	きより (m)	番号	きより (m)	番号	きより (m)
①	27	⑮	23	①	22	⑮	28	①	14
②	17	⑯	20	②	18	⑯	23	②	24
③	20	⑰	35	③	30	⑰	34	③	29
④	22	⑱	14	④	37	⑱	18	④	14
⑤	17	⑲	33	⑤	22	⑲	33	⑤	38
⑥	32	⑳	30	⑥	28	㉑	27	⑥	24
⑦	27	㉑	26	⑦	28	㉒	19	⑦	33
⑧	18	㉒	28	⑧	19	㉓	27	⑧	24
⑨	34	㉓	35	⑨	31	㉔	34	⑨	38
⑩	41	㉔	13	⑩	33	㉕	17	⑩	40
⑪	24	㉕	26	⑪	25	㉖	36	⑪	19
⑫	28	㉖	19	⑫	32	㉗	23	⑫	25
⑬	32	㉗	24	⑬	24			⑬	40
⑭	37	㉘	26	⑭	21			⑭	33

4 研究の実際と考察

(3) ソフトボール投げの記録を分析する (手だて②)

最頻値で判断していたA児

B児「平均値で比べるのが一番いい」

平均値で考え直すA児

D児の平均値に対する指摘を受ける

C児「どのクラスが一番かわからん」

資料10 授業記録1(第4時グループでの対話)

A児：私は最頻値で考えただけで、最頻値が一番高いのは2組だから、2組じゃないかなと思います。

B児：僕も最頻値は見ただけで、平均値で考えた方がいいなと思います。平均値で考えると3組が一番いいから、3組だと思っよ。

C児：でも、最大値で考えたら1組じゃない。

B児：確かにそうだけど、月例テストで学年で勝負したときも平均点で比べたじゃん。だから、平均値で比べるのが一番いいと思う。

A児：確かに。平均値で比べることによってよくあるね。平均値で考えたほうがいいのかな。

D児：でもさ、平均って例えば有り得ないと思うけど、0, 0, 0, 0, 0, 5米とかでも、ものすごく高くなっちゃうよね。平均値だけで見ていいのかな。

C児：ああ、何だかややもやしてきた。これだと、どのクラスが一番かわからんじゃん。

4 研究の実際と考察

(3) ソフトボール投げの記録を分析する (手だて②) 【算数科：第4～6時】

資料11 A児の振り返り(第4時)

(前略) B児の「平均値が高い方が良い記録」という意見はよく分かったけど、C児と同じで答えが見つからず少しもやもやしています。

はつきりとした結論には辿り着かなかったが、自分の考えは「本当に正しいのか」と考え直す内省的な考察は十分にできたといえる

4 研究の実際と考察

(4) 1週間の買い物調べのデータを分析する (手だて①②) 【算数科：第9，10時】

資料 13 授業記録 2 (第10時グループでの対話)

B児：国産と外国産の差が大きいということがいえます。他にはどう？

A児：ドットプロットのことでもいい？ドットプロットで考えると、この辺りにかたまっているよ。平均値とかその辺りが多くなっているよ。

D児：最小値は4000Lだけど、それでもお風呂20杯分はあるよね。

B児：一番少ない人でもそんなにあるんだもんね。

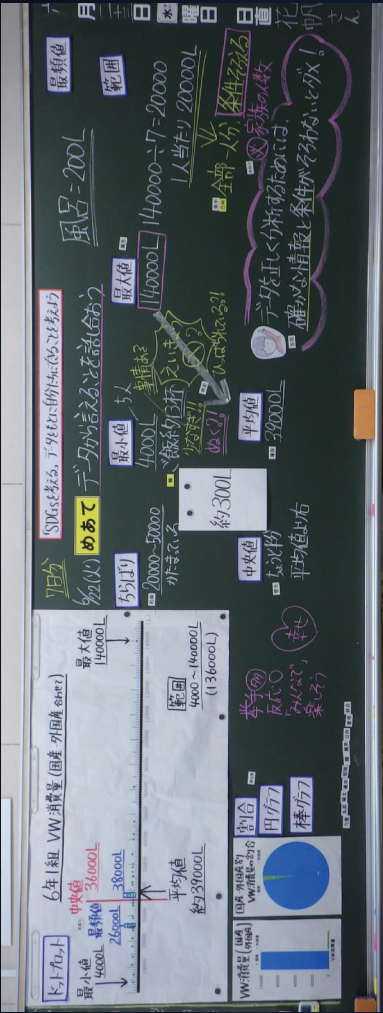
A児：ああ、最小値か。そこはあまり見てなかった。

D児：うん、最小値で考えても多いと思う。

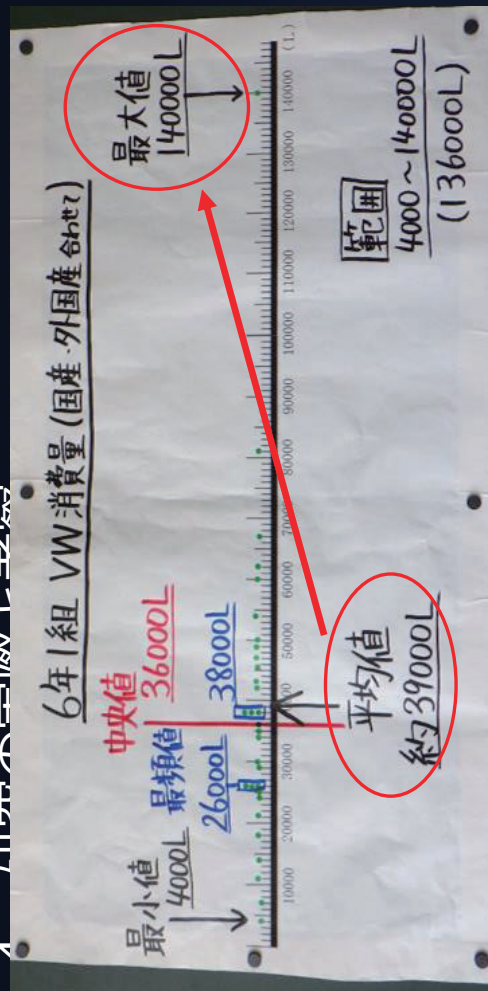
D児の最小値に関する発言を聞き、最小値に関わる内容を追加で記述
→別の観点を取り入れる

4 研究の実際と考察

(4) 1週間の買い物調べのデータを分析する (手だて①②) 【算数科：第9，10時】



資料 14 授業記録 3 (第10時全体での対話)



4 研究の実際と考察

(4) 1週間の買い物調べのデータを分析する

**F児の発言をきっかけに
最大値の不自然さを感じる
A児は条件を揃えて分析する
必要性に気付いた**
→ **多面的・内省的な考察**

資料 14 授業記録 3 (第10時全体での対話)

F児：中央値よりも平均値が大きいので、最大値の14万に引っ張られていると思います。

T：今のF児の、平均値が最大値の14万に引っ張られているという考え、わかる？グループで確認してみよう。

(グループで確認後)

T：確認できた？何かさ、最大値の14万が大きすぎで気になるって声が聞こえてくるんだと思う？

G児：家族の人数が多いからだと思います。

B児：家族の人数を考慮していないから、本当の値が分らないと思います。

多数：ああ。確かに。そういうことか。

T：じゃあ、どうしたらいいの？

H児：家族の人数で割る、一人分に揃えればいい。

：A児は、今H児が言ったことわかる？

A児：全員を一人分にして、条件を整えないといけないということだと思います。

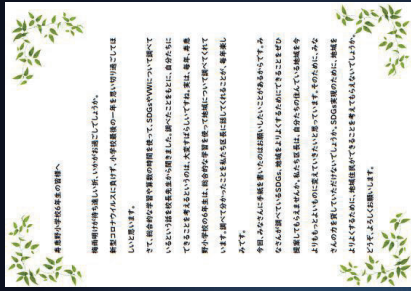
4 研究の実際と考察

(5) 自分たちや地域にできることを考える (手だて③④) 【総合：第8～17時】

算数科で1週間のVWVのデータを分析

自分たちにできることを考える

区長さんからの手紙



4 研究の実際と考察

(6) 近隣店舗の廃棄量が多い日の傾向を分析する (手だて①②) 【算数科：第16, 17時】

セブン-イレブンのある店舗における日ごとのごみの廃棄量(7月)

日付	曜日	前日	当日	7月12日	7月13日	7月14日	7月15日	7月16日	7月17日	7月18日	7月19日	7月20日	7月21日	7月22日	7月23日	7月24日	7月25日	7月26日	7月27日	7月28日	7月29日	7月30日	7月31日
1	月		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	火	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	水	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	木	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	金	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	土	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	日	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	月	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	火	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	水	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	木	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	金	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	土	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	日	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	月	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	火	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	水	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	木	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	金	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	土	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	日	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	月	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	火	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	水	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	木	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	金	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	土	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	日	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	月	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	火	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	水	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

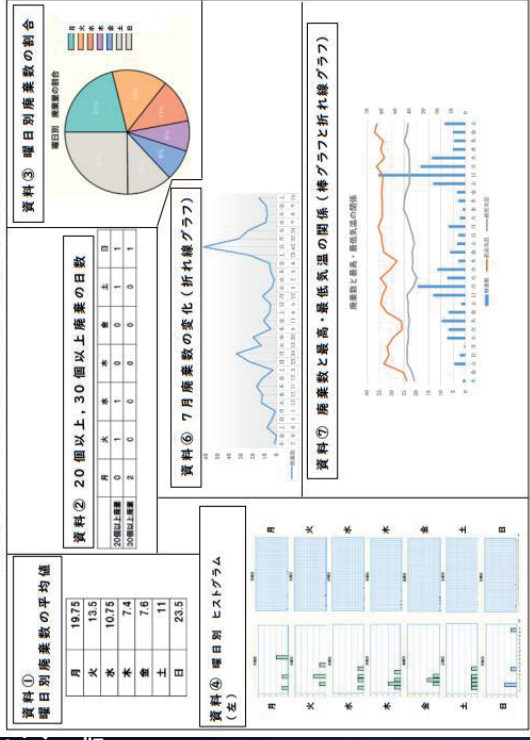
4 研究の実際と考察

(5) 自分たちや地域にできることを考える (手だて③④) 【総合：第8～17時】



地域のコンビニエンスストア7店舗、子どもたちの家庭が普段利用するスーパーマーケット5店舗から、7月の廃棄量のデータを集めた

資料 19 子どもたちが作成した資料



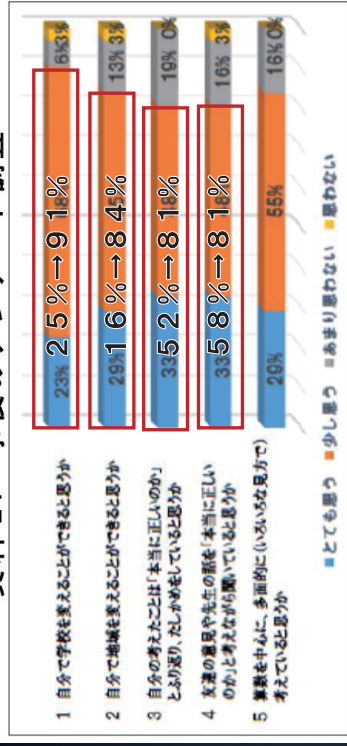
4 近隣店

16, 17時

5 研究の成果と今後の課題

(1) 研究の成果

資料 24 事後のアンケート調査



資料 25 A 児の振り返り (地域への提案会后)

とてもぎんちょうしたいけれど、お店の方や区長さんに提案、考えが伝わってよかった。
 区長さんやお店の方にほめてもらえて、とても自信がついたと思う。
 上り土成のためにがんばれたと思います。私たちの力で少しは土成を変えられたと思います

5 研究の成果と今後の課題

(1) 研究の成果

- ・「スキルBOX」を活用したことで、複数の観点で分析する姿が見られた。また、観点を組み合わせて、より妥当な判断をすることができた。
- ・統計的問題解決の過程や結論を検討するために、グループや全体での対話を行ったことで、結論を再検討したり分析方法を見直したりできた。
- ・SDGs や地域を関連させた単元構想により、日常生活の問題を自分事として捉え、現状を変えるために何かしたいという気持ちを強くした。学習を進めていく中で、実際に行動する姿も見られ、地域への提案会を通して「地域を変えられた」と実感することができた。
- ・VWを取り入れたことで、食品ロスに関する問題を身近な水に置き換えることができ、現状を変えたいという切実感が高まった。

