

統計的探究プロセスとその評価

深澤 弘美¹・櫻井 尚子²・和泉 志津恵³

(受付 2017年6月30日; 改訂 9月19日; 採択 9月20日)

要 旨

Cyber-Physical System(CPS)によるデータ駆動型社会では、あらゆる業種や領域においてビッグデータをはじめとするデータ活用による意思決定プロセスが必須となる。インターネットの普及やモバイル社会、IoT、AIの進化により、データの蓄積は爆発的に増殖を続け、従来、勘や経験に頼っていたあらゆる人々の選択や判断は、データの分析結果を根拠に行われる。このような時代に生きる必要条件として、統計を用いて問題を探究し、科学的に判断する力の重要性はますます高まっている。本稿では、科学的探究力・科学的判断力の育成を目指した統計教育の在り方、具体的な授業のデザイン、評価の方法について、イギリス、アメリカ、ニュージーランド等の海外の事例を含めて調査・研究した結果をまとめる。科学的探究や判断を実験させる授業デザインとしては、統計的探究プロセスに沿って問題解決を行うプロジェクト型の授業を提案する。評価については、統計的探究プロセスの各フェーズにおける生徒や学生の理解や能力を評価する基準として「統計的問題解決評価ルーブリック SPART」を提案する。

キーワード：統計教育，統計的問題解決，PPDACサイクル，探究力，アセスメント。

1. はじめに

Cyber-Physical System(CPS)によるデータ駆動型社会の到来(図1)を受けて、ビジネス界のみならず、医療や環境、行政などあらゆる分野でビッグデータをはじめとするデータ活用による意思決定プロセスが重要視されている。CPSとは、蓄積された実世界のデータを分析・解析し、その結果を再び実世界にフィードバックするサイクリックなシステムを指す(経済産業省, 2017a, 2017b)。インターネットの普及やモバイル社会の到来で、知らぬ間に情報は収集され、自動的にデータベースとして日々刻々蓄積されている。さらなるIoTやAIの進化により、データの蓄積は爆発的に増殖を続け、従来、勘や経験に頼っていたあらゆる人々の選択や判断は、データの分析結果を根拠に行われる新たな時代が到来している。CPSにより変革を余儀なくされるデータ駆動型社会では、様々な分野で従来存在しなかった価値を創造する手段を提供している(図2)。これが進んだ先の社会は、全ての産業間でデータを核としたビジネスモデルの革新を招き、それがいずれ産業構造の大変革を引き起こすことが予想される。この状況を踏まえて経済産業省では、国際競争に遅れない企業のチャレンジ促進環境の整備を大きな課題としてとらえている。その中でコアテクノロジーの進展とともに掲げられているのが、人材育成の強化であり、その一端を担うのがデータサイエンティストである。ITを駆使して大量の

¹ 東京医療保健大学 医療保健学部：〒154-8568 東京都世田谷区世田谷 3-11-3

² 東京情報大学 総合情報学部：〒265-8501 千葉県千葉市若葉区御成台 4-1

³ 滋賀大学 データサイエンス学部：〒522-8522 滋賀県彦根市馬場 1-1-1

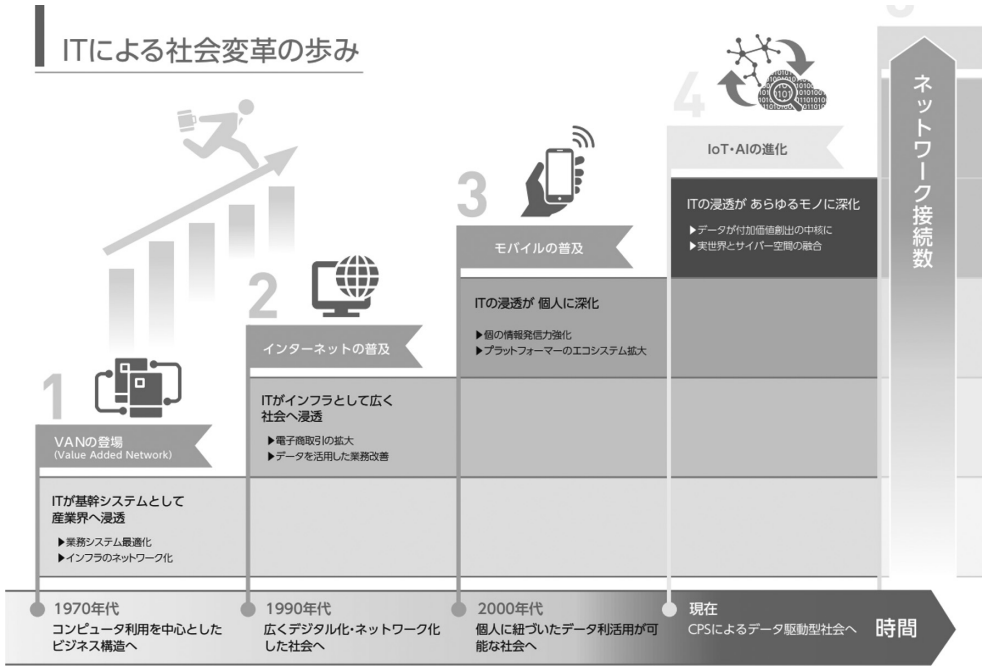


図 1. IT による社会変革の歩み (経済産業省, 2017a).

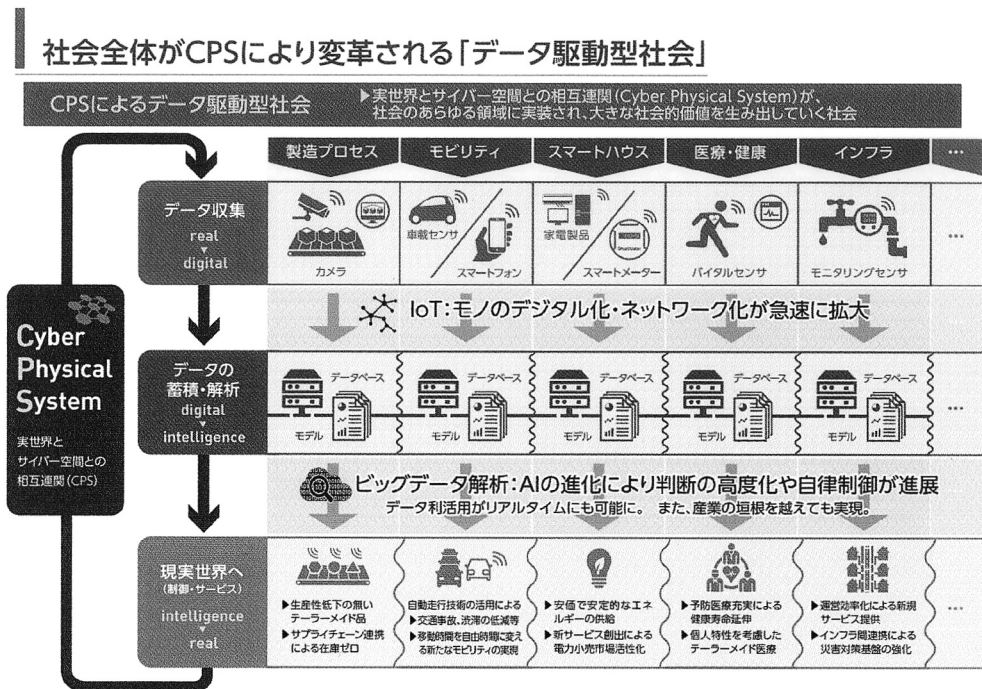


図 2. 社会全体が CPS により変革されるデータ駆動社会 (経済産業省, 2017b).

