

# Python プログラミングで学ぶデータサイエンスのための高校数学基礎に関する オンデマンド教材の構築

## Construction of On-demand Teaching Materials on High School Mathematics Basics for Data Science Learning with Python Programming

章 志華\*      山本 敏幸\*  
Zhihua ZHANG      Toshiyuki YAMAMOTO

### 抄 録

データサイエンスの学習には、基礎数学および基礎情報の知識が不可欠である。データを活用するには、統計の知識が必要であり、機械学習アルゴリズムを理解するためには、線形代数や微分積分などの数学的な手法が求められる。本稿では、文系学生のデータサイエンス教育において、中高で学んだ基礎数学知識や 2022 年から始まった新しい教科「情報 I」の内容とどう関連づけるかを考え、Python プログラミングで基礎知識を復習できるオンデマンド型教材の開発について述べる。

### 1. 背景および目的

データ駆動型社会において、データサイエンス（以下、DS）は、科学およびと産業において変革をもたらす重要な分野として注目を集めており、現在あらゆる分野で DS 人材の育成が喫緊の課題となっている。政府の AI 戦略 2019 において、数理・DS・AI に関する知識や技能は、デジタル社会における基礎的な素養と位置付けられ、大学や高専において、より多くの学生を対象とした全学的な取り組みにより、体系的な教育課程を通じて基礎的な能力を修得することが示されている。また、文部科学省では、国際的に活躍できる科学技術人材の育成を目的として、文理を問わず、数理・DS・AI に関する知識や技術について体系的な教育を行うプログラム（リテラシーレベルおよび応用基礎レベル）の認定推奨を進めている。これを受けて、関西国際大学では、令和 3 年度より学長プロジェクトとして「データサイエンス教育プロジェクト」が発足され、DS 教育を全学横断的に実施するためのカリキュラムを作成した。そして、令和 4 年度より全学生を対象とした DS 教育プログラムをスタートし、DS 教育科目の授業実践に取り組んでいる。

データサイエンスの学習においては、基礎数学および基礎情報の知識が不可欠である。データを分析するためには、確率統計や線形代数、微分積分などの数学的手法を理解し、適用する能力が求められる。また、プログラミング言語やデータベース、アルゴリズムなどの情報技術も必須であり、これらを使いこなすためには論理的思考や問題解決能力も重要である。データサイエンスは数学と情報技

---

\* 関西国際大学社会学部 教育総合研究所学内研究員

術が融合した分野であり、その両方を身につけることが DS 力向上の鍵となる。これまでに文科省が推進する MDASH 認定のための教育プログラムに応じて、全国各大学が盛んに独自の数理・DS・AI 教育カリキュラムを展開しているが、必修科目の検討や DS 入門科目の設置過程において、学生の高校での数学・情報科目の学習歴等を十分に考慮したとは言い難い。数学や IT スキルが苦手な学生が多い本学のような私立文系大学では、基本的な数理・データサイエンス教育をする場合でも、受講生の数学・情報のレベルが問題となる。MDASH リテラシーを重視する文理融合型の大学では、データサイエンス教育を行う際、受講生の数学や IT のレベルが低いという問題がよく起こる。関西国際大学では、筆者が担当する 1 年前期の授業「データサイエンス」でも、データサイエンスの教育に関する課題があることを実感している。この授業では、社会学科や心理学科、英コミ学科の 1 年生を中心に、3 年生や 4 年生も含む履修者が対象となっている。

「データサイエンス教育に関するスキルセットおよび学修目標」-第 1 次報告（リテラシーレベル）（数理・DS 教育強化拠点コンソーシアム・カリキュラム分科会，2019）でも同様な懸念を持っていた<sup>1)</sup>。例えば、現行の学習指導要領では、高等学校の教科「情報」は「社会と情報」と「情報の科学」の分野からの選択式であるが、選択されている「社会と情報」の方が圧倒的に多い。DS に関連する内容は「情報の科学」に多く含まれているが、高校では多くの生徒がこの分野を学ばず、コンピュータとプログラミングを含む科学的理解やデータサイエンスの基本的な考え方が必要とされる時代において、重要な学習項目が不足していると言える。コンソーシアムのモデルカリキュラムは策定されたが、基礎数学等に対する習熟度・意欲が様々である大学生に対して、具体的にどのような数理・DS 教育を行えばよいのかについて、「議論は端緒についたばかりである」と指摘している<sup>2)</sup>。