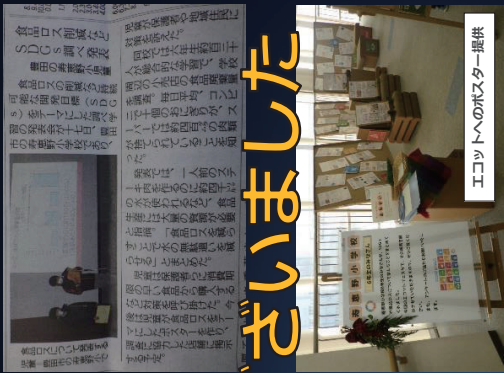
5 研究の成果と今後の課題

(2) 今後の課題

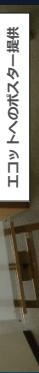
データを分析したことを話し合う際、理解が難しい子の姿も見られた



資料ごとに整理できるように、「**意図的指名**」**「板書の工夫**」などの**対話における教師の支援**が必要だった



ご静聴ありがとうございました



問題解決までの見通しをもち、
適切なグラフや分析結果を
問題解決に活用しようとする子の育成

— 5 年 割合のグラフ（あいさつNo.1プロジェクト）の実践を通して—

児童の実態②
課題解決に適した資料を選べない児童

・国語科「提案分を書こう」の学習において

動物の殺処分数が増えていることをグラフではなく表に表す児童

感染拡大を示すために棒グラフを使う児童

グラフからわかることまで分析できない児童

めざす子ども像②
問題解決のために、適切なグラフや
分析結果を活用しようとする子

児童の実態①
問題解決までの見通しが持てない児童

・「単位量当たりの大きさ」の単元において

答えの導き方を説明できない児童

式によって導かれた値が何を示しているかわからない児童

出てきた数を使ってどうするかわからない児童

めざす子ども像①
筋道立てて考え、問題解決までの見通しがもてる子

統計的問題解決の過程と児童の実態

児童の実態①
問題解決までの見通しが
持てない児童

児童の実態②
課題解決に適した資料を
選べない児童



めざす子ども像①
筋道立てて考え、問題解決
までの見通しがもてる子

めざす子ども像②
問題解決のために、適切なグラフや
分析結果を活用しようとする子

研究の仮説と手立て①

仮説①

目的に応じたデータを集める計画を立てる
場面において、ワークシートを工夫すれば、
筋道立てて考え、問題解決までの見通しがも
てる子が育つだろう。

手立て1

思考の流れを見える化
したワークシートの活用

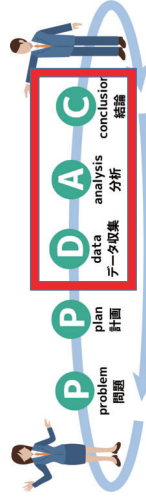
手立て2

話し合いの視点を明確化
したワークシートの活用

単元の構成と手立ての位置づけ①

小単元1

童子山小学校のケガの状況を調べよう



小単元2

あいさつNo.1プロジェクト

研究の仮説と手立て②

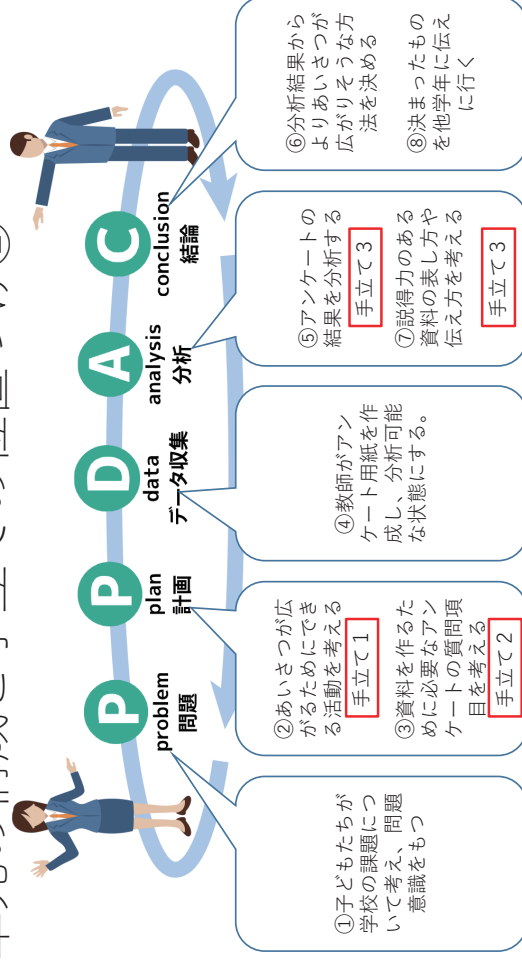
仮説②

集めたデータを使ってどのようなグラフを考え
える場面において、様々なグラフに表して考えら
れる環境があれば、問題解決のために、適切なら
グラフや分析結果を活用しようとする子が育つだろ
う。

手立て3

タブレットを用いて、多様なグラフに表して分
析できる環境を作る

単元の構成と手立ての位置づけ②



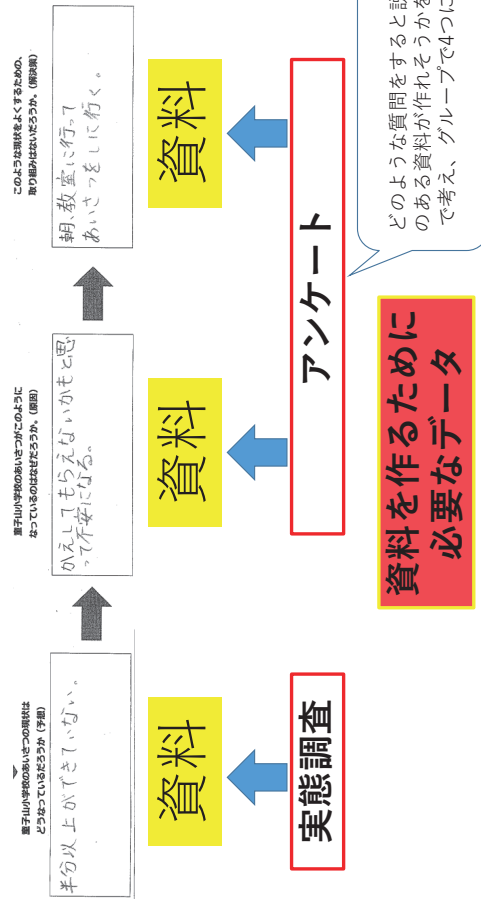
抽出見Aにかけた願い

- 答えが出た後に、自分の考えを説明することが苦手で、式から出てきた値が何を示すか見通しを持っていないことが多い。

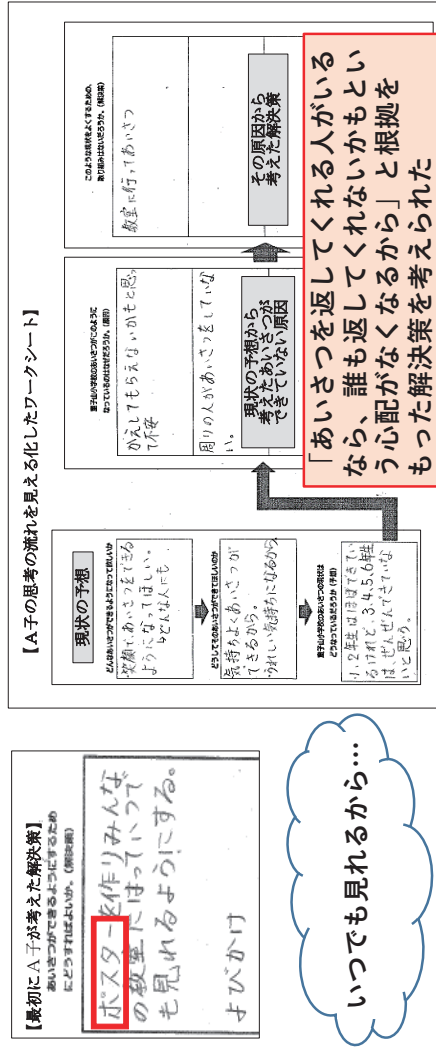
根拠のある筋道だった問題解決までの見通しをもって、何のために必要な質問かを考えたり、得られたデータが活用できそうかを考えたりできるようになってほしい。

- 動物の殺処分数を伝える際、グラフではなく表を使っていた。

問題解決のために、適切なグラフや分析結果を活用できるようにしてほしい。



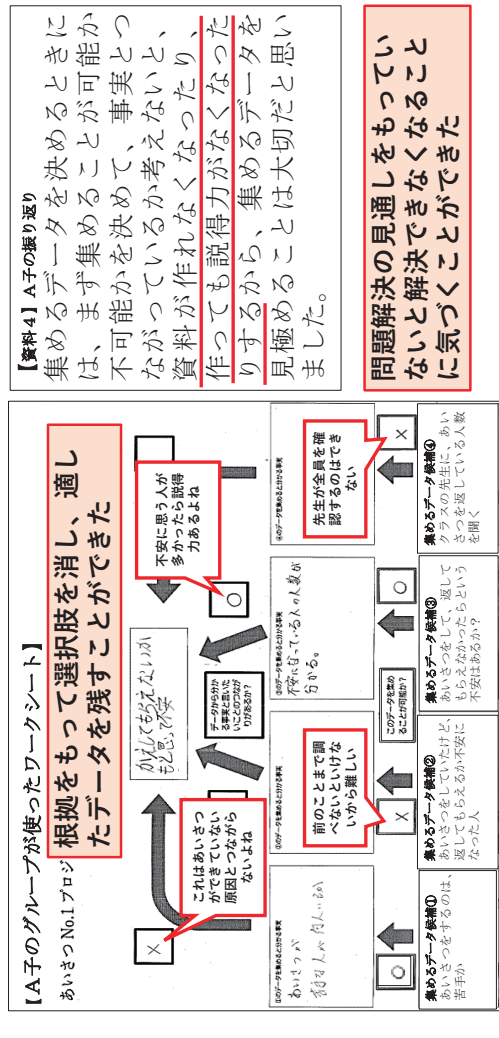
手立て1【思考の流れを見える化したワークシートの活用】について



いつでも見れるから...

「あいさつを返してくれないかものというなら、誰も返してくれないから」と根拠を
もった解決策を考えられた

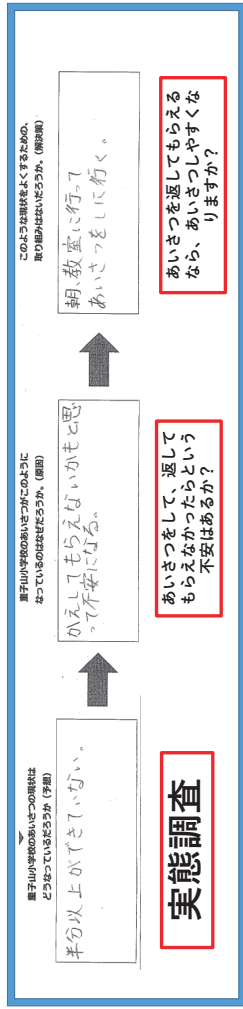
手立て2【話し合いの視点を明確化したワークシートの活用】について①



問題解決の見通しをもっていないと解決できなくなることができた

手だて3

【タブレットを用いて、多様なグラフに表して分析できる環境を作る】について①



9つ→2つの考
え方にクラスで
話し合って絞る

アンケートを教師
が作成し、回答し
てもらおう

教師が分析可能な
状態までExcelの
編集をする

手だて3

【タブレットを用いて、多様なグラフに表して分析できる環境を作る】について③

【タブレットを用いた分析後のA子の振り返り】

学習の振り返り

以外と、自分がイメージしてい
たのちが、たのでびっくり
しました。

例えば、あいさつができて
いない原因は私はけずが
いかなのかかなと思っ
た結果は、できて、みんなが
かたりしたのでおどろきま
した。

タブレットでデータを見るとき
書をよりにくさんか、人に
見れるいろいろなグラフで
表せるのでもいいと思いまし

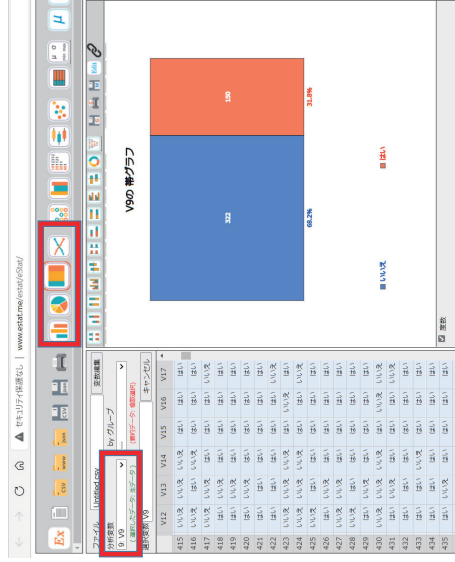
集めたデータの結果を予想しな
がら、確認という流れが1つの
授業内で何回もできた