データから必要な情報を取捨選択し、統計学を用いて問題を解決する作業は、各種アプリケーションプログラムや人工知能とともにチームワークとしての作業量を要求される。各領域のデータサイエンティストに求められる素養は深く、広域にわたることが想定される。専門家の育成が急務であると同時に、学校教育においても情報教育と並行して統計教育の充実を図り、ビッグデータ時代に生きる市民の力を育成していかななくてはならない。

本研究では、諸外国の統計教育およびその評価の事例等を参考に、統計的探究プロセスを体験し、実際にデータを用いて問題解決を行うプロジェクト形式の授業のデザインとその評価方法を提案する。まず2節で統計教育の現状と課題をレビューする。続いて3節では、統計的探究プロセスについて述べる。4節では、3節で述べた統計的探究プロセスの授業デザインについて提案する。5節では統計的探究力の評価をレビューする。6節では、プロジェクト型統計的問題解決学習の評価基準を提案し、最後に7節でまとめを行う。

2. 統計教育の現状と課題

わが国では2020年に向けて新センター試験、次期学習指導要領の検討等が進められ、あらゆる科目においてデータ重視の変革が行われる。イギリスの「Scientific Inquiry」、アメリカの「Scientific Investigation」、ドイツの「Nature of Science」のような数学という科目の枠を超えた統計教育の充実という世界的動向をうけて、日本でも新科目「数理探究(仮称)」の新設が検討されている。統計関連学会連合は「高度情報化・データサイエンス社会における国民の科学的判断力・探究力強化に向けた新科目『数理探究(仮称)』内容への要望書」を文部科学省に提出し、“国際的な共通基準とすべき通用性に照らして内容を充実し、具体的には推測統計を含む科学の手法としての統計的仮説の構築、実験・調査のデザイン、データ処理、モデリング、結果の解釈に至るまでの統計学的体系の学習内容を充実すること”を要望した。具体的な問題に対して客観的なデータを集め、適切な統計の手法を活用して分析し、得られたエビデンスに基づいて判断して問題の解決策を提案するという、統計的探究のプロセスが重要であり、このプロセスをプロジェクトベースで体験させることを繰り返し授業に取り入れることが科学的思考力や問題解決力の育成につながる。入試を含めて初等教育、中等教育、高等教育の統計教育は変革の時であり、教材等授業の充実に向けた研究はますます重要となる。そしてこれらの教育を充実させるためには、その評価方法や評価基準を確立することも喫緊の課題である。そこで本研究では、諸外国の統計教育およびその評価の事例等を参考に、統計的探究プロセスを体験し、実際にデータを用いて問題解決を行うプロジェクト形式の授業のデザインとその評価方法を提案する。

本稿で提案する授業デザインは、平成24年8月の中央教育審議会答申「新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学へ～」に記されている、学士課程教育の質的転換の内容(中央教育審議会, 2012)ならびに、平成29年3月の新学習指導要領改訂のポイントに示された“必要なデータを収集・分析し、その傾向を踏まえて課題を解決するための統計教育の充実”に準じている。すなわち、受動的教育が主流だった従来型授業の枠を取り外し、教員と学生が意思疎通を図りつつ、一緒になって切磋琢磨し、相互に刺激を与えながら知的に成長する場を創り、学生が主体的に問題を発見し解を見出していく能動的学修(アクティブラーニング)を踏まえた統計教育を早急に拡充しなければならない。ビッグデータ時代においてグローバルに活躍する人材を育成するためには、統計で問題を解決するためのプロジェクト型授業を初等教育の早い段階から大学まで継続的に深めていく学びが必須と考える。

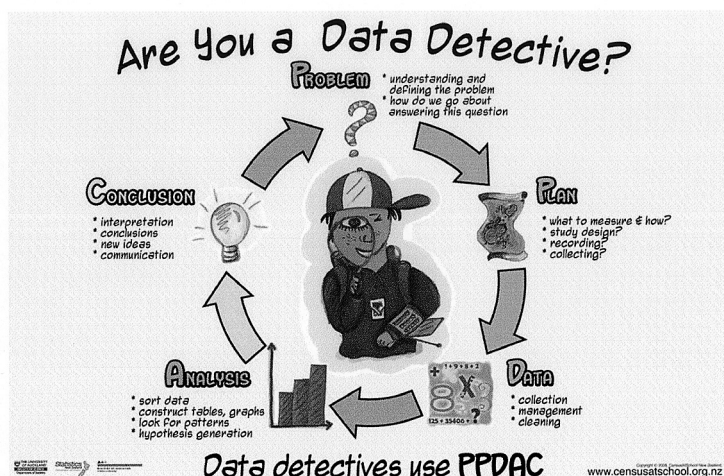


図 3. PPDAC サイクルのポスター(ニュージーランド).

3. 統計的探求プロセス

統計を使って問題を解決するために、問題を捉え、データを分析して結論を導き出す一連のプロセスを統計的探究プロセスという。論理的なプロセスで科学的かつ協働的に問題を解決する力は、21世紀を生きる若者に必須の力であり、世界共通で必要な力としてその教育が世界中で重要視されている。業種を問わずあらゆるビジネスや生活の場面で必要な力である。統計的探究プロセスには、国際的な枠組みとして PPDAC サイクルがある。Problem(問題をとらえる)、Plan(計画を立てる)、Data(データを集める)、Analysis(分析する)、Conclusion(結果を読み取る)の5つのフェーズを繰り返しながら問題を解決するステップを定義したものである。各フェーズの頭文字をとって PPDAC サイクルと呼ばれている。図3はニュージーランドで作成された統計の重要性を子供たちに示すポスターである。ニュージーランドでは、初等教育の初期段階から身近な問題や身の回りの現象に関する疑問や課題を PPDAC サイクルに沿って解決する授業が繰り返し行われている(Education in New Zealand, 2017; Te Kete Ipurangi (TKI), 2017)。

4. 統計的探究プロセスと授業デザイン

Gal and Garfield (1997)は、統計教育の目標を以下の2点であると主張している。

- 不確実性、ばらつきについて理解し、身の回りにある統計情報に正しく対処することによって情報満載の現代社会に効率よく参加できる市民の育成
- 各自が直面するより専門的な課題に関連して、データの提示・説明・伝達に貢献および参加できる市民の育成

我が国の学校教育においても、さまざまな統計的問題解決のプロセスおよびその概念の基礎となる数学的なアイデアや理屈を理解して活用し、分析した結果を説明し伝える力を養うための授業作りが重要となる。統計教育を行う教員は複合的な目標に向けて、概念的に異なる幅広い視点で指導を行う必要がある。生徒及び学生が判断・まとめ・記述・予測等を行うことに慣れるための手助けをする授業デザインが重要であろう。

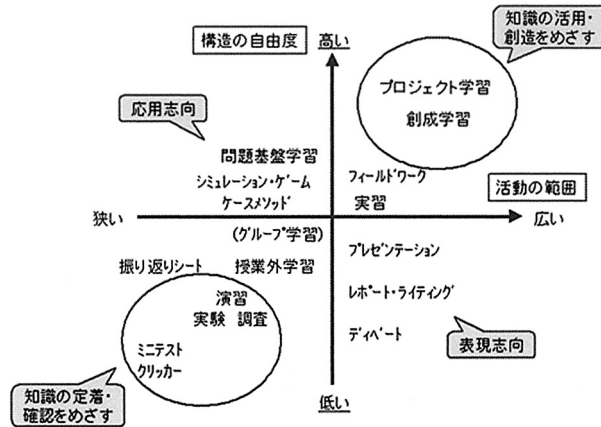


図 4. アクティブ・ラーニングの多様な形態 (山地, 2014).