一方で、データサイエンス・カリキュラム標準の策定に関して、DS・AI 教育先進国アメリカの The National Academies of Sciences, Engineering, and Medicine によるデータサイエンスの学部教育に関する報告書、「Data Science for Undergraduates Opportunities and Options」では、重要な 10 分野として、以下のように示されている。倫理に配慮した課題解決 (Ethical problem solving), データの記述・可視化 (Data description and visualization), データの取得・管理・加工 (Data management and curation), モデリングと評価 (Data modeling and assessment), 領域知識の考慮 (Domain specific considerations), コミュニケーションとチームワーク (Communication and teamwork), ワークフローと再現性 (Workflow and reproducibility) のほか、数学関連の分野には統計基礎 (Statistical foundations), 数学基礎 (Mathematical foundations), 計算基礎 (Computational foundations) の三つが強調されている<sup>3)</sup>。また、欧州の EDISON Data Science Framework より DS-BoK (Data Science Body of Knowledge) および DS Model Curriculum が提案され、統計, AI, データ工学, ソフトウェア工学, コンピュータサイエンス, ビジネス基礎等を含んでおり, ACM および DS スキルチェックリストの範囲をカバーしている<sup>4)</sup>。日本においても、情報処理学会が日本における情報処理の専門家集団としての立場から「データサイエンス・カリキュラム標準 (専門教育レベル)」が提案された。その分類「A: 基礎数学と数理統計学」では、「線形代数学基礎」「微分・積分」「数理統計学」などにおいて、高校数学との接続を考慮し、関連知識とスキルを示している<sup>5)</sup>。

文系大学生は、数学基礎力やプログラミングスキルが不足していることが多いこと、これらの能力は、DS 学習に必要な不可欠であること、また、高校情報基礎との接続も重要であることなどから、本研究は、文系大学生のリメディアル教育を行うことを目的とし、オンデマンド型学習で展開する数学基礎コンテンツを構築する。このコンテンツでは、DS 学習に欠かせない基礎数学および情報基礎の知識モジュールと対応する Python プログラミング活用型の学習コンテンツを提供する。学術的問いとして、本研究では、文系大学生の DS 教育に必要な数学基礎知識とプログラミング技術を融合した教材を開発し、オンデマンド型リメディアル教育モデルの実現可能性を検証する。即ち、教材コンテンツの開発では、数学基礎力および高校情報基礎との接続を考慮している。また、学習者が独自に進むことができるように、オンデマンド形式で構築している。

本稿の主題は、オンデマンド型コンテンツの構築である。このコンテンツでは、学生がデータサイエンス学習に必要な数学や情報の知識やスキルを身につけることを目指す。具体的には、中学校基礎数学と高等学校基礎数学の知識、および新しい必修教科「情報 I」の関連知識やスキルをモジュールとして選定した。それらに基づいた学習ユニットを用いて、オンデマンド型コンテンツを作成した。その過程と成果について報告する。また、オンデマンド型学習のための実装環境についても報告する。その中で、学内で使用するクラウドサービスとの連携、学習者自身によるプログラミング環境、LMS などのプラットフォームについても併せて紹介する。

## 2. 開発教材の概要

### 2.1 研究プロジェクトの位置づけ

これまで筆者らは知識情報処理の立場から対象領域向けの知識モデルの構築に関する研究に取り組んできている。近年新しい情報技術の発達とデータサイエンス教育の必要性の研究<sup>6)</sup>や、デジタル変革時代で求める人材像と文系大学データサイエンス教育に関する研究（第 84 回情報処理学会全国大会、2022）では、AI を代表する情報技術が発達し、DX 時代に求められる人材の特徴やデータサイエンス人材の育成に向けた教育プログラムを構築する方法について論じた<sup>7)</sup>。また、情報教育および教育へのデジタル技術利用に関する代表的な国際会議である WCCE 2022 (World Conference on Computers in Education) において、文系私立大学におけるデータサイエンス領域の教育カリキュラムの開発を論じた<sup>8)</sup>。そして、文系私大生のデータサイエンス教育におけるオンデマンド型授業の実践（第 85 回情報処理学会全国大会、2023）では、関西国際大学 DS 教育科目「データサイエンス」のオンデマンド型授業実践の概要および学習分析について報告した<sup>9)</sup>。

本研究プロジェクトは、これらの研究を基にした知識情報処理モデルリングの視点から、データサイエンスの学習に欠かせない中学校数学科、高等学校数学科、そして高校教科「情報 I」の基礎部を知識モジュールとして選定し、対応するオンデマンド型学習ユニットとして扱う。また関連学習ユニットの内容を Python プログラミング言語で実装し、学習者が基礎数学の知識の理解を深めるとともに、IT の基礎、問題解決、プログラミングスキル、そしてデータ活用などの要素を総合的に学べるオンデマンド型教材として構築する。