たくさんデータをグラフの形
を変えながら見ることができた

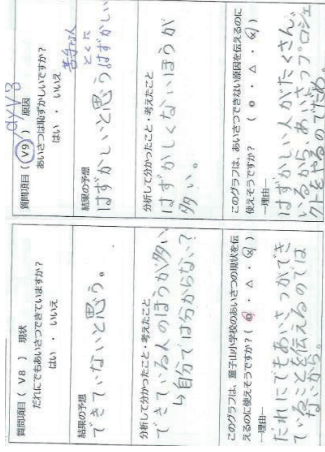
手だて3

【タブレットを用いて、多様なグラフに表して分析できる環境を作る】について②

【授業で使った多様なグラフに表せるサイト】



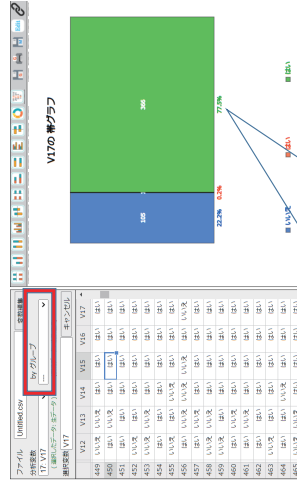
予想→分析
→使えるかどうかの検討



手だて3

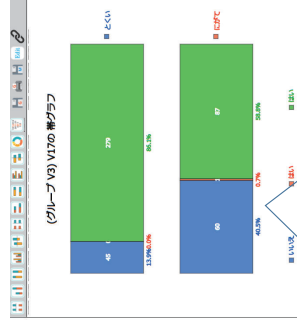
【タブレットを用いて、多様なグラフに表して分析できる環境を作る】について④

【「いいあいさつができたからカードがもらえるなら、
楽しくあいさつできそうですか」に対する回答】



「はい」の人が多いから
カードを渡すとあいさつが
できるようになるという主張に
説得力が付けられる

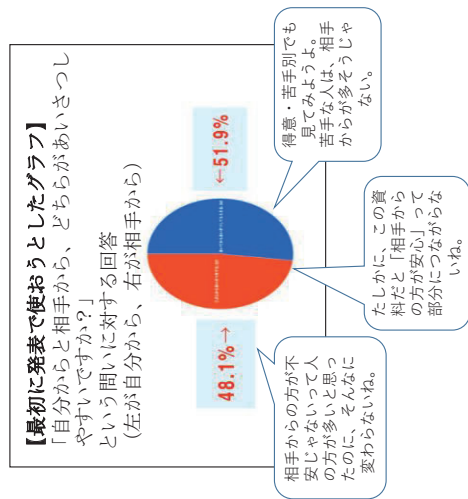
1つのデータを多面的
な視点から分析し、吉
効果的な解決策を選ば
うとする姿が見られた。



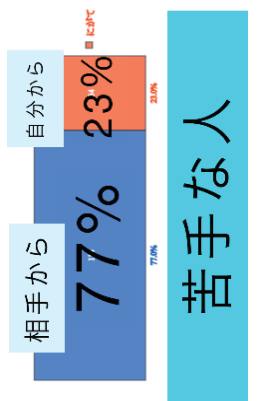
苦手な人が半分くらい
「いいえ」と答えているから
あいさつが広げられなさそう

手立て3

【タブレットを用いて、多様なグラフに表して分析できる環境を作る】について⑤



【実際に発表の時に使ったグラフ】



問題解決のために、より適したグラフを活用しようとして話し合い、説得力のある資料を提示することができた。

研究の成果〈仮説②〉～A子の変容から～

一当初の実態一

動物の殺処分数を伝える際、グラフではなく表を使っていた。

手立て3

タブレットを用いて、多様なグラフに表して分析できる環境を作る

タブレットを用いて、データを多面的に分析できる環境を用意したこと、より説得力のある資料を提示しようとして話し合ったり、問題解決により効果的な解決策を選ぶことができた。

研究の成果〈仮説①〉～A子の変容から～

一当初の実態一

答えが出た後に、自分の考えを説明することが苦手で、式から出てきた値が何を示すか見通しを持っていないことが多い。

手立て1

思考の流れを見える化したワークシートの活用

手立て2

話し合いの視点を明確化したワークシートの活用

あいさつが広がりそうな解決策を、予想したあいさつができていない原因から筋道立てて考え、説明することができた。集めたデータからどのようなことが分かるかを見通し、必要なデータとそうでないデータを見極めることができた。

研究の反省

23時間という長期間の単元になってしまった。

→教科書の配当時間は7時間

算数だけでなく、国語や総合と合わせて進めていく必要がある

子どもの考えた解決策が実行できない児童や、調べたかった質問事項がアンケートに乗らなかった児童が多かった。

→児童の意識・意欲が途切れてしまう恐れがある

児童がせっかく考えた方法を実践できる方法を探し必要がある。

PPDACサイクルの2周目に入ることができなかった。

→計画を練る段階で時間数を使いすぎてしまっていた。

計画を緻密に立てるのではなく、うまくいかなかったから次を考えるという流れなら2周目に自然に入ることができるのではないか。

数学的モデル化を遂行する力を評価するフレームについての研究 —中学校第1学年「データの活用」領域における授業実践—

福井大学教育学部附属義務教育学校 三 上 泰 生

福井大学教育学部 松 本 智 恵 子

hrk.3kaaa3@gmail.com

1. はじめに

現代の社会は VUCA (Volatility:変動性 Uncertainty:不確実性 Complexity:複雑性 Ambiguity:曖昧性) な時代の到来といわれており、社会が急速に変化し、予測することが困難になってきているといわれている。こうした社会が到来する中、学校教育においては、授業と評価をそれぞれ別々に改善させていくのではなく、両輪で行っていく必要性が明示されている(国立教育政策研究所教育課程研究センター(2020))。

現実世界の問題を、数学を用いて解決する力を育成する授業実践は多く発表されている(例えば、藤原(2020))。それらは数学的モデル化を遂行する力の育成を目的としており、数学的モデル化を遂行する力を授業者が評価する方法や、生徒自身が自己評価するための枠組みを生徒とともに創り上げることに注目がされてこなかった。こうしたことを踏まえ、本研究では、西村(2012)の「数学的モデル化を遂行する力」に着目し、「数学的モデル化を遂行する力」を評価するフレームを生徒とともに開発することを目指す。この目標に向け本稿では、西村(2012)の先行研究に基づいて、「数学的モデル化を遂行する力」を育成する授業の開発を行った。その上で、中学校第1学年の統計領域の授業の中で、生徒とともに評価フレームの開発を行った。

2. 数学的モデル化を遂行する力を育成する教材開発

2.1 「数学的モデル化過程」と「数学的モデル化を遂行する力」

数学的モデル化過程について、西村(2012, pp. 22-23)では次の方針のもとで、「数学的モデル化過程」を図1のように規定している。

方針1 Pollak(2003)の規定を基盤として、データの扱いを加味し、問題解決過程を示す。

方針2 学習指導や評価の場面での利点を考え、それぞれの思考の相に「問題の翻訳」「定式化」等の名前を付けるとともに、それらの思考においてめざすものとして「社会の問題」「数学の問題」「数学的結論」を明示する。

本研究では、図1の数学的モデル化過程の定義を援用する。

また、西村(2012, p. 85)では、Wildら(1999)の統計的探究過程で働く考え方や能力を拡張し、数学的モデル化を遂行する力を4つの次元で規定している。

次元1 数学的モデル化過程

次元2 数学的モデル化能力

次元3 批判的サイクル

次元4 数学的モデル化に関する意識・態度

本研究では、数学的モデル化を遂行する力を評価するフレームを、生徒が自分の活動を振り返って行う自己評価や他者の発表を聞いて評価するペア評価後に生徒と創り上げていくことを目的としている。そのため、成果物(レポート)や発表の内容から評価しやすいと考えられる次元1と次元2を、数学的モデル化を遂行する力の対象とした。

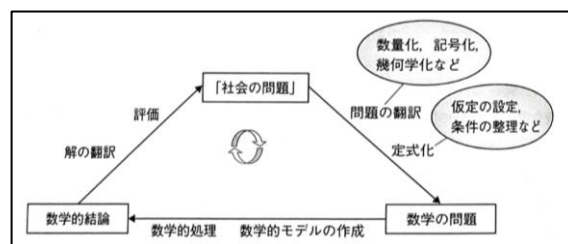


図1: 西村(2012)における数学的モデル化過程

2.2 数学的モデル化を遂行する力を育成する教材の開発

数学的モデル化を遂行する力を育成する教材として、本研究では「福井の冬は昔と比べて暖かくなっているのか」という教材を開発した。この教材は、西村(2012, pp. 254-287)にある「東京の夏は昔と比べて暑くなっているのか」という統計領域における実践を参考に作成している。また、西村(2012, pp. 73-84)も参考にしていたWildら(1999)が規定した統計的PPDACサイクルや、授業実践を行った福井大学教育学部附属義務教育学校の授業づくりに使われている「発意—構想—構築—遂行・表現—省察」の協働探究サイクル(図2)も考慮に入れている。(福井大学教育地域科学部附属中学校研究会, 2010)

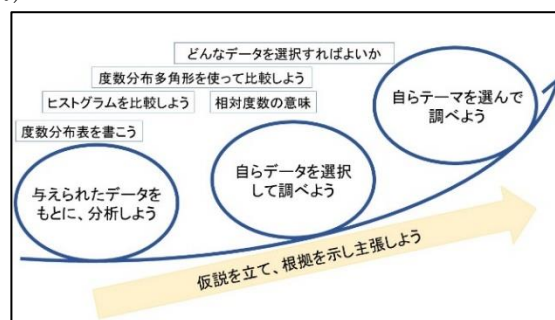


図2: 福大附属の協働探究サイクル

まず、第1次で、教師から与えられたデータをもとに、数学的モデル化過程を通して、結論を導くという授業をデザインした。数学的モデル化過程1周目において、過去の学びを想起させ、小学校算数科で学習した知識・技能を活用し、課題に対する結論を導かせる。次に、数学的モデル化過程2周目において、代表値を数学的モデルとし、それらの比較から課題に対する結論を導かせる。さらに、数学的モデル化過程

3 周目において、ヒストグラムを数学的モデルとし、ヒストグラムの形と代表値を絡めながら課題に対する結論を導かせるという 3 週の数学的モデル化過程を通すようデザインした。第 2 次には、第 1 次と同じ課題であるが、調べるデータは生徒が選択することとした。第 3 次では、自分たちで設定した社会の問題(課題)について、数学的モデル化過程を通して、結論を導くデザインをした。

3. 数学的モデル化を遂行する力を評価するフレームの検討

西村(2012, pp. 254-287)は、「東京の夏は昔と比べて暑くな

っているのか」という中学校第 1 学年の教材を開発し、数学的モデル化を遂行する力を評価するフレームとしてルーブリックを示している。その評価フレームの課題として、西村(2012, p. 201)では、生徒が使用する際、ルーブリックの表現を生徒にわかりやすく改めると留意されているが、教師が表現を改めたとしても生徒が理解しやすいものになるとは限らない。そのため、本研究では、新しい統計領域における数学的モデル化を遂行する力を評価するフレームとして、まず、生徒と評価フレームについて対話する中で改善していくことを前提とした試案を教師側が作成した(表 1)。

表 1:教師が作成した評価フレーム(試案)