知識理解だけを目的とするのではなく、習得した知識を使って主体的かつ協働的に問題を解決する力を身に付けるためには、アクティブラーニングの導入やMOOC(Massive Open Online Courses)やJMOOC(Japan Massive Open Online Courses)を用いた授業外での学習、反転学習など新たな試みも必要となる。

4.1 アクティブラーニング

アクティブラーニングは、教員と学生間での相互刺激による知識の醸成と問題発掘・解決力の育成を目指す学習・指導方法である。この方法が通常授業の中でデザイン出来得る最適な時間、クラスの大きさ、方法の難易度などを考慮して、学生の学びに合った適正な学習指導方法を選択する必要がある。統計教育においてもアクティブラーニングを取り入れ、探求のプロセスを実体験することはとても重要である。アクティブラーニングを実質化するにあたり、キーになる7つの原則が米国で発表され、世界中で参照基準とされている(山地, 2014)。そのイメージを図4に示す。人数や教室の形、授業内容や専門性等、様々な制限は存在するが、図4における活動範囲を広げることで知識の定着や興味・思考の弾力強化を図ることが可能であり、一方向だけの伝達型授業よりアクティブな内容を盛り込める。専門性の高さによって、図4の各象限内の実施を検討する必要がある。そこで、クラス編成のタイプや学年、専攻等を考慮したパイロットスタディモデルを企画した。内容は4.3節以降に示す通りである。

4.2 MOOCとJMOOC

オープンオンライン教育のプラットフォームであるMOOCは、2012年前半の実質的なサービスが開始から急速な発展を続け、特に統計学やデータサイエンス関連は社会の需要にあわせて多種多様なコースが開講されている。情報分野に特化したUDACITYやあらゆる領域をカバーしているCoursera、およびMITとハーバード大学によって設立されたedXはいずれも米国の代表的なMOOCである。特に、2017年6月現在で2,500万の受講者を持つCourseraは、149のパートナー教育機関と連携し、2,000を超す授業を提供している。このうち4つのコースでは、学位の取得が可能である。上記の米国発のMOOCの他にもヨーロッパやアジアで独自のMOOCが誕生しており、言語の壁を乗り越えて特色を生かした統計教育コンテンツが用意されている。MOOCでは特に、データサイエンスを冠した講義・演習が多く、オンライン講座

の特性を生かして企業と結び付いた実践的課題演習内容を課しており、アクティブラーニング型授業スタイルを通して実質的なデータサイエンティストの養成を手がけている。MOOCでは、オンラインで公開された無料の講座を受講し、修了条件を満たすと修了証が取得できる。ただし、修了証(Verified Certificate)には、Courseraのように有料なオプションとして提供されているものもある。

CourseraのData Scienceというカテゴリーはさらに、Data Analysis, Machine Learning, Probability and Statisticsの3カテゴリーに分類され、細分化された授業が用意されている。Data Analysisでは100の授業が開講されており、Introduction to Data Science in PythonやApplied Machine Learning in Python等Pythonの教育が一角を占めている。Big Dataもキーワードである。

MOOCでは言語の壁が問題であったが、2013年にJMOOCが設立され、2014年4月からgaccoの運用が本格化している。JMOOCが公認した配信プラットフォームには、ドコモgacco社が提供するgacco、ネットラーニング社が提供するOpenLearning, Japan、放送大学が提供するOUJ MOOCがある。gaccoにおいて統計関係では、「統計学Ⅰ：データ分析の基礎(日本統計学会)」「統計学Ⅱ：推測統計の方法(日本統計学会, 日本計量生物学会)」「社会人のためのデータサイエンス入門(総務省統計局)」「社会人のためのデータサイエンス演習(総務省統計局)」が開講されている(政府統計の総合窓口, 2017; 竹村 他, 2015)。また、2017年6月に初開講された「統計学Ⅲ：多変量データ解析法(日本統計学会, 日本行動計量学会)」が加わり、学修の幅が拡大している。MOOCやJMOOCによるオンライン授業においては、対面学習やオンラインディスカッションによりアクティブ性を保持している。

4.3 中等教育(高等学校)における授業のデザイン(90分*1コマ)

滋賀大学において高校生を対象に2016年10月に実施した高大接続授業の取り組みを紹介する。受講者は高校1年生から3年生の36名、担当者は教員1名と学生アシスタント6名である。授業名は「お茶の間にやってきたデータサイエンス」といい、90分1コマで、統計的探究プロセス(PPDACサイクル)を体験し、実際にデータを用いて問題解決を体験するプロジェクト型の授業である。授業の目標は、データサイエンスにおける身近な事例を用いて「PPDACサイクル」の活用の仕方を習得することであり、課題を解決する流れ「問題の発見(Problem)→調査の計画(Plan)→データ収集(Data)→分析(Analysis)→結論(Conclusion)」を学ぶ。授業の流れを次に示す。

- (1) イントロダクション
- (2) PPDACサイクルとは
- (3) 事例紹介

例題：育児休業、資料：内閣府男女共同参画白書

- (4) グループワーク

課題：エネルギー消費、資料：資源エネルギー庁エネルギー白書

- (5) まとめ

グループワークでは、受講者を6名ずつのグループに分け、各グループに学生アシスタント1名をつける。グループごとに、与えられた課題についてPPDACサイクルを組み立てA1サイズのポスターにまとめる。学生アシスタントは、グループでの議論を見守り、活動を補助する。なお、学生アシスタントは、この授業内容について事前学習をし、例題と課題に取り組んでいる。各テーブルには、インターネット接続が可能なノートPCとiPadを用意する。

例題と課題に取り上げたテーマは、身近な事例とし、統計グラフ全国コンクールの入賞作品

を参考とする。たとえば、例題では、第 63 回(2015 年)の「増やそうイクメン！日本のパパの働き方」、課題では、第 60 回(2012 年)の「あかりをかえて節電へ」のポスターを選ぶ。グループごとに、エネルギー消費に関する問題を見つけ、その問題を解決するために、PPDAC サイクルをまわす。データのフェーズでは、内閣府男女共同参画白書や資源エネルギー庁エネルギー白書などの公開データも用いる。分析のフェーズでは、適切な統計グラフを選んで、データをどのように視覚化するかも議論する。ポスター発表では、PPDAC の各フェーズが適切にまとめられサイクルが回ったかを、他のグループが確認する。

中学校や高等学校での 50 分*2 コマの授業に適用する場合、①イントロダクション、②PPDAC サイクルとは、③事例紹介を 1 コマ、④グループワーク、⑤発表・まとめを 1 コマとしてこの授業デザインを調整することが可能である。

4.4 高等教育(大学)における授業のデザイン(90 分*15 コマ)

東京情報大学で平成 29 年 6 月現在実施中の授業の取り組みを紹介する。授業概要は以下のとおりである。

対象：学部 3 年生(総合情報学部・総合情報学科，社会情報系データを扱う研究室の中で専攻の異なる 2 ゼミの合同)計 18 名

授業名：プロジェクト研究

授業時間：週 1 回 90 分 2 コマ(連続)

教員：2 名

取組内容：地元千葉市との共同研究プロジェクトを続ける中で、自治体からの課題を受け、それをデータ分析によって解決する、というサイクルを試行している。学生たちと同世代による課題解決の糸口発見という自治体の希望により、授業の中で下記のプランを立てて実行中である。課題の対象は、自治体が運営する大学近くの直売場と連結したレストランであり、知名度向上と売上の上昇、および、新たな層の開拓である。これらを元に、提供された売上データを自ら整理・集計・解析して知見を発見する作業を継続中である。

授業計画：

- (1) イントロダクション：共同研究による課題解決の意義を説明
- (2) PPDAC サイクルの解説と事例紹介
- (3) データの提示とその説明
- (4) 試食とその評価データの解析
- (5) 個人およびグループによるデータ解析(自治体提供の売上データ)
- (6) 上記(4)(5)の解析結果の発表
- (7) グループワークによる相互評価とディスカッション
- (8) 問題の所在確認と提示およびディスカッションと発表のサイクルを繰り返す
- (9) 相互評価のまとめ
- (10) 解析結果の最終提示
- (11) 発表資料の作成と発表
- (12) まとめと反省

経過と評価体制：

売上データについて変数整備後、視点を変えて探索的データ解析の実施を繰り返し、そのたびに相互評価を取り入れ、解析の方向性を議論している。そのプロセスを辿ることで、解析結果に対する自分とは異なる考え方に触れ、その理由を議論しながら別の解析手法や変数の入替等に考えが及び、知識の醸成を促す効果が生じる。同時に、データ解析のためのコンピュータ操作技術が向上し、データサイエンス実施過程では結果の視覚化や中途データのストック技

術, およびチーム間でのデータ共有の方法等, たくさんの関連する知識と技術の必要性に気付くことができる。学生に対する評価としては, 毎回の提出物の精査に加えて相互評価を重要視し, ディスカッションとともに評価表を点数化して分析し, その結果を再びPPDCAサイクルにのせる試みを実行している。本授業は継続中であるが, この教育結果については機会をみて発表する予定である。