昨年度より開始された先行研究では、全学生対象のデータサイエンス教育プログラムを実現するために、図1に示す通り、データサイエンス教育カリキュラム構成モデルを作成した。このモデルでは、下から上へ順番に DS 基礎、DS 入門、DS 実践、DS 発展という4つの段階がある<sup>10)</sup>。

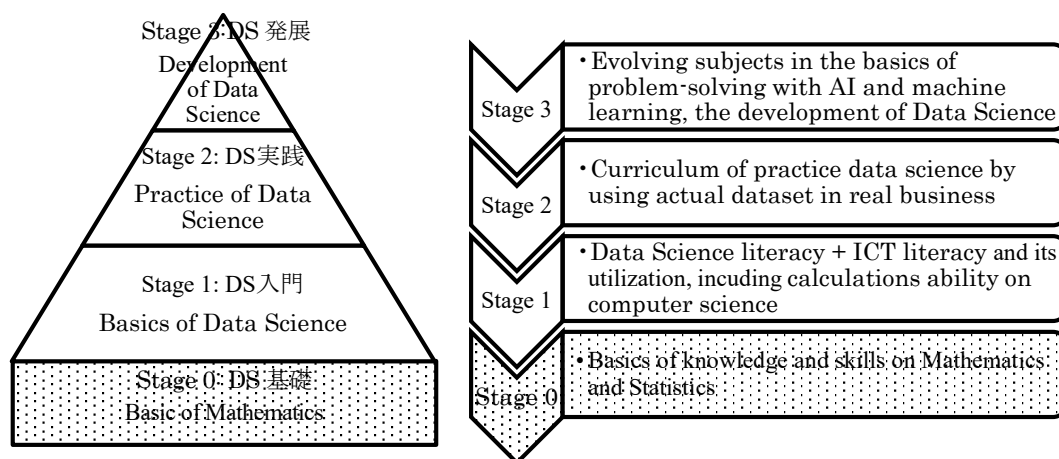


図1 KUISs データサイエンスの教育カリキュラムの構成モデル

ここでは、Stage 0: DS 基礎では、データサイエンスに興味を持ってもらうために、数学やIT技術の基本的な知識を身につけるカリキュラムが用意される；次に、Stage 1: DS 入門では、データサイエンスの基礎的な知識を学ぶカリキュラムがある；Stage 2: DS 実践では、実際のデータを使ってデータ分析の方法を実践的に身につけるカリキュラムが用意される；そしてStage 3: DS 発展では、AIや機械学習などの高度な技術を用いたデータ分析の手法を学ぶカリキュラムがある。本モデルはSRM (Stagewise Refinement Model)<sup>8)</sup>と呼ばれるもので、カリキュラムは階層的に分かれている。各階層では、簡単な内容から難しい内容へと徐々にレベルを上げられるように設計されている。

今回の研究プロジェクトは、Stage 0: DS 基礎としてデータサイエンスの入門を目的とし(図1の影部)、数学やIT技術の基本的な知識を扱っている。このプロジェクトでは、中学校や高等学校で習う数学や高校で必修となる情報Iの基礎的な内容を用いる。これらの知識は、文系私立大学でデータサイエンスを教える際に必要な前提知識であり、学生達にとっては、データサイエンスを深く理解したり、その分野で活躍したりするための大切なスキルとなる。

## 2.2 教材コンテンツの構成

データサイエンスは、データをもとにした意思決定や施策作りを助ける方法である。データを理解したり、見せたり、説明したりすることや、データを「予測」「発見」「分類」などで解析することは、すべて数学的な知識に基づいている。従って、基本的な数学(線形代数・微積分・確率統計)はデータサイエンスの土台の一つである。そういう意味では、DSの基礎理論にあたる「数学」を学ぶ意義は、言うまでもなく重要である。すでに述べたように、DSの学習は基礎数学のほかに、DSの学習に

欠かせないコンピュータ科学 (CS) , もっぱらプログラミング技法やアルゴリズムの学習も重要である。両方を学習することは、CS の理解を深めるだけでなく、数学的思考力や論理的思考力を養うことにもつながる。このように、基礎数学とプログラミングの知識は相補的な関係にあり、両方をセットで学ぶことはこの分野に対する知識が不足している文系学生に非常に有益であると言える。