	非常によい(S)	よい(A)	普通(B)	もう少し(C)
問題の翻訳		「暖かくなっている」ことに関する定量的な基準を設けて仮説を立てる。	「暖かくなっている」ことに関する定量的な基準を設けずに仮説を立てる。	B に至らない。(「暖かくなっている」ことに関する仮説を立てていない。)
定式化	仮説に基づいて、複数年のデータを複数の観点で収集しようとする。	仮説に基づいて、最低気温について複数年のデータを収集しようとする。	仮説に基づいて、最低気温について 2 か年のデータを収集しようとする。	B に至らない。(仮説に基づいていない、データを集めようとししない)
作成・数学的モデルの処理	代表値、ヒストグラムの形状(広がり、山の形)を総合的に分析して、結論を出そうとする。	代表値、ヒストグラムの形状を分析し、結論を出そうとするが、つながりは見いだせない。	代表値、ヒストグラムなどを 1 つだけ使って結論を出そうとする。	B に至らない(何も使おうとししない)
・解の翻訳 ・評価	仮説に基づいて、適切な解釈及び判断をし、その根拠を適切に示し、他者に伝えるように説明できている。	仮説に基づいて、飛躍はあるものの解釈及び判断をし、他者に伝えるように説明できている。	仮説に基づいて、適切な解釈及び判断をしている。	B に至らない。(解釈及び判断ができていない。)

表 1 は、「福井の冬は昔と比べて暖かくなっているのか」という今回の実践課題における数学的モデル化を遂行する力を評価するフレームの試案である。ただし、課題として挙げた「生徒が理解できる評価フレームになっているか」については生徒自身が判断するものである。そのため、授業実践の中で、表 1 の評価フレームを使用した自己評価とペア評価を生徒たちに実践させ、その後評価フレームの検討を行い、生徒たちの意見をもとに表 1 を改善することにした。

4. 授業実践

4.1 授業実践の目的とその評価方法

本授業実践の目的は 2 点ある。まず、第 2 節で提案した教材で「統計領域における数学的モデル化を遂行する力を育成する」ことができるかどうかを確かめることである。次に、第 3 節で提案した評価フレームの試案(表 1)を改善し、数学的モデル化を遂行する力を評価するフレームを生徒たちとともに開発することである。

以上の目的が達成されているかについては、統計領域にお

ける数学的モデル化を遂行する力を授業実践内で生徒が作成したレポート・ワークシートで評価する。

4.2 実践方法

対象：福井大学教育学部附属義務教育学校後期課程 7 年生 1 クラス(男女 36 人)

時期：令和 3 年 1 月 8 日(金)～令和 3 年 1 月 28 日(木)
(第 1 次・第 2 次) 計 9 校時
令和 3 年 3 月 18 日(木)～令和 3 年 3 月 22 日(月)
(第 3 次) 計 2 校時

※第 2 次と第 3 次の間に課題テーマを考える時間を確保するために、別の単元を行っている。

4.3 授業実践の様相

4.3.1 第 1 次(第 1 時～第 4 時)

第 1 次の導入では、授業前に行った認識調査で判明した、生徒が持っている地球温暖化に対するイメージ例として「地球全体の平均気温が上がること」「二酸化炭素が原因であるこ

と」を取り上げた。しかし、この授業を行った日が大雪であったことから、生徒は地球温暖化のイメージとのずれを感じている様子だった。ここで「福井の冬は昔と比べて暖かくなっているのか」を問うと、大半の生徒は「暖かくなっている」に手を挙げた。少数派の「暖かくなっていない」に手を挙げた生徒は「今の外見たらなっていない！」と意見を述べた。この時の意見は経験に基づいたものであり、データによる根拠はない。そのため、数学的モデル化過程の1周目として、小学校算数科で行ってきた過去の経験・体験を想起させ、データを用いて課題に対する結論を導かせた。

子どもたちに課題を解決するためにどのようなデータが欲しいかを問うと、「過去の気温」「今の気温」「積雪量」が欲しいといった。データとして、教師が用意した「2007年（子どもたちが生まれた年）と2019年（子どもたちが暖冬といった年）の12月の最低気温」を表にまとめたものを配布した。4人班で課題に対する結論をホワイトボードに書くよう指示し、班活動に入った。

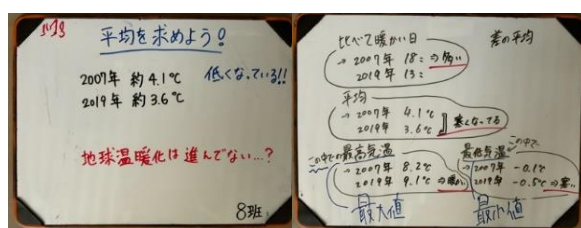


図 3: 班で出したテーマに対する結論

課題に対する結論を導くために2007年と2019年の12月の日別最低気温を比べたとき、すべての班が平均値を用いていた。これに対し、1つの班だけが、平均値のみではなく、最大値、最小値を比べている。この班は『平均値はデータを比べる際の1つの指標でしかない』と判断したようである。実際、平均値だけを比べてみると、2007年のほうが2019年よりも高い値となっているため、地球温暖化は進んでいないという結論に至る。しかし、最大値で比べると2019年の方が高いことから、暖かくなっているという結論も導ける。この班活動によって、平均値、最大値、最小値で比較するという考えが共有され、代表値の学習を行った。代表値は小学校第6学年で学習するが、本時では出てこなかった中央値と最頻値の意味も復習してまとめた。その後、実際に2つのデータから代表値を求めさせ、生徒は代表値という数学的モデルを使って結論を導くことを体験した。このようにどの値を根拠に結論を導くかで結論が真逆になることを実感した生徒たちは『2つのデータだけでよいのか?』ということに着目しはじめた。ここで、もう一度提示した課題に戻り、「福井の冬は昔と比べて暖かくなっているのか」の定性的な言葉に着目する機会を設けた。ある生徒が「冬は12月だけではないと思う。」と発言し、続いて「自分が生まれた年より前のデータと比べてみたい。」と発言した。教師は、「では、皆さんにとって冬とは何月のことですか?」と問うと、大半の生徒が、12月・1月・2月が冬であると答えた。しかし、11月も冬に入ろうという意見から、生徒全員を巻き込んだ議論になった。定

性的な「冬」という言葉の定量的な基準に11月を入れるかどうかは、全体では統一せず、結論を導く際どのようなデータを根拠にしているかを明示することを指示した。これによって、班活動でもう一度11月を冬とするかを議論し、合意形成をはかりながらデータを根拠に探究していくこととした。また、「冬」と同様に、「昔」とは何年前のことなのか、「暖かくなっている」とはどのような状態で、どのように定量化するのかを考えさせる機会を設けた。

代表値を用いて結論を導く活動を行った生徒たちは、最初の班活動の際に書いたホワイトボードにあった度数分布表という用語に着目した。生徒に度数分布表とヒストグラムを作成する経験をさせるために、このタイミングでワークシートを配布した。度数分布表を作成しつつ、階級や階級値、度数などの用語を確認した。その後、2つのデータの度数分布表からヒストグラムを作成し、結論を導く活動へ入った。

生徒たちはヒストグラムによって視覚化されたデータの形に着目する。「山の形」「階段状」などの言葉を用いて形の違いから結論を導こうとしていた。ここである生徒が「ヒストグラムを重ねたい」という発言をした。これは前単元の関数の学習で比例と一次関数のグラフを比べる際、同じ座標軸にグラフをかくことで、2つが平行なグラフであることに気づいた体験から出てきた発言と考えられる。関数の学びを思い出させつつ、ヒストグラムを重ねるためにOHP透明シート(以下、透明シートとする)を配布した。透明シートに2007年のヒストグラムを写し取り、それを2019年のヒストグラムに重ねてみると、ヒストグラムは幅を持った柱状のグラフであるため違いが見えづらく、比べることが困難であった。この困難を解消するために、生徒は柱の真ん中に点を打ち、線分で結び始めた。これは度数分布多角形である。生徒はヒストグラムより度数分布多角形を重ねて比べたほうが分かりやすいと判断した。これは課題に対する結論を導くためにデータを比べる際、判断しやすい数学的モデルを生徒たち自身が作り上げた瞬間である。

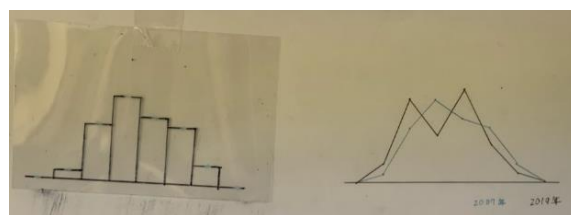


図 4: 透明シートに度数分布多角形を描く

ヒストグラムや度数分布多角形を比べ、課題に対する結論を導いたのだが、その大半は「どちらともいえない」であった。度数分布多角形の形状は明らかに異なっているが、生徒はその違いを「暖かくなっている」根拠になり得ないと判断したと考えられる。それは生徒の「暖かくなっている」の基準は、気温が高くなっていることであったためだと考えられる。実際は、2019年のヒストグラムの形がM字形で、12月の日別最低気温の寒暖差が激しくなっている。そのことと地球温暖化という異常気象とをつなげることができなかったの

あろうと予想される。

4.3.2 第2次(第5時～第9時)

第2次では、まず、2007年と2019年の12月の日別最低気温を比較した学びのサイクルを振り返った。

第1次では、生徒は2つのデータを比較し、課題に対する結論を何度も導いた。その際、根拠とする数学的モデルが代表値やヒストグラムなど多様であったため、同じ課題に対しても、意欲を持って取り組むことができていた。

次に自分たちで必要なデータを集め、数学的モデルを作成し、比較することで課題に対する結論を導く段階に入った。主には、第1次でも検討する場面があった「冬」、「昔」、「暖かくなっている」についての基準を班で合意形成しながら決定していく活動となった。

次に、定式化の相へと移った。問題の翻訳の相で合意形成した基準に合わせて、結論を導くための根拠となるデータを検討した。第1次の活動の際、2007年と2019年における12月の日別最低気温の出典が、気象庁のホームページであるとワークシートに記載していたため、生徒は気象庁のホームページから自分たちに必要な年のデータを収集していた。その後、数学的モデルの作成・数学的処理の相に入った。生徒は第1次で学習した代表値やヒストグラムを作成し、その比較から結論を導こうとしていた。また、月別降雪量のデータを根拠に結論を導こうとしていた班もあった。その班は、小学校算数科で学習した折れ線グラフを数学的モデルとして作成し、月別降雪量のグラフを重ねて比較するなど、データに合わせた数学的モデルを作成しようとする態度がみられた。

解の翻訳・評価の相において、生徒は数学的モデルを根拠とした結論が書かれたレポートを作成し、別の班の生徒とペアを組み、発表会を行った。

4.3.3 第3次(第10時・第11時)

第3次では、テーマ自体を生徒が考え、似たテーマの生徒とグループを組み、グループごとに数学的モデル化過程を通して結論を導く活動を行った。生徒は、代表値やヒストグラムなどの数学的モデルを作成し、結論を導くという数学的モデル化過程を進めていった。解の翻訳・評価の相において、生徒は数学的モデルを根拠とした結論をレポートにまとめ、別のグループの生徒とペアを組み、1対1で発表会を行った。

5. 子どもたちとともに考える評価フレームの提案

5.1 評価フレームを修正する

第3節で「福井の冬は昔と比べて暖かくなっているのか」という課題における数学的モデルを遂行する力を評価する試案(表1)を提案した。表1の評価フレームはあくまで教師が作成したものであり、生徒が理解できる記述となっているのかどうかを生徒自身が判断し、改善していく必要がある。その判断を行うために、実際に生徒がこの評価フレームによる自己評価やペア評価を行い、評価フレームについて検討する

時間を設けることが必要である。このような活動によって、表1を統計領域における幅広いテーマに順応する評価フレームとしていく。そのために、第2次、第3次の発表会后に評価フレームについて生徒の意見を聞きながら、改善していくこととした。