5. 統計的探究力の評価

統計教育における生徒や学生の理解度を評価することは, 成績管理のみならず教育の改善においても大変重要である。また, 前述の授業デザインでも示したような, 実際にPPDACサイクルに沿って問題を解決する学びの成果として統計的探究力を評価するには, 一般的な知識理解とは異なる「思考」「判断」「批判」「提案」などを評価する新しい基準が必要となる。海外ではすでに統計的探究力の評価についての研究や実施の実績があり, イギリス義務教育修了試験(GCSE試験), アメリカAP試験, 国際バカロレア試験などの試験問題でも, 思考力等を問う出題が工夫されている。諸外国のペーパー試験では記述式問題も多く出題され, その有用性が広く認識されている。さらには, アクティブラーニング型の教育にも対応し, 統計的探究プロセスに沿った新たな評価基準も取り入れられている。

Davies と Marriott は統計教育の評価には以下7つの視点が含まれていなくてはならないと指摘している (Davies and Marriott, 2010)。

- (1) 統計的問題解決のプロセスの理解
- (2) 統計の技術的視点など授業内容の理解
- (3) 分析についての理解(パラメータの推測など)
- (4) 統計ソフトなどの活用力
- (5) 応用可能な力
- (6) 現実な視点で批判的にみる力
- (7) 結果を伝える力

これらの能力の評価においては, 単純な複数選択問題や短い答えを求める質問はむかない。

イギリス・アメリカ・ニュージーランドなど, 20年以上前から国を挙げて統計教育に注力してきた諸外国のカリキュラムでは, Davies と Marriott の指摘する7つの視点が盛り込まれた評価がすでに実施され, ペーパー試験とレポートの提出の両方を要件とする科目が多くある。国際バカロレア試験では, プロジェクトと呼ばれるレポート課題の提出が必須である (Bedding et al., 2007)。AP試験の後半問題では, 与えられた課題に対し, 解決に至るプロセス全てについて, その論拠を示しながら解答者に記述させる (Legacy, 2008)。海外での事例からもわかるように, 統計的探究プロセスを用いた統計教育の成果を評価するためには, プロセスとその実行結果をまとめたレポートを含めた演習課題とペーパー試験の両面から評価することが必須といえる。次節にイギリス, ニュージーランド, アメリカの事例をまとめる。

5.1 イギリス中等教育における統計教育の評価

5.1.1 科目「数学」のコースワークの評価

イギリスでは, 16歳の時(日本の高1に相当)にGCSE試験を全高校生が受験する。その後の学習や進路を決める大切な試験である。GCSE試験では, ペーパー試験に加えてコースワーク (Coursework) と呼ばれるレポート形式の課題が課す科目も多く, 数学も例外ではない。コースワークの採点は公表された明確な評価基準をもとに各学校で行われる。コースワークの配点

表 1. 数学のコースワークの評価シート(イギリス).

数学の内容	コア科目	応用科目
数学的な根拠がほとんどない。 文あるいは図での説明は多い。	0	0
概念や手法を表面的あるいは反復的に使っただけである。	1	0
基本的な概念や手法は使われている。 発展的試みがなされているが不適切、もしくは不成功である。	2	1
通常使われる関連概念と手法が活用でき、それらの適性を示せた。	3	2
コアシラバス内の広範な数学的内容を適切に使用し、さらにそれらを超えたいくつかの数学も使用した。数学的に話題を発展させた。数学的な説明は簡潔である。	4	3
応用科目のレベル以上の概念や方法を活用し、相当量の仕事をこなした。 通常の方法を超えた別の概念や手法も試している。数学的に獨創性がある。	4	4

表 2. 演習課題の評価シート「仮説と戦略」(イギリス).

仮説と戦略	点数
仮説も戦略も述べられていない	0
仮説が述べられているが、戦略はほとんどない。 正確な用語、表現が部分的には使われている	1
仮説が戦略とともに述べられている。作業の進め方や手法の採用について詳細に記述している。正確な用語、表現を常用できた。	2

は数学科目全体の 20%、ペーパー試験は 80% である。コースワークに関しては、各学校で 10 から 20 時間の問題解決型の授業が行われ、生徒は各自テーマを設定して問題を解決し、レポートを作成する。数学のコースワークは、1)問題を解決する力、2)数学を活用する力、3)自主的に学ぶ力、4)カリキュラムを横断して数学を活用する力、以上 4 つの能力を「計画と戦略」、「数学の内容」、「正確性」、「議論とプレゼンテーションの内容」、「個人の理解度」の 5 項目各 5~6 段階で採点する。数学を正しく活用することはもちろんのこと、問題解決の計画や戦略、議論や発表の明快さも評価の対象となる。例えば「数学の内容」の採点は表 1 の評価シートをもとに行われる。科目のレベル(コア科目と応用科目)により配点は異なる。また、個人の理解度については個別に口頭試問を行い判断する(University of Cambridge International Examinations, 2009)。

5.1.2 科目「統計」の演習課題の評価

イギリスでは、2014 年からは統計単独の科目も開講され、ペーパー試験と演習課題(Statistics in Practice)での評価がはじまった(AQA, 2012)。演習課題の配点は 25% で、採点は公表された評価基準をもとに各学校で行っている。演習課題は、表 2 から表 5 に示す「仮説と戦略」、「データの選択と収集」、「統計グラフの活用」、「統計指標の活用」の 4 項目で評価される。

また、イギリスでは継続してカリキュラムの見直しが行われている。2017 年度からは GCSE 試験の新科目として「問題解決(数学的推論、意思決定、数学的コミュニケーション)」という名前の新試験がはじまり、GCE-A 試験(17 歳)では「統計」が必修となった。さらには、GCSE 試

表 3. 演習課題の評価シート「データの選択と収集」(イギリス).

データの選択と収集	点数
適切なデータが収集できなかった	0
適切なデータ収集についてそのおおもとから丁寧に述べられたサンプリングについても若干の言及があった用語, 表現の使用は完全ではない	1-2
データを選択した理由とおおもとについて正確な記述があり, 適切なデータが収集された. サンプリング手法は適切であり, 再現可能な形で細かい記述がなされた. 適切なサンプルサイズを設定できた. 正確な用語, 表現を常に使うことができた.	3-4

表 4. 演習課題の評価シート「統計グラフの活用」(イギリス).

統計グラフの活用	点数
グラフや図を活用することができなかった	0
簡単ではあるが適切なグラフと図を活用できたグラフや図表を活用したり, グラフや図に基づいてコメントを作成したりすることがほとんど, もしくは少ししかできなかった	1-2
高い精度で適切な種類のグラフと図が作成できたグラフと図に基づく適切なコメントを作成した結論を引き出すために正しくグラフと図を活用できた	3-4
データをまとめるために正しいグラフや図を選択し, 正確に作成できた. より高度なグラフや図を作成した仮説の妥当性に関する意思決定のために推論を行い, 結論を導き出すためにグラフや図を最新の方法で活用した結論は継続しており, 最終的な要約へつながっている	5-7

表 5. 演習課題の評価シート「統計指標の活用」(イギリス).

統計指標の活用	点数
指標を求めている	0
簡単ではあるが適切な指標を計算できた計算結果を活用したり, 計算結果に基づいてコメントを作成したりすることがほとんど, もしくは少ししかできなかった	1-2
高い精度で適切な指標の計算が作成できた計算結果に基づく適切なコメントを作成した結論を引き出すために正しく計算結果を活用できた	3-4
データをまとめるために正しい指標を選択し, 正確に計算できた. より高度な指標も計算した. 仮説の妥当性に関する意思決定のために推論を行い, 結論を導き出すためにグラフや図を最新の方法で活用した. 結論は継続しており, 最終的な要約へつながっている.	5-7

験の学習内容と, 進路に合わせてより専門的に学習する GCE-A レベル試験の学習内容の差が見直され, その中間に位置する科目として理系に進学しない人向けの統計重視の数学科目が新設されるなど, 統計教育の充実が加速している (西村, 2016). 科学でも統計を活用した問題解決サイクルを実行する「科学的探究」の授業がカリキュラムに組み込まれている (イギリス教育省, 2015).