このような発想を基にして今回の研究プロジェクトでは、開発するオンデマンドリメディア教材の内容は、現段階では主に下記三つの知識モジュールから構成されている。

(1) 中学校基礎数学部

扱ったテーマとして、中学校で習う「一次関数」「関数とグラフ」「方程式で図形を描く」などがあり、学習ユニットとして「直線を表すグラフ」「直線の式の求め方」を取り上げた。

(2) 高校基礎数学部

扱ったテーマとして、高校で習う「2 次関数」「図形と計量」(三角比など)「データの分析」(データサイエンスに必要なデータの整理・データの代表値・データのばらつき度合、傾向・関連、分散と標準偏差)、「確率と統計」(割合・順列・組合せ、乱数と値、共分散と相関係数、平均移動、回帰直線)、「線形代数」(ベクトルの演算、行列の演算)、「微分・積分」(曲線とグラフ、指数・対数関数、道具としての微積分・輪郭の抽出) などがあり、関連する学習ユニットとして取り上げた。

(3) 高校教科「情報Ⅰ」部

扱ったテーマとして、高校で習う「コンピュータとプログラミング」(コンピュータの仕組み・アルゴリズムとプログラミング・モデル化とシミュレーション) などがあり、関連する学習ユニットとして取り上げた。

情報分野から見た高大接続という視点から、三番目の知識モジュールを選んだ。その理由は、2022 年度から実施される新学習要領に従って、高校側では新しい教科「情報Ⅰ」と「情報Ⅱ」がスタートしたからである。「情報Ⅰ」はプログラミング技術を含む情報の科学的理解とデータサイエンスの基本的考え方が必修になる科目とし、さらにその発展的内容(ほぼデータサイエンスに傾けた内容)を扱う「情報Ⅱ」を選択科目として設置してある。その内容は表 1 に示すとおりである。

表 1 高校教科「情報Ⅰ」と「情報Ⅱ」の内容

「情報Ⅰ」(共通必修科目, 2 単位)	「情報Ⅱ」(選択科目, 2 単位)
1) 情報社会の問題解決	1) 情報社会の進展と情報技術
2) コミュニケーションと情報デザイン	2) コミュニケーションとコンテンツ
3) コンピュータとプログラミング	3) 情報とデータサイエンス
4) 情報通信ネットワークとデータの活用	4) 情報システムとプログラミング
	5) 情報と情報技術を活用した問題発見・解決の探究

「情報 I」の 3) のプログラミングでは、アルゴリズムの表現にあるフローチャートやアクティビティ図、典型的なソートとサーチ、問題の発見と問題解決に応じた計測・画像処理・シミュレーションなどが扱っており、WebAPI の利用や人工知能の活用といった内容まで及ぶ。また、4) のデータの扱いでは、分散・標準偏差・相関係数などの統計指標、散布図、仮設検定の考え方；クロス集計、仮設検定、単回帰分析およびこれらを通じたデータの可視化；そして量的データ・質的データ・データの整然性といった内容がある。この辺りは本研究で扱う知識モジュール（1）中学校基礎数学部の「D データの活用」および（2）高校基礎数学部「数学 I」の「（4）データ分析」と関連している。

これらの事情から、冒頭の研究背景でも言及した大学生の数理・DS・AI の習得に向けて、教科「情報 I」は大きな力になることは間違いなく期待される。またこのことから今後大学初年次教育も大きく変わる必要があると言えよう。しかしながら、新教科を教える教員には情報学全般について幅広い知識や技術が求められ、高校側の実情から考慮すると情報科教員配置にばらつきがあることも想定される。従って、今後大学初年次教育や情報リテラシー教育へ対応策から考えても、本研究プロジェクトで取り上げた内容は意義深いもので、情報分野高大接続展開へ寄与したことが期待される。

### 3. システムプラットフォーム

開発教材の学習内容は基礎数学およびプログラミング技法に関わることから、LMS（Learning Management System：学習管理システム）である WebClass のほかに、構築する際のコンテンツを扱うシステムプラットフォームは複数に渡る。Python プログラミング環境は Anaconda、そして、情報共有およびファイルの保存編集などは Microsoft 365 ユーザー（本学の学生の殆ど）にとって使用が便利なクラウドサービスである SharePoint（Web サイト作成・管理ツール）と Teams（チャット・ビデオ会議・協働作業ツール）と連携している。