5.2 評価フレームの検討

第2次の終わりに、生徒たちは「福井の冬は昔と比べて暖かくなっているのか」に対する結論を導く根拠となる数学的モデルをレポートにまとめた。その上で、生徒たちは、異なる班員同士のペアで、レポートに書かれた数学的モデルやその処理を根拠として、課題に対する結論を相手に伝える活動を行った。その後、生徒たちに、生徒自身の発表とペアの発表について、教師が作成した評価フレーム(表1)を使用して、自己評価とペア評価をさせた。その上で、評価フレームについて生徒たちが感じたことを聞き出し、教師とやり取りをする場面を設けた。

定性的		
7C 数学 「福井の冬は昔と比べて暖かくなっているのか」		
	非常によい(S)	よい(A)
問題の翻訳	「冬、昔」にも基準 1～2 50年前 12～2 20年前	「暖かくなっている」 ことに関する定量的な 基準を設けて仮説を立 てる。具体的な 数値が基準をつくる

図 5: 問題の翻訳に関する生徒のワークシート

まず、生徒たちは、理解しづらい言葉として「定量的」という言葉を挙げた。確かに「定量的」という言葉は生徒にとって馴染みがなく、イメージが浮かばないと考える。しかし、生徒は第1次で「冬」に11月を入れるかを議論する活動などから、ある程度まではその意味を推察できていると考えられる。そのため、教師は生徒たちに「では、どのような言葉を用いればよいか」と返した。すると、生徒は「具体的というか…」や「数値を使って基準を設ける」などの的確なつぶやきがあがってきた。そのつぶやきが全体に共有されると、「定量的」を理解していなかった生徒からは感嘆の声があがった。教師は復習を兼ねて、「福井の冬は昔と比べて暖かくなっているのか」の課題を「定量的」に翻訳することはどういうことかを問うと、「何度暖かくなっているのか」「何年前なのか」という反応があった。また、問題の翻訳の相は、教師が作成した評価フレームではS評価の部分が空欄となっていた。すると、生徒たちは「冬、昔、暖かいに対して基準を設けたらS評価にしてもよい」、「暖かいに基準を設けないと、結論を導くことができないから、暖かいに基準を設けることをAとする」など、3つの定性的な言葉にそれぞれ重みづけをする発言も出てきた。問題の翻訳の相の対話のように、各相において生徒たちが理解しづらい言葉を教師が補足したり、改善

したりする活動を行った。

第 5.2 節の評価フレームに関する生徒の意見を反映して、修正・整理した評価フレームを表 2 に示す。この表 2 を、新たな評価フレーム（本案）とする。

5.3 子どもたちと創り上げた評価フレーム

表 2:生徒たちと創り上げた評価フレーム(本案)

	非常によい(S)	よい(A)	普通(B)	もう少し(C)
問題の翻訳	課題の中にある複数の定性的な言葉に対して、具体的な数値による基準を設けて仮説を立てる。	課題の解決に重要な定性的な言葉に対して、具体的な数値による基準を設けて仮説を立てる。	課題の解決に重要な定性的な言葉に対して、具体的な数値による基準を設けずに仮説を立てる。	B に至らない。(定性的な言葉に対して、仮説を立てていない。)
定式化	問題の翻訳の相で設定した仮説に基づいて、複数年のデータを複数の観点で収集しようとする。	問題の翻訳の相で設定した仮説に基づいて、複数年のデータを収集しようとする。	問題の翻訳の相で設定した仮説に基づいて、特定のデータを収集しようとする。	B に至らない。(仮説に基づいていない、データを集めようとしない)
数学的モデルの作成・数学的处理	代表値（平均値、中央値、最頻値、最大値、最小値）、相対度数、ヒストグラムの形状（広がり、山の形）など複数の数学的モデルを総合的に分析して、結論を出そうとする。	代表値（平均値、中央値、最頻値、最大値、最小値）、相対度数、ヒストグラムの形状など複数の数学的モデルをそれぞれ分析し、結論を出そうとする。(分析結果のつながりを見出すまでには至らない)	1 つの数学的モデルを使って結論を出そうとする。	B に至らない(何も使おうとしない)
解の翻訳・評価	問題の翻訳の相で設定した仮説に基づいて、データに対して適切な解釈及び判断をし、他者に伝わるように説明できている。	問題の翻訳の相で設定した仮説に基づいて、段階を踏んでいない部分はあるもののデータに対して適切な解釈及び判断をし、他者に伝わるように説明できている。	問題の翻訳の相で設定した仮説に基づいて、適切な解釈及び判断をしているが、他者に伝わるように説明できない。	B に至らない。(解釈及び判断ができていない。)

5.4 評価フレームの本案を用いて評価する

第 3 次では、生徒が考えた課題テーマの発表会を行った。その自己評価とペア評価において表 2 の評価フレームを使用した。その上で、評価フレームについて生徒たちが感じたことを聞き出し、教師とやり取りをする場面を設けた。

生徒からは 2 つの指摘がなされた。1 つは定式化の相において、複数年（3 か年以上）のデータを集めることは当たり前ではないかということである。この指摘から、生徒は複数回の数学的モデル化過程を通す中で、現実の世界における問題は特定の 2 か年について数学的モデルを使って比較し、結論を導いたとしても、その結論には妥当性があまりないということが認識されてきていることが分かった。その認識から、特定のデータを収集することが普通(B)という評価であることを改善してほしいと要望された。もう 1 つは解の翻訳・評価の相において、他者に伝わるように説明できているかは自分では判断できないということである。この指摘はもっともである。表 2 の評価フレームに、数学的モデル化を遂行する力を自己評価する機能を持たせるためには、解の翻訳・評価

の相に関する再定義の必要性が示唆された。

6. 考察

6.1 生徒たちの「省察シート」から考察する

対象のクラスでは、単元が終わるごとに、生徒が学びを俯瞰しながら探究したことや仲間とともに学んだことのよさ、学んだことが数学の他の単元や他教科でどう生かされると思うかについて書き記す「省察シート」を作成している。この省察シートを利用して、生徒の記述を見取ることで、授業実践の評価を行った。

この省察シートをみると、ある生徒は複数の数学的モデルを作成することはできていたが、課題を解決する際どの数学的モデルが適切かという視点が欠けていたことを反省している。その反省を踏まえ、次回は数学科以外の教科での探究で生かしていこうとする意欲もみられた。また、評価フレームを生徒同士で使うことに対し、すぐにフィードバックが得られることや評価する側に立って評価フレームを見ることで、その記述を理解していないと、適切な評価ができないことに

多くの生徒は気付いていた。これより、数学的モデル化を遂行する力は生徒間で差があるものの、この授業実践を通じて着実に身につけていると判断した。

6.2 生徒の自己評価から考察する

生徒たちは第2次と第3次において、評価フレームを使った自己評価を行った。評価フレームを使ったS、A、B、C評価の人数の推移から、本実践が数学的モデル化を遂行する力を育成する授業実践であったかを判断することにした。生徒36人の第2次と第3次における各相のS、A、B、C評価の人数を、それぞれ表3、表4で示す。(便宜上、①問題の翻訳、②定式化、③数学的モデルの作成・数学的处理、④解の翻訳・評価とする。)

表3: 第2次における自己評価

	S	A	B	C
①	0人	24人	11人	1人
②	9人	11人	16人	0人
③	9人	16人	11人	0人
④	6人	14人	16人	0人

表4: 第3次における自己評価

	S	A	B	C
①	2人	24人	10人	0人
②	8人	18人	10人	0人
③	1人	26人	5人	4人
④	12人	15人	9人	0人

表3、表4の評価フレームでは、A評価以上は活動がよくできていたと判断される。A評価以上の人数の推移をみると、すべての相で人数が多くなっていることが分かる。これより、数学的モデル化を遂行する力を育成する授業であったと判断することができる。

6.3 成果と課題

本授業実践は、中学校第1学年の統計領域における課題に対し、数学的モデル化過程を通して結論を導く活動を行った。こちらから課題を与えたのは、「福井の冬は昔と比べて暖かくなっているのか」のみで、数学的モデル化過程を進めていく中で、疑問が次々とわき授業が進んでいったため、第2節で作成した授業教材は、授業者として実践しやすいものとなっていた。また、生徒から疑問が次々とわく状態が、多様な数学的モデルを意欲的に作成し、結論を導いている状態だと判断できる。さらに、第6.1節、第6.2節の考察からも、本実践は数学的モデル化を遂行する力を育成する授業であったと判断する。

評価フレームづくりの留意点として、この実践で作成した評価フレーム(表2)は、授業実践を行った生徒たちとともに創り上げたものである。目の前の生徒が変われば、着目する言葉も異なってくると予想される。そのため、この評価フレ

ームを基本的な評価フレームとし、学校現場の中学校第1学年の統計領域における数学的モデル化を遂行する力の評価フレームとしてアレンジし続けていくとよいだろう。

また、評価フレームの改善を生徒とともにに行った際、改善点をワークシートに書かせるだけでなく、自己評価やペア評価を行った直後に教師と生徒が評価フレームについて対話する時間を設けることで、第5.2節のように素直な意見が引き出せた。ただし、教師は、生徒が「わからない」といった部分のすべてを改善しようとするのではなく、生徒の思いを受け入れつつも、教師の意図を説明し、生徒とともに評価フレームを創り上げることが重要である。

本実践において2つの課題があった。1つは、評価フレームにある問題の翻訳の相の「具体的な基準」が客観性を持ち、妥当性のある数値である必要性を生徒に認識させることができなかったことである。数値によらない具体的な基準を設定したとしても、発表を聞く側やレポートを読む側が納得できないものである可能性がある。そのため、評価フレームを使う際は「妥当性のある具体的な基準」を検討しなければならない。もう1つは、生徒が評価フレームにある記述を読み替えられるよう促すことができなかったことである。第5.2節にある定式化の相において、「複数の観点」について具体的な例示をしてほしいという指摘があったが、これは第3次の評価フレームを使う際にも指摘があった。しかし、課題テーマによって用いるデータは異なってくるため、評価フレームに具体例を記述することは難しい。そのため、生徒に課題テーマに沿った観点が複数浮かび、それが課題テーマに適切かどうかを判断させる機会を設ける必要がある。

今後、2つの課題の克服と、多くの生徒を対象に授業実践を行うことで、評価フレームの精度を高めていきたい。

参考・引用文献

- 国立教育政策研究所教育課程研究センター(2020),「指導と評価の一体化」のための学習評価に関する参考資料【中学校 数学】, 東洋館出版社.
- 西村圭一(2012), 数学的モデル化を遂行する力を育成する教材開発とその実践に関する研究, 東洋館出版社.
- 福井大学教育地域科学部附属中学校研究会(2010), 学びを拓く〈探究するコミュニティ〉第1巻, エクシート.
- 藤原大樹(2020), モデリングと統計的探究のサイクルにもとづく数学科授業研究の諸相, 第8回春期研究大会論文集, pp. 179-186.
- Pollak, H. O. (2003), A History of the Teaching of Modelling, A History of School Mathematics Volumes, NCTM, pp. 647-671.
- Wild, C. J. & Pfannkuch, M. (1999), Statistical thinking in empirical enquiry, International Statistical Review, 67(3), pp. 223-265.