表 6. ニュージーランドの演習課題の評価基準(レベル1).

普通	高い	非常に高い
<p>生徒は統計的探究サイクルを使用して多変量データを探究した</p> <p>生徒は</p> <ul style="list-style-type: none"> 統計的探究サイクルを使用して、多変量データセットを調査した 対象となるデータを選定・収集できた 与えられた多変量データセットを使用して適切な比較の問い(仮説)を提示した 適切な規算表現を選択して実行した 5 数要約(最小値, 最大値, 中央値, 四分位数)などの統計指標を示した 分布の特徴(形状, 中心, 偏り, 重なり, 広がりなど)を比較し, 通常でない興味深い現象について言及した。 簡単な推論や裏付けの証拠をもとに結論をまとめた <p>例えば</p> <p>Problem(問題)</p> <p>母集団や標本, 変数を明確に特徴づける比較の問い(仮説)を設定した。「傾向がある」「一般的に」などの表現を用いて比較の問いを設定した。</p> <p>Analysis(分析)</p> <p>ドットプロット, 箱ひげ図を作成した。要約統計量を求めた。少なくとも 2 つ以上の分布の特徴を比較する議論ができた。</p> <p>Conclusion(結論)</p> <p>母集団や, 標本, 変数, 推論の質を含めて, 簡単な推論などの結果が述べられた。推論は箱ひげ図に示された中央値の位置と一致した。</p>	<p>生徒は理由で探究した</p> <p>生徒は</p> <ul style="list-style-type: none"> 統計的探究サイクルの構成要素を文脈や母集団に関連付ける根拠を示す答えを書き出し, 統計量, データの値, またはグラフや図の特徴などの根拠も参照した。 調査で得られたデータを探究した 与えられた多変量データを使って適切な比較の問い(仮説)を提示した 適切な規算表現を選択して実行した 5 数要約(最小値, 最大値, 中央値, 四分位数)などの統計指標を示した 分布の特徴(形状, 中心, 偏り, 重なり, 広がりなど)を比較し, 通常でない興味深い現象について言及した。 簡単な推論や裏付けの証拠をもとに結論をまとめた <p>例えば</p> <p>Problem(問題)</p> <p>母集団や標本, 変数を明確に特徴づける比較の問い(仮説)を設定した。「傾向がある」「一般的に」などの表現を用いて比較の問いを設定した。</p> <p>Analysis(分析)</p> <p>ドットプロット, 箱ひげ図を作成した。要約統計量を求めた。少なくとも 2 つ以上の分布の特徴を比較する議論ができた。議論はグラフや要約統計量などの根拠を参照して正当化された。</p> <p>Conclusion(結論)</p> <p>母集団や, 標本, 変数, 推論の質を含めて, 簡単な推論などの結果が述べられた。推論は箱ひげ図に示された中央値の位置を参照して正当化された。</p>	<p>生徒は, 統計的探究サイクルを使用して多変量データを統計的な洞察をもとに探究した</p> <p>生徒は</p> <ul style="list-style-type: none"> 統計的探究サイクルを通して統計の知識と調査対象に関する知識を統合した根拠を示す答えを書き出し, プロセスを振り返り, 結果に関する他の解釈についても考えたりすることができた。 調査で得られたデータを探究した 与えられた多変量データを使って適切な比較の問い(仮説)を提示した 適切な規算表現を選択して活用した 5 数要約(最小値, 最大値, 中央値, 四分位数)などの統計指標を適用した 分布の特徴(形状, 中心, 偏り, 重なり, 広がりなど)を比較し, 通常でない興味深い現象について言及した。 簡単な推論や裏付けの証拠をもとに結論をまとめた <p>例えば</p> <p>Problem(問題)</p> <p>母集団や標本, 変数を明確に特徴づける比較の問い(仮説)を設定した。「傾向がある」「一般的に」などの表現を用いて比較の問いを設定した。</p> <p>Analysis(分析)</p> <p>ドットプロット, 箱ひげ図を作成した。要約統計量を求めた。少なくとも 2 つ以上の分布の特徴を比較する議論ができた。議論はグラフや要約統計量などの根拠を参照して正当化された。議論は適切な対象の背景知識ともどもにまとめられた。</p> <p>Conclusion(結論)</p> <p>母集団や, 標本, 変数, 推論の質を含めて, 簡単な推論などの結果が述べられた。推論は箱ひげ図に示された中央値の位置を参照して正当化された。結論は調査対象に関する知識を統合してまとめられた。結論について他の解釈を検討する余地があるかもしれない。</p>

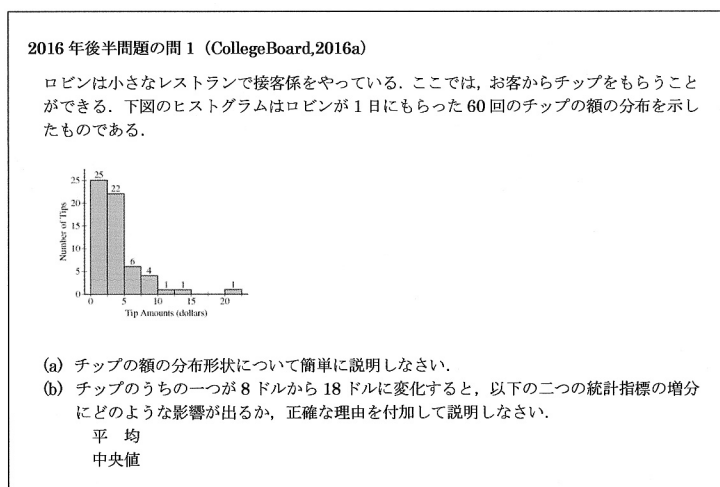


図 5. AP Statistics 2016 後半問題.

5.2 ニュージーランド中等教育における統計教育の評価

ニュージーランドの初等教育はほぼ 5 歳から始まり、初年度から「数学と統計：Mathematics and Statistics」という名称の科目の中で充実した統計教育が行われている。第 9 学年から中等教育がはじまり、第 11 学年(日本の高 1 に相当)で義務教育が修了する。第 11 学年から第 13 学年で全国共通学力試験 NCEA (National Certificate of Educational Achievement) のレベル 1～3 を受験し、これらの成績を大学入試の出願時に提出する。「数学と統計：Mathematics and Statistics」の試験は、ペーパー試験(外部評価)と演習課題(内部評価)で採点される。例えば 2017 年のレベル 1(第 11 学年)の演習課題では「センサス@スクールの 13 人分のデータを使って、11 年生全体に関する仮説をたて、92 人分のデータでその仮説を検証しなさい。少なくとも 2 つのグラフを作成し、グラフ、統計指標をもとに分布を比較して結果をまとめなさい。」という統計的探究力を問う課題が出題された (NCEA, 2017)。PPDAC サイクルを用いて、多変量データセットを調査・分析し、統計的な洞察を得る力が問われる。全体的な評価のみならず PPDAC の各フェーズについて、評価のポイントが示された評価基準も公開されている(表 6)。

5.3 アメリカ中等教育における統計教育の評価

アメリカでの統計教育評価事例として、AP Statistics と PLTW を取り上げる。AP は Advanced Placement の略であり、非営利団体 College Board が運営する、高校で学べる大学 1-2 年生用プログラムである。高校教員が行う授業と年度末一斉試験により評価が決まり、その評価内容は大学への進学を判定する要素の一部となっている。AP Statistics の認定試験は前半の択一解答方式と後半の自由記述方式で構成されている。自由記述式問題は、前半後半合計 6 問が出題され、前半の 5 問は計 65 分が解答に充てられている。6 問目は総合問題としてさまざまな知識を統合して論述する解答内容が要求され、25 分の解答時間が割り当てられている。これら後半問題の記述式の採点に関しては、出題の趣旨が明確に示され、一問ずつ詳細な採点ガイドラインが公表されている (AP Statistics 2016 Scoring Guidelines: College Board, 2016b)。一例として 2016 年後半問題の問 1(図 5)とその採点ガイドライン(図 6)を示す。