#### 3.1 Python プログラミング環境

学習内容を理解するためには、基本的に学習者が Python プログラムを実行し、動作確認および理解する必要がある。そのため、学習者は Python の利用環境をインストールすることが第 1 歩となる。Python プログラミング環境は、Google 社が提供するクラウドベースの Jupyter Notebook 環境である Google Colab がある。インストール不要ですぐに Python プログラミング言語を使ったデータ分析や機械学習の学習を行うことができる便利性的がある。しかし、Google アカウントが必要となるほか、データファイルを扱う際にクラウド環境と利用者 PC 環境との間でのやり取りは、文系の学生にとって少し複雑で、難関である。一方、Python 標準パッケージ（公式版）は、Python プログラミングそのものの学習環境として素早く利用可能であるが、使用中に DS 学習に必要なライブラリ（特定の目的や機能を持つプログラム集）を追加することが学生にとって容易ではない。本研究では、コンテンツの構築に Anaconda 環境を使用した。Anaconda とは、データサイエンスや機械学習のための Python パッケージを集めたプラットフォームである。Python 本体のほかに、プログラムを作成する際に NumPy, SciPy, matplotlib などの DS 学習に必要なパッケージが含まれている。Jupyter Notebook のほか、Spyder

や VS Code など、よく使うエディターも同梱されている。Jupyter Notebook は、Web ブラウザ上で動作するインタラクティブなプログラミング環境で、Python プログラムを作成・実行でき、ソースコードとその実行結果を素早く確認できるツールである。コンテンツはマークダウン形式の説明文などを組み合わせて、PDF 形式や HTML 形式などでダウンロード可能、データ分析や機械学習の学びに優れた環境といえる。筆者らは実際の授業で Jupyter Notebook を用いて、演習問題を学生に解かせたところ、学生が素早く慣れることが確認されている。図 2 はその一例を示す。図中 (1) は標準データセット iris (花弁の長さや幅などからアヤメの種類を分類するデータセット) を取り上げて解説し、(2) は量的データの分布の様子を確認するために、Python でヒストグラムの描画方法を紹介している。

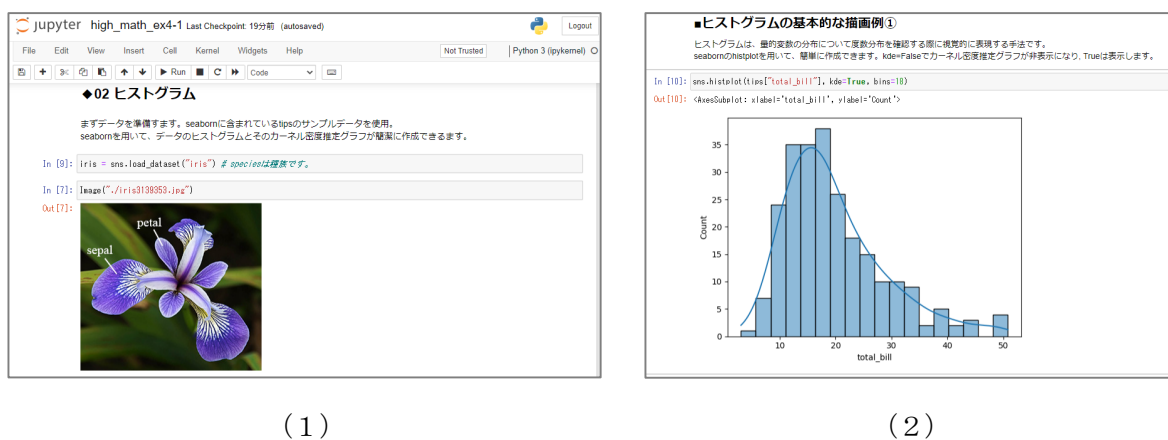


図 2 Jupyter Notebook 用いたヒストグラム描画の例

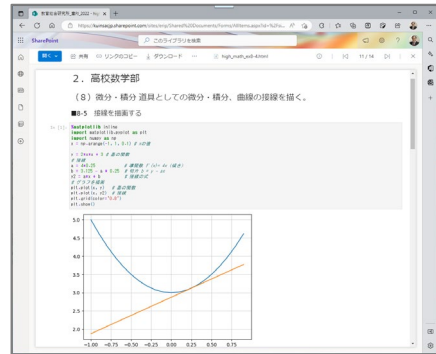
### 3.2 SharePoint サイト

SharePoint は、Microsoft 社のコラボレーションプラットフォームで、組織内のグループに文書共有やタスク管理、イントラネットなどを提供する。本研究プロジェクトでは、コンテンツ共有のために使用している。SharePoint サイトには、複数のページやリスト、ドキュメントライブラリ、カレンダーなどを含み、学習者 (Microsoft 365 のユーザー) が共有・編集できる。ブラウザで情報共有できるため、前述の Jupyter Notebook で作成した学習コンテンツを HTML 形式に変換すれば、簡単に学習問題バンク (本研究では、学習ユニットのデータベースを意味する) として活用できる。