課題研究で使えるデータサイエンスの授業実践

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を目指し、「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(以下 AW)』と『データサイエンス (以下 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 (以下 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) と Analysis (分析) 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

DS 前期では、次年度より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) で講義と演習形式の両方で実践した。データの記述と可視化、統計的推定と検定等の内容を取扱った。

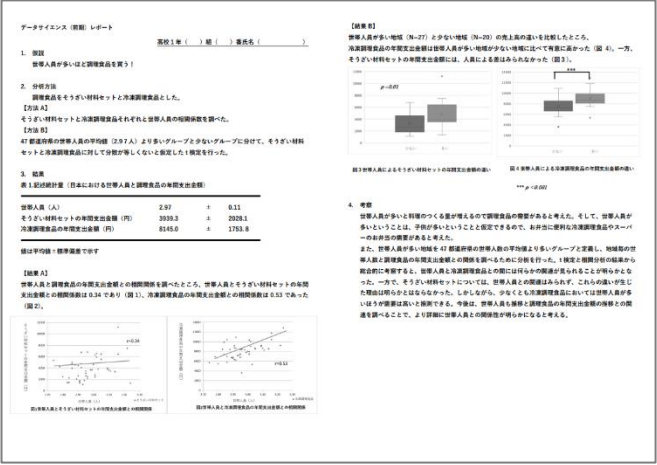
表2. DS 前期計画

回	内 容
1	基礎 データの種類・構造
2	基礎 統計図表・代表値
3	基礎 散布度
4	基礎 変数間の関係
5	演習 エクセルの基本操作
6	演習 度数集計
7	演習 クロス集計
8	演習 グラフ・ヒストグラム
9	基礎 統計的推測・有意性
10	理論 基礎 回帰分析
11	基礎 統計的検定
12	演習 相関、Excel 検定分析ツール
13	まとめ 演習 分析・レポート作成 1 回目
14	実践 演習 分析・レポート作成 2 回目
15	演習 レポート作成時の注意・まとめ

前期のまとめとして、教育用標準データセット SSDSE「都道府県庁所在市別、家計消費データ」から、生徒各自でデータの特徴を読み取りテーマ (仮説) を決め、統計分析と考察を行うレポート課題を設定した。生徒達は学んだ統計手法 (相関分析や t 検定) を用いて分析を行った。また、レポート内に記載

する図表、結果や考察の記述方法についての講義を行った。

図 1. 生徒のレポート例



4. 評価

レポート課題の評価にはルーブリックを用いた（図 2）。ルーブリックは育成目標である 4 つの力の評価に対応する形で作成した。特にレポート課題においては、分析作業を行う上での Excel の操作を含めた知識・技能がレポートの完成度に大きく関わってくるが、このような作業スキルが評価に直接的に影響しないよう調整を行った。分析する際のデータ整理や前処理段階でのアイデア、仮説や考察を設定する際の着眼点などが優れている場合に、評価として汲み取ることが可能となるようルーブリックの約 60%が「気づき」、「思考」を拾い上げる項目となっている。

5. 成果と課題

レポート課題について、各クラス 3 時間の作業時間と、授業後の自由参加の作業時間を 2 時間とした。この時間内でレポート完成に至った生徒は全体の 79%であった。また、自宅に PC 環境のない生徒のレポート完成度は PC 環境のある生徒に比べ 7%低かった。

さらにアンケートでは、「この授業によって事象を分析するための基本的なスキルが向上した」と 93%の生徒が回答し、「この授業の必要性を感じた」と 93%の生徒が回答した。一方で、「言葉の意味を理解するのが難しかった」など、統計の専門用語や独特な考え方に抵抗感を示す生徒がいた。統計理論の单元では、数学 I との接続を意識した為、分析に用いる式の構造を取り扱った。このことが生徒の苦手意識を発現させ

図 2. 課題の評価に使用したルーブリック

目標	つけたい力	評価項目
仮説を設定して研究計画を作成し、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 扱うデータに対し適切な分析方法を導き出す力	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 扱うデータの特性を理解し、適切な分析方法を導くことができる
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 他者へ論理的に説明する力	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 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
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 課題の本質を理解する力	他者の研究を理解し論理的な助言をすることができる 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
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	新たな価値を見出す力	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

評価 A：大変満足 B：概ね満足 C：要努力

た可能性が考えられる。また、「Excel などの使い方がすごく難しかった」と答えた生徒が一定数存在した。授業外の自習時間では、操作に慣れる時間としては十分でなかった可能性がある。

今後の課題として、取り扱う用語や複雑な式、理論的内容の取捨選択、説明方法の再検討が必要である。また、PC の操作が慣れない生徒や環境が無い生徒については授業外でのフォローの時間確保、貸出用 PC の用意等が必要である。

高校生のためのデータ駆動型の授業デザインと実践

ー 地域の課題から気づき生まれるデータサイエンス教育 ー

松宮敬広¹ 和泉志津恵² 畑山満則³

滋賀県立虎姫高等学校¹ 滋賀大学データサイエンス学部² 京都大学防災研究所³

¹ 〒529-0112 滋賀県長浜市宮部町 2410 番地, Email: matsumiya-takahiro-4192@pref-shiga.ed.jp

1. はじめに

1-1 背景

平成29・30年改訂学習指導要領（新学習指導要領）では、「総合的な学習の時間」が「総合的な探究の時間」と改められ、教科「理数」が新設されるなど、探究活動の拡充が強く意図されている。一方で、先進的に探究活動の取組を進めてきた全国のスーパーサイエンスハイスクール（以下SSH）指定校からは、課題研究における時間不足や教員への負担過多が、共通の課題として報告されている（国立研究開発法人科学技術振興機構，2019）。

また、新学習指導要領では、統計教育の充実が謳われている。とくに統計的な知識を現実の問題に活用することが重視されており、例えば数学Ⅰの内容には、「目的に応じて複数の種類のデータを収集し、適切な統計量やグラフ、手法などを選択して分析を行い、データの傾向を把握して事象の特徴を表現すること。」などの記述が見られる（文部科学省，2018）。

本稿では、新学習指導要領のねらいに即しながら、より小さな負担で実施できる事例として、データ駆動型の探究授業のデザインを提案し、実践結果を報告する。なお、本稿の内容は、第16回統計教育の方法論ワークショップにおいて報告した授業デザインを改善し、アンケート調査の結果分析を加えたものである。

1-2 研究の実施校

本研究は、滋賀県立虎姫高等学校（以下虎姫高校）の第2学年生徒を対象としたデータサイエンス教育の授業デザインと指導方法の開発を目的とする。虎姫高校は、平成24年度（2012年度）に文部科学省からSSHの研究指定を受け、学校設定科目「究理Ⅱ」を開設し、実施してきた。「究理Ⅱ」は、1年間かけて研究活動を行う、いわゆる課題研究の科目である。SSH指定1期目（2012年度～2016年度）には、理系の希望者を対象に選択科目として実施してきたが、2期目（2017年度～2021年度）からは「究理Ⅱ」の対象を理系生徒全

員へと広げ、データ駆動型の探究活動を導入した授業を実施している。そして、2018年度から「究理Ⅱ」をサイエンスコースとデータサイエンスコース（以下Dコース）の2つに分けた。また、2期目の2017年度に、滋賀県高等学校教育研究会数学会会長 三上保彦氏が校長として赴任した。

滋賀大学は、平成29年（2017年）4月に日本初のデータサイエンス学部を開設した。それに1年先立ち設置されたデータサイエンス教育研究センターでは、データサイエンス教育の教材や教授法の開発、そしてデータサイエンスに関する価値創造プロジェクト研究などの活動を推進してきた。さらに、滋賀大学は、平成28年（2016年）12月に、文部科学省の数理及びデータサイエンス教育の強化に関する懇談会において、東京大学などを含め、データサイエンス教育の拠点校に選定された（数理及びデータサイエンス教育の強化に関する懇談会，2016）。我が国におけるデータサイエンス教育研究を先導する拠点校として、協力校とも密接な連携を図り、数理・データサイエンス教育の全国の大学への普及・展開の加速化に向けて今後とも取り組む。滋賀大学データサイエンス学部が育成したい人材は、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である（大学ポートレート，2017；竹村ら，2018）。大学生向けのデータ駆動型のPBL演習の授業デザインと教材開発についての共同研究を和泉志津恵教授（データサイエンス学部、以下DS学部）が畑山満則教授（DS学部特別招聘教授、京都大学防災研究所）の協力のもと2017年度にたちあげた。データサイエンス学部1年生向けに「統計学要論」の授業デザインと教材を開発し、授業での実践を始めた。

1-3 高大連携

虎姫高校と滋賀大学は、教育や研究の分野で連携を深めることを目的に、2016年度に連携協定を締結した（滋賀大学，2016）。この協定に基づいて、2017年度からSSH運営委員会委員として和泉が派遣された。また、三上校長と和泉が企画

し、2017年6月に滋賀県高等学校教育研究会数学会 平成29年度総会を滋賀大学にて開催した。総会では、和泉と田中琢真准教授(DS学部)が「データサイエンス学部での初年次教育および高大接続授業について」講演を行った。また、滋賀大学と総務省統計研究研修所は2017年から毎年8月に滋賀大学大津サテライトにて「教育関係者向けセミナー」を共催した。虎姫高校の教員らは、このセミナーにも参加した。また、ほぼ同時期に虎姫高校は国際バカロレア(IB)ディプロマプログラムの候補校となり、2019年3月にIBワールドスクールの一員として認定された。このようにして、虎姫高校と滋賀大学との双方において、データサイエンス教育の普及について素地が整ってきた。

一方、SSH運営委員会の定期会合での意見交換を重ね、松宮敬広 教諭(2012年度から2019年度までSSH運営委員会主任)と和泉とが、高校生向けのデータ駆動型のPBL演習の授業デザインと教材開発について共同研究を開始した。滋賀県HPには、全国に先駆けて県全域での水害リスクを示した地先の安全度マップを公開している。2017年8月の台風5号により滋賀県長浜市内を流れる姉川が氾濫し、虎姫高校の通学路の一部が閉鎖された。自然災害の被害を受けた虎姫高校の生徒や周辺住民は自然災害の減災に興味がある。この地域の課題に注目して、防災教育とデータサイエンス教育を組み合わせた授業を開発することとした。

高大連携授業を設計するため、授業のレベルのすり合わせを双方の教員で行った。まず、滋賀大学の学部1年生の「統計学要論」の授業と虎姫高校の2年生の「究理Ⅱ」の授業の様子を双方の教員が見学した。次に、「統計学要論」で扱うデータ活用演習の内容を高校生向けに簡素化し難易度を調整することにより、「究理Ⅱ」Dコースの授業を設計した(Izumi & Hatayama, 2018; Takata, et al., 2018; 和泉ら, 2016; 竹村ら, 2018)。そして、2018年度から「究理Ⅱ」Dコースにおいて、データ駆動型の授業デザインを実践し始めた。

この他にも、虎姫高校2学年生徒の希望者を対象として、データサイエンスに関わる講義形式の講座(文系生徒向け)や、野外実習を含む講座(理系生徒向け)「地先の安全度マップのための河川地形モデルの作成」など、両者が連携した取組を2018年度からサマーセミナーとして実施している(和泉ら, 2019)。野外実習を含む2日間の講座では、生徒らは、実習で作成した画像データと滋賀県が公表している地先の安全度マップのデータをGISで

重ね合わせ、水害リスクについて議論する。この議論からも防災や減災について新たな気付きや解釈が生まれる。ドローンを用いた野外調査の実施にあたり、滋賀県土木交通部流域政策局、長浜市役所、長浜市大井自治会に協力いただき、高校生のデータサイエンス学習を見守っていただいている。

2. 研究の目的

データ駆動型の探究活動を導入することで、高等学校において、より多くの生徒を対象として実施可能な探究授業カリキュラムを構築する。あわせて、データ駆動型の探究活動によって、探究に必要な力や態度についての生徒の理解の深まりを分析・検証する。

なお、本稿では、虎姫高校における学校設定科目「究理Ⅱ」Dコースのカリキュラムに関わる内容を報告する。

3. 取組の展開と評価の手法

3-1 取組の展開

「究理Ⅱ」Dコースの単位数は1単位で、例年、2～3クラスに分かれて、各クラスを担当者3名の内2名の教員でペアになり担当している。担当教員の専門は、理科(松宮、責任者)、数学、地歴公民である。履修人数は、2021年度では68名である。

授業は「情報処理実習」、「統計的探究基礎実習」、「探究活動」の3つの段階にわけて展開した(表1)。

表1 「究理Ⅱ」Dコースの授業展開

月	授業内容
4～6月	情報処理実習 ・ガイダンス ・表計算ソフトウェアの使用法
6～9月	統計的探究基礎実習 ・基本統計量 ・標本調査、信頼区間 ・相関関係、散布図 ・研究倫理
10～2月	探究活動 ・テーマと仮説の設定 ・データの収集と分析 ・ポスターの作成と発表

情報処理実習では、表計算ソフトウェアを用いた図表の作成や関数の入力など、データ処理の基本操

作の授業を行った。教員の説明に従って、Microsoft Excel 2010（以下 Excel）ファイルで与えられるワークシートを完成させていくスタイルで授業を進めた。

統計的探究基礎実習では、データ分析の基本的概念のうち、後半の探究活動で活用する可能性が高いと予想されるものを選び、それらの理解を深めるために実習形式の授業を行った。この実習では定義や原理の確認よりも、データ分析の基礎知識を身に付け、身近な問題解決に生かす力を養うことに主眼を置いて、授業を行った。生徒が Excel ファイルで与えられる課題に取り組み、教員は生徒の解答に対して解説を加えるという形で授業を進めた。