また、工学予備軍として PLTW(Project Lead The Way)という中学・高校を対象とした STEM(Science・Technology・Engineering・Mathematics)教育プログラムが実効性を発揮して

AP Statistics 2016 の採点ガイドライン (CollegeBoard, 2016b)

Part(a): ロビンのチップ額の分布は左に集中が偏り右に長く裾を引く形状である。最大額と二番目の額の差は大きく、最大額は外れ値に相当する。中央値は、2.5 ドル～5 ドルの間にある。チップの額は 0～2.5 ドルから 20～22.5 ドルと幅広い範囲の値をとるが、78%は 0～5 ドルに間にある。

Part(b): 8 ドルが 18 ドルになった場合、平均は、増分の 10 ドルを総計 60 で割った値ずつ増加する。(1/6 ドルずつ、または、17 セントずつ)。一方、中央値は、変化しない。中央値は、2.5 ドル～5 ドルの間にあるので、それより大きい値の変化に影響されない。

スコアリング (採点)

Part(a)も Part(b)も、正解(E)、部分的に正解(P)、不正解(I) で採点される。

Part(a)の採点根拠は以下のとおりである。

正解(E)は、解答内容が以下の 5 つの要素を含んでいることが必要である。

1. 分布の形状 (右に裾を引く偏った形状)
2. 外れ値の指摘、または、上位額間のギャップへの言及
3. 中心化傾向 (平均、中央値) の説明
4. ちらばり具合への言及について、特に最大から最小までの範囲の広さや 0～5 にほとんどが集中していること、あるいは分散について近い値で具体的な数値を提示していること、等が挙げられる。
5. チップ額の状況や前後関係を説明できる

部分的に正解(P)は、上記のうち、3～4 の要素を含んでいることが必要である。

Part(b)の採点根拠は以下のとおりである。

正解(E)は、解答内容が以下の 4 つの要素を含んでいることが必要である。

1. 平均は増加すること
2. 何故、平均が増加するのかについて説明できること
3. 中央値は変化しないこと
4. 何故、中央値が変化しなのかについて説明できること

部分的に正解(P)は、上記のうち、2～3 の要素を含んでいることが必要である。

不正解(I)は、上記のうち、少なくとも 1 つの要素を含んでいることが必要である。

まとめると、採点は以下のようなになる。

- | | |
|--------------|-------------------------------|
| 4 点: 正解 | Part(a)も Part(b)も正解 |
| 3 点: ほぼ正解 | どちらかの答えがやや正確性に欠ける |
| 2 点: ある程度正解 | 片方は正解だが一方が不正解、または、両方とも部分的には正解 |
| 1 点: 最小限度の正解 | 片方が部分的に正解だが、他方は不正解 |

図 6. AP Statistics 2016 の採点ガイドライン。

いる。PLTW は、コンピュータサイエンスの知識・技術の修得に力を入れており、その実践力を用いて工学的な問題解決力を磨く点に特徴がある。AP+PLTW のプログラムは、1) College (AP Courses), 2) Career (PLTW Courses) の 2 つのフィールドの組合せで構成され、1) は大学の単位認定につながるコースであり、2) はキャリアアップを認定するコースである。現在 AP Statistics は、1) の領域に入っている。1) で AP Statistics および AP Computer Science 関連科目、2) でコンピュータを使った問題解決関連科目 (例: Computer Integrated Manufacturing) を履修すればデータサイエンス教育の強力なツールになる。AP+PLTW の評価に関しては、AP のみ、PLTW のみ、に加えて、AP または PLTW の選択があり得る。AP または PLTW のコースでは、AP 試験の評価 3 以上および PLTW の期末評価における Proficient 以上を取ることで対応できる (College Board, 2017)。

6. プロジェクト型統計的問題解決学習の評価基準

これまで本研究グループでは Statistical Literacy, Statistical Reasoning, Statistical Thinking の 3 つの視点に対応した試験問題やワークシート課題の調査研究を通して、統計教育の評価について検討を行ってきた。海外では、プロジェクト型の授業も多く行われ、その評価においては、基準を明確にして記述式問題にも対応している。わが国でも現行の学習指導要領のもとで

表 7. 統計的問題解決評価ルーブリック SPART (高校版).

	1	2	3	4	5
P(問題)	問題を理解することができなかつた	提示された問題を解くことはできた	提示された問題を自らの問題として扱うことができた	自らテーマを設定し、そこでの課題を考へ、具体的な統計の問題として捉えることができた	自らテーマを設定し、そこでの課題を考へ、具体的な統計の問題として捉えることができた
P(計画)	必要な計画を理解できなかつた	プロジェクトの計画を立てることができた	計画の概略を作ることができた データの収集、クリーニングに関する計画を立てた	適切で根拠のある計画を立てた	適切で根拠のある計画を立てた
D(データ)	必要なデータを理解できなかつた	問題に合ったデータを収集できなかった	データが収集できない、不適切なデータの取り分けがなかった	問題解決につながるデータを集めることができた	問題解決につながるデータを集めることができた
	クリーニングの必要性を理解できなかつた	問題に合ったデータのクリーニング実行できなかった	データのクリーニングが不適切なものであった	問題解決につながる適切なクリーニングを実行した	問題解決につながる適切なクリーニングを実行した
A(分析)	統計的内容を活用できなかつた	平均、中央値、範囲、四分位数、標準偏差、分散など、公式通り計算することができた	平均、中央値、範囲、四分位数、標準偏差、分散など、中心、分散について計算し、データに合わせて使い分けられた	必要となる統計的観念と手法を考へ、広範囲にわたる統計の内容を適切に使った	広範囲な統計の内容が問題解決に活用され、それ以上の独自の活用が見られた
	統計グラフ	統計グラフのタイトル、ラベルに不備、ヒストグラムと棒グラフの違いが区別できていなかった	棒グラフ、ヒストグラムなど分布を示すグラフ、構成比率を示す円グラフなどを用いてデータを得意化するものが試みられた	分布を示すグラフ(ヒストグラム、バーストグラフ)を適切に活用し、正しい解釈を示した	グラフと指標を正しく活用し、条件の違いやグループでの比較、対象の分類にも試みを行った
C(まとめ)	結論に至らなかつた	不十分な箇所もあり、うまく適用されたとはいえない	分析を試みたが背後にある問題と合っていない箇所が見られた	2変量の関係について検討し適切に分析した	適切に分析し、予測、問題解決に活用した
	結論に至らなかつた	不十分な点があるが結論をまとめることができた	結論を適切にまとめることができた	問題の意味を広く認識し、結論をもとに考察や考察を明確に説明できた	適切に分析し、予測、問題解決に活用した
プレゼンテーション	計画性がなく発表や議論が継続できなかつた	計画性がなく発表や議論が継続できなかつた	発表の準備ができていなかった	発表の準備ができていなかった	発表の準備ができていなかった
総合的達成度	学んだ理論や手法をほとんど使っていない	学んだ理論や手法をほとんど使っていない	学んだ理論や手法をほとんど使っていない	学んだ理論や手法をほとんど使っていない	学んだ理論や手法をほとんど使っていない

グループ評価・個人評価

実データを分析してまとめ、発表するなどの統計的探究のプロセス(PPDAC サイクル)に沿ったプロジェクト型学習も行われてきているが、その評価は教員個々の裁量に任されていることが多い。そこで、本研究グループでは、プロジェクト型授業の評価のための統計的探究力評価ルーブリックを開発した。