さらに、SharePoint は、外部からのアクセスを許可すれば、VPN なしで学外のユーザー (学習者) にも利用できる。SharePoint は、Microsoft 365 のクラウド版に含まれている。また SharePoint サイト自体は Web サイトであり、Web ページの作成やファイルの保存・共同作業が可能である。このため、HTML 形式の学習コンテンツは SharePoint のニュースとして公開することにより、外部のユーザーである学習者にオンデマンド学習教材を提供するとともに、Teams の開発チームメンバーと開発状況の連携・共有や内容の共同編集・確認なども行える。図 3 に SharePoint で学習問題を共有する一例を示す。図中 (1) は本研究プロジェクトの SharePoint のトップページで、(2) は高校数学部の「微分・積分の応用例：曲線の接線」の学習ユニットの内容である。



(1)



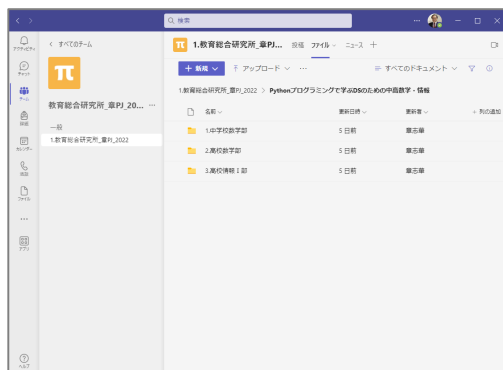
(2)

図 3 SharePoint での学習コンテンツの共有

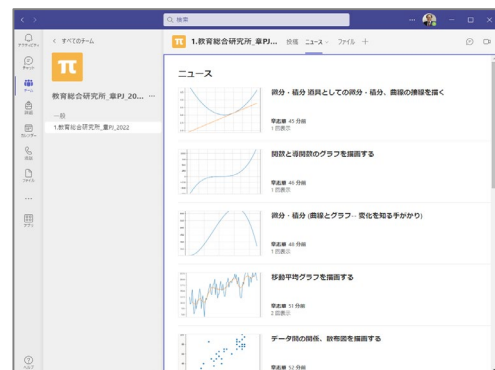
### 3.3 Microsoft Teams

Microsoft Teams は、Microsoft が提供するコラボレーションプラットフォームで、Microsoft 365 の一部である。チームワークに必要なツールがまとまっており、チャットや通話、ビデオ会議、ファイル共有、Office アプリとの連携などができる。また、統合されたアプリと連携することで、他のサービスともつながることができる。チームとは、同じ目的の業務やプロジェクトを行う人たちのグループで、Teams 内でコミュニケーションや情報共有をする単位である。各チームは、SharePoint サイトに接続されており、ファイルやデータを保存できる。また Teams のチャンネルごとにサイト、ファイル、フォルダーを共有するしくみがあり、開発者同士以外に、学習者グループごとの共有実現もできる。本研究プロジェクトでは、現在は開発者間での情報共有に Teams を利用しているが、今後はコンテンツを充実させてオンデマンド学習者とも共有したり、学習管理にも活用したいと考えている。

図 4 に Microsoft Teams における本研究プロジェクト様子を示す。図中 (1) は本研究プロジェクトの Teams の専用チャンネルで、(2) は SharePoint のニュースと連携する様子が確認できる。



(1)



(2)

図 4 Teams におけるコンテンツと SharePoint のユースページ連携



### 3.4 LMS (WebClass)

WebClass は日本の大学向けに開発された国産の LMS (学習管理システム) で、授業資料の配布やレポート提出などができるシステムである。科目やクラスごとに学習コンテンツのアレンジや小テスト、レポート提出などの機能が揃っている。本研究プロジェクトにおいて、今後オンデマンド履修者の学習履歴管理や学習効果の評価手段、LA (学習分析) の支援手段としても活用する予定である。

## 4. 学習コンテンツの紹介

本研究プロジェクトでは、オンデマンド型学習コンテンツの開発と構成設計を以下のように行っている。

1) 知識モジュールに基づいて、学習主題を選び、一問題を一ユニットとして学習資料を作成する。学習資料は、学習主題に関する説明文や図表などで構成される。学習資料は HTML 形式と PDF 形式の両方でウェブで閲覧できるようにし、SharePoint の学習問題バンクに保存する。また、HTML 形式の資料はニュースとして公開し、Teams で共有する。