探究活動については、前述の通り、滋賀大学 D S 学部の 1 年次で開講されている「統計学要論」を参考に設計した（Izumi & Hatayama, 2018；Takata, et al., 2018；和泉ら, 2016；竹村ら, 2018）。「統計学要論」における探究活動は、市町村ごとにウェブ上で一般公開されているテキストデータ（地域防災計画）から、文字数や語句数などの量的データを抽出し、それを基本データとしてデータ解析を行う点が特徴的である。この授業展開にならって、以下のように授業を進行した。

(1) ガイダンス

「地域防災計画の特徴を探れ！」をミッションとして掲げ、探究活動の目的や地域防災計画について簡単な説明を行った。

(2) 基本データベースの作成

教員が示した一定の基準にもとづいて、47 都道府県から 1 つずつ市町村を抽出し、それを全員の生徒に割り当てた。各生徒は、担当市町村のテキストデータをダウンロードし、Microsoft Word（以下 Word）のファイルへと変換した。その後、Word の検索機能等を用いて、ファイルの最新更新年、テキストの総文字数、指定した語句（「河川」「地震」「トラフ」「噴火」「雪」「津波」「停電」「地区防災計画」「隣接」「ハザードマップ」の 10 種類）の出現数等のデータを収集した。全員のデータを 1 つにまとめて、探究の基本資料となるデータベース（基本データベースと呼ぶ）を作成した。なお、初年度（2018 年度）にこの方法を実施したところ、データベース作成に時間がかかり、後半の探究活動に十分な時間を費やすことができなかった。そこで、2019 年度以降は、初年度の生徒が作成したデータベースを使って探究活動を進めた。ただし、2019 年

度以降についても、探究班のテーマによっては、基本データベースを使わずに生徒がその都度データを集めたケースもあった。

(3) テーマの設定

生徒 4 ～ 5 名からなる探究班を編成し、班ごとに地域防災計画の課題に即したテーマを設定した。限られた時間で焦点化されたテーマを設定できるように、「テーマ探索シート」と「研究計画書」の 2 つのワークシートを作成して活用した（表 2）。

表 2 テーマ設定を支援する工夫

ワークシート	内容
テーマ探索シート	個人単位でテーマと仮説を 3 つ以上考えて記入する。適切なテーマや仮説の立て方についてのレクチャーを行った後、アイディアの掘り起こしのために用いた。
研究計画書	具体的な記入項目からなるワークシートで、これに書き込むことで探究の段取りが自然と見えてくるように項目を設計した。記入項目として、「研究テーマ」「疑問」「目的」「仮説」「仮説の根拠」「必要なデータ」「データの入手先」などを設定した。班単位で作成させた。

(4) データの収集・分析

情報処理実習や統計的探究基礎実習で学習した内容を活用しながら、班ごとにデータの収集と分析を行った。また、データの入手先として e-Stat や気象庁のデータベースを紹介、解説した（総務省統計局, 2008；国土交通省気象庁, 2018）。生徒らは、地域防災データと他の関連データを組み合わせ、Excel でデータ加工や解析を行った。

(5) ポスターの作成と発表

生徒らは、データ解析の結果から防災や減災について新たな気付きや解釈を生み出し、探究内容をまとめてポスターを作成した。個人の技能向上という観点から、個人単位での作成とした。完成したポスターは教員がルーズリーフにもとづいて採点し総括評価に用いた。班内で個人ポスターの相互発表を行い、それを踏まえて、班代表のポスターを 1 つ決

定し、最終的には班単位で発表を行った。2時間連続の授業時間を設定して、発表会を開催した。発表会には滋賀大学の教員や学生らが参加し、双方の教員や生徒らが交流できる機会とした。

生徒らは、この「究理Ⅱ Dコース」をとおして、身の回りにあるデータを意識する、異なる分野のデータの組み合わせを考える、解析方法の組み合わせを考える、データや解析結果の視覚化を工夫することを経験した。また、虎姫高校の教員らも、新たな教育分野であるデータサイエンス教育を実践する経験を積み重ねた。

3-2 アンケートによる分析

探究活動を通じて、探究に必要な力が身についたかどうかの検証は、生徒の自己評価によって行われることが多い。まず、2018年度の生徒(83名)を対象に、情報処理実習や統計的探究基礎実習の内容を探究活動に役立てることができたかどうかのアンケートを実施した。具体的には、以下の5つの項目ごとに4件法(そう思う、どちらかといえばそう思う、どちらかといえばそう思わない、そう思わない)のアンケート調査を行った。

- ・表計算ソフトの使い方
- ・基本統計量(平均値、標準偏差など)
- ・標本平均の信頼性(標準誤差、信頼区間など)
- ・関係性の分析(相関関係、回帰直線など)
- ・研究倫理(個人情報、著作権、引用など)

しかしながら、これまでの虎姫高校におけるSSHの取組とその評価の中で、探究的な能力を高いレベルで発揮した生徒が、比較的低い自己評価をつける傾向が観察された(虎姫高校、2018)。これは、高いパフォーマンスを発揮する生徒は、探究に関わる能力についての評価基準をより明確に持っており、そのため自己評価が厳しくなると考えたと説明がつく。

そこで、探究活動の経過とともにどのように探究活動に対する理解が深まるのか、また、評価基準が形成されるのかについての手がかりを得るために、生徒の自由記述文の分析を行った。具体的には、2020年度の生徒(68名)を対象に、探究活動の実施前(9月)と実施後(3月)に、1)研究課題の設定、2)研究の進行、3)研究結果の発表に関する質問項目について自由記述式のアンケートを行った。アンケートの時間はいずれも15分間とした。自由記述のテキストデータの解析では、単語の抽出と計数にKH Coder(樋口、2020)を用いた。

4. 取組の分析

4-1 実習や講義が探究の役に立ったか

実習が探究活動に役立ったかを調べた2018年度のアンケート結果では、5つの項目のうち4つにおいて、肯定的回答(そう思う、どちらかといえばそう思う)の割合が90%を超えた(図1)。このことから、探究活動前の実習が探究活動を進める上で、有効に機能していたことが推察できる。ただし、標本平均の信頼性については他の項目と比べて肯定的回答の割合が低かった(図1)。これは、平均値の比較を行うような研究計画が立てにくかったことが原因として考えられる。多くの班が相関分析を行っていたことから、関係性の分析については肯定的回答の割合が非常に高かった(図1)。

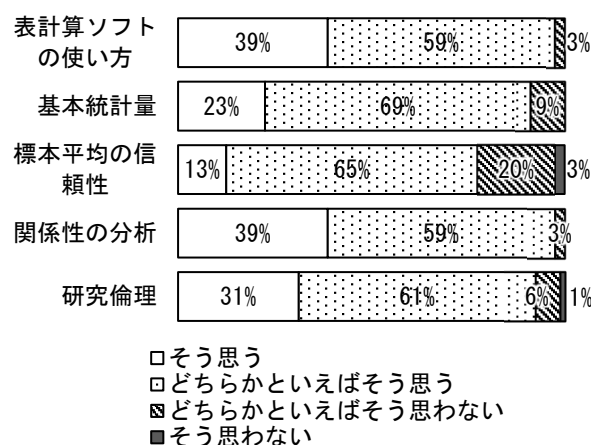


図1 実習内容を探究活動に役立てられたかについてのアンケート結果

4-2 探究に必要な力に関する生徒の認識の変容

探究に必要な力に関する生徒の自由記述アンケートにおいて、すべての質問項目について、探究活動の実施前よりも実施後の方が生徒の記述における総抽出語数や異なり語数(単語の種類数)が増加した。このことから、探究活動を経験したことで、生徒は探究に必要な資質・能力について、より詳細あるいは多面的に説明できるようになったことが推察できる。3つの質問項目に関する詳細な解析結果について、著作権の関係から本原稿では省略する。

5. おわりに

本研究において、大学で開発された授業をアレンジすることで、高校でも実施可能なデータ駆動型授業の展開例を構築することができた。また、生徒アンケートの結果から、学習した内容を役立てながら

生徒が探究活動を進められたことが分かった。さらに、自由記述アンケートの分析からは、探究活動を経験することで、探究に必要な力や態度に対する生徒の理解が深まった可能性が示唆された。

防災とデータサイエンスの専門家が直接、高校の教育現場にてデータ駆動型の授業や実習を行うことや通年授業の支援を行うことは、虎姫高校や滋賀大学としても過去に例のない初の試みとなった。データサイエンス教育の共同研究に発展したこの活動は、大学と高校と自治体の社会連携の一つのロールモデルとなる。

現在、この授業を実施して4年目であるが、毎年70名程度の生徒を3名の教員で指導しており、教員が入れ替わりながら、過度の負担なく継続的に実施できている。このことは、本授業デザインに次の2点の特徴があるためだと考えられる。

- (1) データ駆動型の探究活動はインターネットにつながったコンピュータがあれば実施できるので、活動に場所を選ばない。放課後や自宅で生徒が自主的に探究を進めることもできる。
- (2) テーマの設定において、使用するデータなどを部分的に固定して自由度を少し減らすことと、研究計画書等のワークシートを導入することによって、探究の指導にかかる負担が軽減されている。

一方で、このような指導の効率性と引き換えに、探究テーマの広がりや、専門的な深まりについては、必ずしも十分とはいえない側面がある。持続可能な探究活動プログラムを開発する上で、指導の効率性は重要な観点ではあるが、このことと、探究の広がりや深まりをどのようにして両立していくかは、今後の課題である。

虎姫高校と滋賀大学は、地域の実課題に挑戦する実践的なデータサイエンス教育、共同研究への参画、提案、データ活用の高度化によりデータサイエンスの社会実装に貢献していく。

謝辞

本研究は、統計数理研究所・共同研究利用・重点研究「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), 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」

(2021-統数研-共同研究集会-00008, 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, 代表者 和泉志津恵)の助成を受けた。サンプルデータの整備については、滋賀大学大学院経済学研究科やデータサイエンス学部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. 国立研究開発法人科学技術振興機構. (2019). 『平成30年度スーパーサイエンスハイスクール情報交換会』
2. 文部科学省(2018). 高等学校学習指導要領.
http://www.mext.go.jp/component/a_menu/education/micro_detail/_icsFiles/afieldfile/2018/07/11/1384661_6_1_2.pdf
3. 滋賀大学(2016). 『県内高校(彦根東高校・虎姫高校)との高大連携協定を締結しました』
<https://www.shiga-u.ac.jp/2016/12/12/44901/>
4. Izumi, S., Hatayama, M. (2018). Overview of an educational model of elementary statistics for the data science program of Shiga University. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018) 1A1: 1-6.
5. Takata, S., Izumi, S., Takemura A. (2018). Education of data science in Japan - Shiga university model -. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018) 5C3: 1-6.
6. 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.
7. 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.
8. 和泉志津恵, 畑山満則, 松宮敬広. (2019). 地域に根ざすデータサイエンス教育 -SDGsの展開-. 「統計教育実践研究」 Jissen Women's Educational Institute (JWEI) (Special Issue) 21-26.
9. 松宮敬広, 和泉志津恵, 畑山満則. (2019). 高校生のためのデータ駆動型の授業デザイン -防災エビデンスの探索-. 「統計教育実践研究」 Jissen Women's Educational Institute (JWEI)