プロジェクト型統計的探究力の評価ルーブリック(SPART: Statistics Project-based Assessment Rubric Table, 以下 SPART という)は、PPDAC サイクルを使って行ったプロジェクト型学習を PPDAC のフェーズごとに評価するルーブリックである。表 7 に一例として高校版の SPART を示す。①P (Problem: 問題), ②P (Plan: 計画), ③D (Data: データ), ④A (Analysis: 分析), ⑤C (Conclusion: まとめ), ⑥プレゼンテーション, ⑦総合達成度の 7 項目 5 段階で評価する。高校版の SPART では③D (Data: データ)を「収集」と「クリーニング」, ④A (Analysis: 分析)を「統計指標」「統計グラフ」「2 変数の関係」に細分化し, ①～⑥はグループ評価及び個人評価, ⑦総合達成度評価は個人評価とした。SPART を用いることにより, 評価者は学習者の到達度を公平に評価できる。また, 学習者自身も本ルーブリックを使って自己評価したり, グループ間で相互評価したりすることが大変有効である。SPART を用いることで学習の到達目標が明確になり, 自らの学習を振り返ることもできる。各フェーズの採点ガイドライン(高等学校版)を次節(表 8～表 17)に示す。中学校版, 大学版についても, 合格ラインやカリキュラムの内容を変更することにより同じ枠組みが利用可能である。

6.1 Problem: 問題フェーズの評価

問題フェーズでは, 解決すべき問題を捉えることができたかが評価のポイントである。統計で解決可能な問題を捉えることができる力が求められる。対象を明確にして, その対象の課題を捉え, 課題を評価するための指標と要因から仮説をたてる力も必要となる。採点基準は表 8 の通りである。

授業では, 各グループで対象を明確にしたうえでその対象の課題を明らかにするための議論が必要となる。評価者は, これらの議論をファシリテートしながらグループまたは個人の評価を行うことが必要である。

6.2 Plan: 計画フェーズ

計画フェーズでは, 問題フェーズでの議論をもとに具体的な計画を立てる力を評価する。必要となるデータを明確にして, アンケートや実験の詳細を決定し, 如何にデータを入手しクリーニングするかなどの具体的な計画を立てる力, 分析の見通しをたてる力も求められる。採

表 8. 問題フェーズの採点基準。

達成度	点数
問題を理解することができなかった	1
提示された問題を使うことはできた	2
提示された問題を自らの問題として使うことができた	3
自らテーマを設定し, そこでの課題を考え, 具体的に問題を捉えることができた	4
自らテーマを設定し, そこでの課題を考え, 具体的に統計の問題として捉えることができた 仮説を立てることができた	5

表 9. 計画フェーズの採点基準.

達成度	点数
必要な計画を理解できなかった	1
プロジェクトの計画を立てることができなかった	2
計画の概略を作ることができた データの仕様に関して言及できた データの収集, クリーニングに関する計画を立てた	3
適切で根拠のある計画を立てた	4
適切で根拠のある計画を立てた 分析の見通しを立てることができた	5

表 10. データの収集フェーズの採点基準.

達成度	点数
必要なデータを理解できなかった	1
問題に合ったデータを収集できなかった	2
データが収集できない, 不適切などの理由により問題の解決につながらない不十分な部分があった	3
問題解決につながるデータを集めることができた	4
問題解決につながるデータを集めることができた, データの不備や欠損, 次への課題についても発展的に考えることができた	5

表 11. データのクリーニングフェーズの採点基準.

達成度	点数
クリーニングの必要性を理解できなかった	1
問題に合ったデータのクリーニング実行できなかった	2
データのクリーニングが不適切などの理由により問題の解決につながらない不十分な部分があった	3
問題解決につながる適切なクリーニングを実行した	4
問題解決につながる適切なクリーニングを実行し, データの不備や欠損, 次への課題についても発展的に考えることができた	5

点基準は表 9 の通りである.

6.3 Data : データの収集フェーズの評価

問題解決のために必要なデータを取捨選択し, 収集できることが求められる. データの不備や欠損にも対処でき, データが収集できない際は計画フェーズに戻るなどの検討も必要とな

表 12. 分析フェーズ(統計指標)の採点基準.

達成度	点数
統計的内容を活用できなかった 中心の位置(平均値, 中央値, 最頻値), ばらつき(範囲, 四分位数, 標準偏差, 分散)などの記述的要約もできなかった	1
平均, 中央値, 範囲, 四分位数, 標準偏差, 分散など公式通り計算することはできたが, 表面的な理解にとどまり公式を反復的に使っただけだった	2
平均, 中央値, 範囲, 四分位数, 標準偏差, 分散など, 中心, 散らばりを示す指標を計算し, データに合わせて使い分けることができた	3
必要となる統計の概念と手法を考え, 広範囲にわたる統計の内容を適切に使った 統計的な説明は簡潔であった	4
広範囲な統計の内容が問題解決に活用され, それ以上のものも適用された 一般には適用されない手法をも試みた オリジナリティが見られた	5

表 13. 分析フェーズ(統計グラフ)の採点基準.

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

る. 採点基準は表 10 の通りである.

6.4 Data : データのクリーニングフェーズの評価

データのクリーニングフェーズでは, アンケートや実験を実施し, データセットを整備する力が必要となる. 1 件 1 レコード(1 行)になるようにデータを整え, データの不備も判断できることが望ましい. 採点基準は表 11 の通りである.

6.5 Analysis : 分析フェーズ(統計指標・統計グラフ)の評価

分析フェーズでは, 統計指標やグラフを適切に活用する力を評価する. 統計指標は, 正しく計算し, その意味を理解して使い分ける力も求められる. 中心の位置やばらつき(範囲)を示す指標の値を正しく解釈して分布の特徴や違いを説明できることが大切である. また, グラフと指標を合わせて活用する力も必要である. 採点基準は統計指標(表 12), 統計グラフ(表 13)に

表 14. 分析フェーズ(2 変量の関係)の採点基準.

達成度	点数
2 変量の関係について分析できなかった	1
不適切な箇所もあり, うまく適用されたとは限らない	2
分析を試みたが背後にある問題とかみ合っていない点が見られた	3
2 変量の関係について検討し適切に分析した 帰帰式, 予測, 標本からの推測に触れるなど若干のカリキュラムを超えた内容にも言及した	4
適切に分析し, 予測, 問題解決に活用した 標本に基づいて全体を推測した	5

表 15. まとめフェーズの採点基準.

達成度	点数
結論に至らなかった	1
不十分な点があるが結論をまとめることができた	2
結論を適切にまとめることができた	3
問題の意味を広く認識し, 結論をもとにさらにそれを広げようとした 結果を明確に説明できた	4
一般化することにより問題を解決した 仕事は完結し, 問題やデータの背景および自らの結論の意味, 影響についても説明した	5

各々に定義する.

6.6 Analysis : 分析フェーズ(2 変量の関係)の評価

高等学校の科目「数学」のデータの分析領域では, 2 変量の関係を正しく分析する力求められる. 本評価基準においても, 高校生に対しては問題解決につながる 2 変量の分析及びその解釈を評価する. 問題フェーズでたてた仮説の検証が行われなければならない. 採点基準は表 14 の通りである.

6.7 Conclusion : まとめフェーズの評価

まとめフェーズでは, 求めた指標, グラフ, 分析結果を的確に整理し, 単なるまとめにとどまらず, 問題解決のための提案や結果の一般化に言及することが求められる. 採点基準は表 15 の通りである.

6.8 プレゼンテーションの評価

プレゼンテーションに関しては, 統計指標やグラフを活用したわかりやすい説明が必要である. 作業したことを順に話すだけではなく, 論理的な構成が求められる. また, 一方的な報告

表 16. プレゼンテーションの採点基準.