2) 配布資料をもとに音声付きの解説動画を制作し、ストリーミングで配信・共有して視聴可能にする。

3) 学習内容に沿って Jupyter Notebook で Python コーディングをし、実行できるように何度も挑戦する。難しい場合は資料の解答コードを見て実践する。

4) 各学習ユニットで理解を深めるために、課外学習用の練習問題を課す。学習者が自分でチャレンジする。

5) オンラインチャットや Q&A 掲示板を用いて、学習者がコミュニケーションを取り、自分と他者の考え方を比較できるようにする。また、学習支援センターアワーと電子メールで、学習者の質問や相談に対応する。

表 2 は一つの主題に関わる学習ユニットの概要を示す。

表 2 オンデマンド型学習ユニットの概要

	内容	学習時間配当	説明
C-1	学習内容の文字解説	10 分	学習主題に関する HTML&PDF 資料
C-2	動画解説	10 分	内容に関する動画解説を視聴
C-3	演習 Python コード	30 分	Jupyter Notebook を用いてプログラムを実行
C-4	理解度確認	自由	学習内容に関連した理解度確認練習問題
C-5	学びの振り返り オンラインチャット	オンデマンド	学びの振り返りのために、オンラインチャットによる自他理解を共有する。

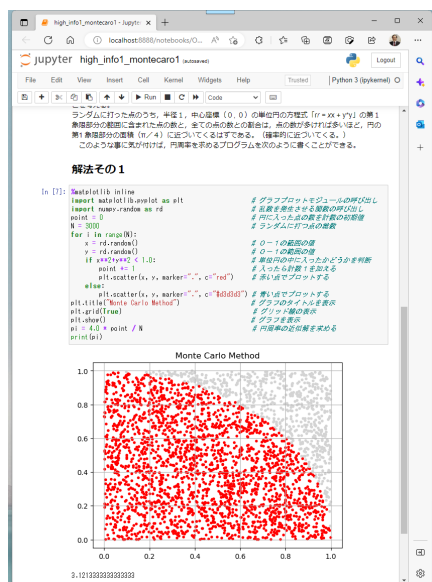
## 4.1 学習ユニットの紹介

ここでは、オンデマンド型学習ユニットの一例として、円周率を求める問題を紹介する。この問題では、モンテカルロ法という乱数を用いた数値計算手法を使って、円周率の近似値を求める問題である。この問題は、情報の高大接続に関わる知識モジュールである。

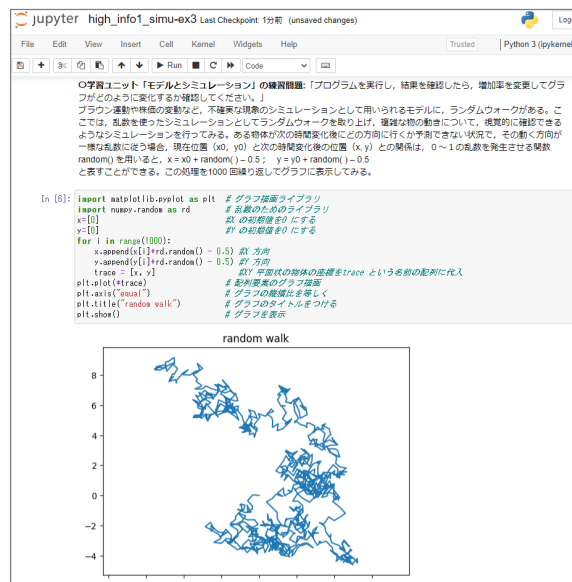
この問題では、「自然現象のモデル化とシミュレーション」というテーマについて考える。高校「情報 I」では、まず、「モデル化」と「シミュレーション」の意味や「確定的モデル」と「確率的モデル」の違いなど、基礎概念を学んだ上、学習目標として、1) 自然現象のモデルの構造を決定し、コンピュータを活用してモデルを動かしてシミュレーションを行う方法を理解する；2) 数式モデルを作成する活動を通じて、生徒に数式モデルを考えさせる授業ができるようになる；3) プログラミング言語などを活用してモデルを動かしてシミュレーションする方法について理解する；4) モデルの妥当性について、検討する活動を通して、生徒に考えさせる授業ができるようになるなどがある。この問題は、文部科学省高等学校情報科「情報 I」教員研修用教材<sup>1)2)</sup>を参考にして作成した。