- (Special Issue) 47-52.
10. 総務省統計局. (2008). 『e-Stat 政府統計の総合窓口』
<https://www.e-stat.go.jp/>
 11. 国土交通省気象庁. (2018). 『各種・データ資料』
<https://www.jma.go.jp/jma/menu/menureport.html>
 12. 滋賀県立虎姫高等学校. (2018). 『平成29年度指定スーパーサイエンスハイスクール研究開発実施報告書(第2年次)』2018年3月発行.
 13. 滋賀県立虎姫高等学校. (2019). 『平成29年度指定スーパーサイエンスハイスクール研究開発実施報告書(第3年次)』2019年3月発行.
 14. 滋賀県立虎姫高等学校. (2020). 『平成29年指定スーパーサイエンスハイスクール研究開発実施報告書(第4年次)』2020年3月発行.
 15. 滋賀県立虎姫高等学校. (2021). 『平成29年指定スーパーサイエンスハイスクール研究開発実施報告書(第5年次)』2021年3月発行.
 16. 滋賀県立虎姫高等学校. (2022). 『平成29年指定スーパーサイエンスハイスクール研究開発実施報告書(第6年次)』2022年3月発行予定.
 17. Izumi, S., Hatayam, M., Matsumiya, T., Ishii, H. (2021). Data science education practice that fosters zest for living - Disaster prevention data analysis exercise - . Proceedings of Disaster Prevention Research Institute (DPRI), Kyoto University Annual Meeting 2021. E306: 1-3.
 18. Izumi, S., Hatayama, M. (2020). Utilization of large-scale disaster prevention data in data science education and research. Proceedings of Disaster Prevention Research Institute (DPRI), Kyoto University Annual Meeting 2020. E27: 1-4.
 19. 高大接続, 滋賀県立虎姫高等学校 理系サマーセミナー. (2020). 滋賀大学 DS センター報「Data Science View」, Vol.4, P.39.
<https://www.ds.shiga-u.ac.jp/dscenter/about/#report>
 20. 高大接続:SSH 連携校, 滋賀県立虎姫高校. (2021). 滋賀大学 DS センター報「Data Science View」, Vol.5, P.48.
<https://www.ds.shiga-u.ac.jp/dscenter/about/#report>
 21. 高大接続:SSH 連携校, 滋賀県立虎姫高校. (2022). 滋賀大学 DS センター報「Data Science View」, Vol.6 (予定).
<https://www.ds.shiga-u.ac.jp/dscenter/about/#report>
 22. 滋賀大学 (2019). 滋賀県立虎姫高校「究理 II データサイエンスコース」発表会にて講評.
https://www.ds.shiga-u.ac.jp/dscenter_info/p4125/
 23. 滋賀大学 (2020). データサイエンス学部の和泉志津恵教授による実践教育が、@DIME (2020年6月3日付) に掲載されました.
<https://www.ds.shiga-u.ac.jp/news-faculty/p5206/>
 24. 樋口耕一. (2020) 『社会調査のための計量テキスト分析 ―内容分析の継承と発展を目指して― 第2版』. ナカニシヤ出版.

1. はじめに

新学習指導要領により数学Ⅰ「データの分析」において仮説検定の考え方が扱われることとなった。既に教育現場には各教科書出版社から見本本が届けられ、学校現場での教科書選定が済んだ。過去の数学教育において仮説検定が扱われていた時期もあり、その変遷は既によく知られているところである。今般の仮説検定の取り扱いの最も重要な点のひとつは、数学Ⅰという必修科目で扱われるという点である。将来的な入試への出題の見込みは未知数であるものの、すべての高校生がそれを学ぶこととなる。また、これまで推測統計の分野を担当していた教員は現場でもごく僅かと推測されるが、今次の取り扱いにより多くの数学教員が指導する。「中心極限定理は数学だが、仮説検定は手続きであって数学でない」という言葉を若い数学教員から聞かされたことがあるが、尤もな感覚であろうと思われる。本発表では筆者が勤務校で学校設定科目としてこれまで推測統計を扱った授業を実践してきた経験を踏まえて、これら仮説検定の扱いにかかわる期待と懸念を記す。

2. 数学Ⅰ「データの分析」における「仮説検定の考え方」

学習指導要領が改訂され、総合的な学習の時間は総合的な探究の時間となった。探究活動の中で、先行研究の読みや、エビデンスベースの主張に仮説検定の考え方の履修が有用である。

2022 年より使用される教科書での仮説検定の考え方の扱いは、数学 A における確率の学習も前提とせず、反復試行の確率や確率分布の知識を仮定しない。こうした制約の下で教科書の記述も出版社により扱いに共通点と違いがある。

今般主に 5 社の教科書について内容の記述を見たところ、ある状況場面での比率に関する主張の正当性を相

対度数により確率的に判断する点が共通である。そして、通常確率分布に代わるものとしてコイントスなどに基づく分布が示されていることも共通である。状況設定は様々であるが、良い・悪い、正しい・正しくない、の 2 値の評価の比率の場面としている。取り扱う頁数は出版社や同じ出版社でも教科書のレベルによって 1 頁～3 頁までの幅がある。これらの内容は学習指導要領解説で例示された「新素材枕の評価」の事例に倣っているものとも言える。統計的な用語の扱いは各社で差異があり、「仮説」について概ね「仮説」に留めているが、欄外で「帰無仮説」や「対立仮説」が記されているものもある。また、ある社では「棄却域」が記載されていたり、発展として反復試行の確率がとりあげられている。正確確率の仮説検定では、めったに起こらない確率や相対度数について、単純にその場合のみを考えるのではなく、その場合を含む上位和をとることの説明や理解が難しいが、これについて状況設定で最も極端な場合を例題とする工夫も見られる。

仮説検定そのものの定義の説明は「あるデータが与えられたとき、仮説を立て、それが妥当かを判定する手法を仮説検定という。」「実際に起こったことがらから、ある仮説のもとで起こり得るかどうかを考えることによって、仮説が誤りかどうか検証する方法を仮説検定という。」など端的なものであるが、およそその手順を紹介しているものもある。

3. 数学 B「統計的な推測」における「仮説検定」

2021 年現在見本となる教科書も未版のため学習指導要領解説の記述と過去の指導内容を参考に推測するに留まる。学習指導要領解説では知識・技能に関して「(エ) 正規分布を用いた区間推定及び仮説検定の方法を理解すること。」と書かれており、Z 検定による母平均の検定及び母比率の検定が対象となることが推測される。また、係る用語を抜き出してみれば、仮説検定・有意

水準・母平均に関する検定・対立仮説・帰無仮説などが並ぶ。また仮説検定の前提としての標本抽出に関して、無作為性を担保する例として「クラスター抽出法・2 段抽出法」が挙げられている。

4. 懸念

4-1 数学Ⅰの「データの分析」で扱われる仮説検定の考え方の指導上の懸念

- ・内容とその理解について（確率的判断の基準）

有意水準がなぜ 5%や 1%なのかを説明することは難しい。数字の根拠は明確に説明されず、5%と 1%の使い分けも経験的である。

- ・内容とその理解について（確率和について）

二項分布などの直接確率検定では当該のことがらの起こる確率のみならず、その発生頻度がより小さいであろう上位の確率和を計算して判断する。なぜ和をとるのかの説明や理解が難しい。

- ・内容とその理解について（信頼性について）

各社教科書での仮説検定の考え方は、例題に始まり、検定の手法や手続きが説明され、若干の問いや練習問題で終わる。例題での状況設定では標本の大きさは必ずしも大きなものでない。また必ずしも無作為に抽出された標本が対象とはなっていない。また、履修上の制約から、「起こりにくさ」の判断は、コイントスなどの事例との対比で行われる。こうした具象を基にした判断について、信頼性に疑問が生じる懸念がある。

- ・入試の取り扱いについて

数学Ⅰにおいて扱われる内容となることから、大学入学共通テストなどでの出題が可能となる。教科書の練習問題は例題を踏襲したものでその数も僅かである。さらに、数学 B「統計的な推測」の履修を考えれば、数学 B でより深く学ぶ学生とそうでない学生の受験上の有利不利も考えられる。入試の出題状況が将来の現場での取り扱いの軽重に大きな影響を及ぼす懸念が生じる。

- ・指導経験不足の懸念

必修履修科目である数学Ⅰで取り扱うことから、現場の多くの数学科教員が携わることになる。仮説検定の内容は空白期が長かったこともあり多くの教員が指導に不慣れであろうことが予想される。

- ・情報科とのかかわり

従前より統計教育の内容は数学科とともに情報科でも扱われてきた。今次の学習指導要領の改定により情報科の科目再編が行われ従前の必修履修科目は「情報Ⅰ」に再編された。具体的な指導項目として「名義尺度、順序尺度、間隔尺度、比例尺度」「質的データ、量的データ」「データに含まれる欠損値や外れ値」「テキストマイニング」「仮説検定の考え方」「相関係数、相関関係や因果関係」「交絡因子、単回帰分析」などが挙げられている。これによれば、情報科においても仮説検定が扱われる可能性がある。

4-2 仮説検定を扱った授業実践からの経験的な懸念

- ・ロジックに慣れない

仮説検定における、「仮説の下ではほとんど起こり得ないことが起こったことになり、仮説が誤りであったと判断する」というロジックに慣れない。

- ・仮説の採択

帰無仮説を棄却し、対立仮説を採択する場合と対になる場合、帰無仮説を採択するのか、対立仮説を棄却するのか、結論を留保するのか、の定着が難しい。

- ・推定との混同

仮説検定と同時期に統計的推定を学習した場合に、しばしば仮説検定との混同が見られる。信頼区間に入るかどうかによって検定しようとする。

5. まとめ

高校生が仮説検定の考え方を学ぶ時代の到来の期待と共に、ここに記した懸念の他にも実際に生徒のレポート課題などからいくつかの指導の難しさを知ることができる。実際の指導をもとに様々な経験が交流されることが望まれる。

附記：本研究の一部は JSPS 科研費 JP21H03958 の助成を受けている。

参考文献：塩澤友樹「高等学校数学科における新内容「仮説検定の考え方」の扱いと今後の展望」第 17 回 統計教育の方法論ワークショップ・理数系教員授業力向上研修会,2020
・文部科学省.「高等学校学習指導要領（平成 30 年告示）解説数学編理数編」,2019

・稲葉芳成,酒井淳平「高校生を対象とした推測統計の指導上の留意点について―課題学習を通して見た推測統計指導におけるいくつかの課題―」,数学教育学会誌 / Vol.59 / No.3・4,2018

神戸大学のデータサイエンス分野における高大連携の取り組み

神戸大学数理・データサイエンスセンター¹センター長

齋藤政彦(mhsaito@math.kobe-u.ac.jp)

神戸大学附属中等教育学校²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

神戸大学附属中等教育学校³数学科教諭

林 兵馬(h-hayashi@people.kobe-u.ac.jp)

1. はじめに

神戸大学では、「学理と実際の調和」を理念として、先端研究・文理融合研究で輝く卓越研究大学として、世界最高水準の教育研究拠点の構築と新たな価値の創造に挑戦し続けています。数理・データサイエンスセンターは、企業や自治体との包括連携や共同研究により社会実装や価値創造を進め、関西地区・兵庫地区・神戸地区でのデータイノベーションを推進します。データサイエンスに基づいた課題解決や価値創造を可能にするデータイノベーションの拠点を構築します。さらに、国内外の研究機関、他大学との連携を進め、数理・データサイエンスの世界的な拠点を目指します。

また、数理・データサイエンスセンターでは、わが国の発展を支えるデータサイエンス人材を輩出するため、中高生を対象としてその所有する3つの資源（データサイエンスに関する知(Knowledge)、情報基盤(Information)、技術(Technology)）を有効活用し、大学における学習の基礎となる統計的思考力、判断力、新たな発想力を養う機会を提供しています。

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などを通して、文部科学省より認定されたスーパーサイエンスハイスクール事業(SSH 事業)の運営にも協力し、その成果を教員対象の研修会や授業研究会の実施、生徒対象のPBL コンテストの開催などを通じて、全国の高等学校に成果を還元していく予

定です。

R3 年度については、高大連携事業では全国の生徒・教員に対して、「第1回中学生・高校生データサイエンスコンテスト」「高校教員向け統計研修会」を行いました。

また、附属中等教育学校の希望者の生徒を対象に「神戸大学1年生向け「データサイエンス基礎」の受講」「Python 講習会」

を提供しました。本ワークショップでは、上記の神戸大学数理・データサイエンスセンターが提供した高大接続事業に関して説明を致します。

第1回 中学生・高校生データサイエンスコンテスト

令和3年度 高等学校教員向け統計研修会

¹ <http://www.cmds.kobe-u.ac.jp/>

² <http://www.edu.kobe-u.ac.jp/kuss-top/index.html>

「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統計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のデータを貸与しますので、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。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。

■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 人でお願いします。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、24:00】までにご提出をお願いします。

2 月：「統計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

登録後に生成される誓約書およびメンバー表に署名と代表者（教員）の捺印の上（生徒の捺印は自筆で代替）、事務局まで郵送。

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 14 巻 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 457

Research on best practice in teaching statistics
Vol. 14

March, 2022

The Institute of Statistical Mathematics
10-3 Midori-cho, Tachikawa
Tokyo Japan 190-8562