達成度	点数
計画性がなく発表や議論が継続できない つながりのない短い議論や発表であった 結果がほとんどまとめられていないもしくはその試みすらない	1
十分に表現できなかった 論理的な展開が欠けていた 比較的重要ではない事柄に焦点が当てられ、重要な事柄に関する議論や発表がなかった 作業の順に発表するだけで、その順番を再度論理的に構築し直していない	2
努力を継続し、適切なプレゼンテーションだった 理にかなったまとめができたが、若干、明快でないところや強調すべき所を誤っている箇所があった	3
明快なまとめ、満足できるプレゼンテーションだった 論理的な順番で構成され、結果の一般化に対して十分に正当性がある	4
プレゼンテーションは、わかりやすく、図表、グラフが適切に活用された 結論とその一般化は、理にかなった文章で表現された まとめは、作業の主要部分で得た結果を反映した内容であった 異なる意見についても、まとめに盛り込まれていた	5

ではなく、議論の様子や内容を含めること、他者の意見を盛り込むことも重要である。採点基準は表 16 の通りである。

6.9 総合的達成度の評価

総合的達成度は、個人の知識理解の程度およびプロジェクトへの貢献度を評価する。採点基準は表 17 の通りである。

7. まとめ

本稿では、統計的探究プロセスを使った問題解決を実行する授業のデザインとその評価方法について論じた。統計教育の内容が質的に変化・拡大し、データをベースに問題解決へ導くプロセスが重要視される社会的な価値基準下では、生徒・学生の一人一人がデータを操作し、納得のいく分析結果を自ら手に入れる必要がある。それに不可欠なアクティブ型・プロジェクト型統計教育に対する評価は、従来の点数のみの評価体系では対応できない。本稿で展開した PPDAC サイクルに丁寧に対応した評価を行うことで、生徒・学生の理解度やデータサイエンス対応能力を測り、その結果が再び統計教育やデータサイエンス教育のカリキュラム構築に生かされるようなダイナミックな学びのサイクルが設計され始動することを目指したい。

本稿で提案した評価ルーブリック SPART を用いることにより、授業内で学生が自らを評価したり、グループ間で相互評価したりすることも可能となる。評価基準を学習者に明確に提示することは、学習者自身の目標の設定、理解度の確認、振り返りに効果的であり、教育の質向上、質保証に寄与するものと考えられる。今後の課題としては、PC 上で簡単に評価を入力でき、学習者間で共有できるアプリケーションの開発を進めたい。ルーブリックの内容をカスタマイズできる機能も盛り込み、ディスカッションを通して多様な場面、基準での評価を可能としていきたい。

表 17. 総合的達成度の採点基準.

達成度	点数
学んだ理論や手法をほとんど理解できていない 学んだことをほとんど使うことができない 多くの助けをかりた	1
理論や手法を十分には理解できなかった 基本的な方法を真似して使うことができるが、これを超えるためには助けが必要 先生やグループの仲間、お手本（書籍、ネット上の情報など）の助けを借りた ほとんど自分では作業をできなかったが、いくつか問題解決のために挑戦できた	2
当該のプロジェクトについて、そこで使われた手法について理解している プロジェクト全体を理解し学んだことを実行できる 先生や仲間から多少の助けをかりた 多少は自身の考えをもとに、形式的なやり方を実行できた	3
当該プロジェクトで使われた統計的手法を異なる状況への対応、応用、拡張はできない プロジェクトの全体を理解して実行でき、詳細な議論やまとめを作成できる 問題解決のサイクルをうまく実行できた 多少の助けをえながら、自ら実行し発展させた	4
当該プロジェクトで使われた統計的手法を異なる状況への対応、応用、拡張のアイデアを持ち、さらなる実行を試みた 当該のプロジェクトについて流暢に説明、文章化できた よく計画され問題解決のサイクルが（ときには複数回）うまく実行できた 自ら作業できた 参考書など資料を有効に活用できた	5

謝 辞

本研究は、統計数理研究所・共同研究利用・重点型研究(27-共研-4105, 27-共研-4108, 28-共研-2072, 28-共研 2073, 29-共研-4306, 29-共研-4308)および共同研究集会(27-共研-5016, 28-共研-5016)の助成を受けた。また、統計数理研究所の田村氏、中野氏から貴重なコメントを頂いた。ここに記して感謝の意を表したい。

参 考 文 献

- AQA (2012). GCSE Specification Statistics 4310, <http://www.aqa.org.uk>.
- Bedding, S., Forrest, J., Coad, M., Paula Tokman, W. and Fussey, B. (2007). *Mathematical Studies: Course Companion*—IB Diploma Programme, Oxford University Press, Oxford, UK.
- 中央教育審議会 (2012). 新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学へ～, 2012.8.28.
- College Board (2016a). AP Statistics 2016 Free-Response Questions, https://secure-media.collegeboard.org/digitalServices/pdf/ap/ap16_frq_statistics.pdf.
- College Board (2016b). AP Statistics 2016 Scoring Guidelines, https://secure-media.collegeboard.org/digitalServices/pdf/ap/ap16_statistics_sg.pdf.
- College Board (2017). AP + Project Lead the Way, <https://apstudent.collegeboard.org/exploreap/the-ap-experience/ap-project-lead-the-way>.
- Davies, N. and Marriott, J. (2010). Assessment and feedback in statistics, *Assessment Methods in Statistical Education: An International Perspective*, 1–19, Wiley, West Sussex.

- Education in New Zealand (2017). ニュージーランド教育省の Web サイト, <https://education.govt.nz/>.
- Gal, I. and Garfield, J.B. (1997). Curricular Goals and Assessment Challenges in Statistics Education, *The Assessment Challenge in Statistics Education*, 1–13, IOS Press, Amsterdam.
- イギリス教育省 (2015). National curriculum in England: Science programmes of study, <https://www.gov.uk>.
- 経済産業省 (2017a). 情報経済小委員会中間取りまとめ報告書について, http://www.meti.go.jp/committee/sankoushin/shojo/johokeizai/report_001.html.
- 経済産業省 (2017b). 産業構造審議会商務流通情報分科会情報経済小委員会中間取りまとめ～CPS によるデータ駆動型社会の到来を見据えた変革～.
- Legacy, M. (2008). AP[®] Statistics Teacher's Guide, <https://apcentral.collegeboard.org>.
- NCEA (2017). Internal Assessment Resource Mathematics and Statistics 1.10A v4 for Achievement Standard 91035, <https://ncea.tki.org.nz>.
- 西村圭一 (2016). イギリスの後期中等教育段階の数学に関する改革について, *日本数学教育学会誌数学教育*, **98**(7), 20–23.
- 政府統計の総合窓口 (2017). e-Stat, 総務省統計局, 東京, <https://www.e-stat.go.jp/SG1/estat/eStatTopPortal.do>.
- 竹村彰通, 椎名洋, 和泉志津恵, 松田安昌, 佐藤俊哉 (2015). 『gacco「統計学 II : 推測統計の方法」オフィシャルスタディノート』, 日本統計協会, 日本.
- Te Kete Ipurangi (TKI) (2017). ニュージーランドの教育用サイト, <https://www.tki.org.nz/>.
- University of Cambridge International Examinations (2009). Cambridge IGCSE Mathematics Syllabus code 0580 Cambridge IGCSE Mathematics (with coursework) Syllabus code 0581 For examination in June and November 2012, <http://www.cie.org.uk>.
- 山地弘起 (2014). アクティブラーニングとはなにか, *大学教育と情報*, 2014 年度 No.1 (通巻 146 号), 2–7.

Statistical Inquiry Process and Assessment

Hiromi Fukasawa¹, Naoko Sakurai² and Shizue Izumi³

¹Faculty of Healthcare, Tokyo Healthcare University

²Faculty of Informatics, Tokyo University of Information Sciences

³Faculty of Data Science, Shiga University

In the data-driven society using Cyber Physical System (CPS), decision-making processes using large data sets, including so-called ‘big data,’ are essential in all industrial and social endeavors. Due to the expansion of the Internet and the evolution of the mobile society, the Internet of things (IoT), and artificial intelligence (AI), the size of the available data continues to grow without limit. People who had traditionally relied on intuition or experience have shifted their decision-making to depend instead on the results of data analysis. Consequently, the importance of the ability to explore problems in a statistical manner, and to answer questions scientifically, has increased dramatically. In this paper, we surveyed and studied various methods of statistical education, with the goal of nurturing both scientific exploratory power and judgment ability. We also have summarized class design and performed evaluations including overseas cases, e.g., in the UK, USA, and NZ. As a class design that will foster empirical scientific abilities such as exploration, we propose a project-oriented practical lecture in which students can solve multiple kinds of problems by statistical inquiry and analysis. For evaluation, we propose “Statistics Project-based Assessment Rubric Tables: SPART” as a criterion for recognizing students’ abilities in each phase of the statistical inquiry process.