一方、モデル化とシミュレーションの学習ユニットは、数学科の第 2 款第 4「数学 A」2 (2)「場合の数と確率」と密接に関係している。さらに、アルゴリズムとプログラミングは、他の教科で必要な論理的思考力や表現力とも深く結びついている。データサイエンス専攻において、コンピュータプログラムにおける乱数を発生させる関数の理解にも役に立つ。本稿では、高校基礎数学と「情報 I」の関連するテーマを典型例として紹介する。

図 5 は、Python 開発環境 Jupyter Notebook で作成した学習ユニットである。(1) は説明資料で、コードと実行結果が見られる。(2) は関連する内容の練習問題で、コードと実行結果が確認できる。



(1)



(2)

図 5 Jupyter Notebook を活用した学習ユニットの説明資料

Jupyter Notebook に基づく学習ユニットの説明資料は、次のようになっている。

### 3. 高校情報 I 部

問題分類：高等学校「情報 I」－「コンピュータとプログラミングとデータ活用」

関連内容：◆コンピュータの仕組み； ◆アルゴリズムとプログラミング； ◆データ活用；  
◆モデルとシミュレーション

#### ○ユニットの目標

(1) 社会現象や自然現象に関するモデル化の方法とシミュレーションについて理解し、これをプログラミングするための条件分岐や関数化などの技能を身に付けている。

(2) シミュレーションを実際の現象と結び付けて考え、これをプログラミングし、出力が正しいかどうかを判断し、プログラムを修正する力を養う。

(3) 自然現象や社会現象などの事象とモデルとの関係を適切に結び付け、シミュレーションを用いたプログラミングで記述し、それを更に発展させようとする態度を養う。

#### ○主な学習活動

- 社会現象や自然現象におけるモデル化とそれをルールに基づいて実行する「モンテカルロ法で円周率を計算する」を通してシミュレーションについて理解する。
- ルールをどのようにプログラミングするか。関数化や条件分岐を用いて、試行錯誤しながら、プログラミングし、目的とするシミュレーションをプログラミングする。
- シミュレーションを行った結果を評価し、適正であるかどうか、より分かり易い表現や可視化の方法はないか、また他へのどのような応用が考えられるかなどについて考察する。

#### ○学習問題説明

【問題名】モンテカルロ法で円周率を計算する問題

この問題では、現実の現象を数学的なモデルにしてシミュレーションする方法を学ぶ。今回は、偶然や不規則さが関係する現象を扱うため、確率的なモデルというものを使う。このようなモデルでは、ランダムな数字（乱数）を使ってシミュレーションする。

モンテカルロ法の特徴は対象のモデルに乱数を大量に生成して入力し、近似解を得ようとする手法である。解法 1 は、平面の第 1 象限部分の X 方向に 1、Y 方向に 1 の長さをもつ正方形内に、ランダムに点を打つことを考える。ランダムに打った点のうち、半径 1、中心座標 (0, 0) の単位円の方程式「 $r^2 = x*x + y*y$ 」の第 1 象限部分の範囲に含まれた点の数と、全ての点の数との割合は、点の数が多ければ多いほど、円の第 1 象限部分の面積 ( $\pi/4$ ) に近づいてくる（確率的に近づいてくる）。コードは解法その 1 とその 2 のように書くことができる。

プログラムを入力・実行して、試行回数が結果に与える影響について考察してみてください。コンピュータの乱数の発生がどの程度の試行回数で妥当と判断できるでしょうか。

解法その 1（第 1 象限）の Python コードは次の通りである。実行結果は図 6 に示すとおりである。

```

%matplotlib inline
import matplotlib.pyplot as plt          # グラフプロットモジュールの呼び出し
import numpy.random as rd              # 乱数を発生させる関数の呼び出し
point = 0                               # 円に入った点の数を計数の初期値
N = 3000                                # ランダムに打つ点の総数

for i in range(N):
    x = rd.random()                    # 0 - 1 の範囲の値
    y = rd.random()                    # 0 - 1 の範囲の値
    if x**2+y**2 < 1.0:                # 単位円の中に入ったかどうかを判断
        point += 1                     # 入ったら計数1を加える
        plt.scatter(x, y, marker=".", c="r") # 赤い点でプロットする
    else:
        plt.scatter(x, y, marker=".", c="#d3d3d3") # 青い点でプロットする
plt.title("Monte Carlo Method")        # グラフのタイトルを表示
plt.grid(True)                         # グリッド線の表示
plt.show()                              # グラフを表示
print(4.0 * point / N)                 # 円周率の近似解を求める

```

注) #から右は、コードのコメントの表記

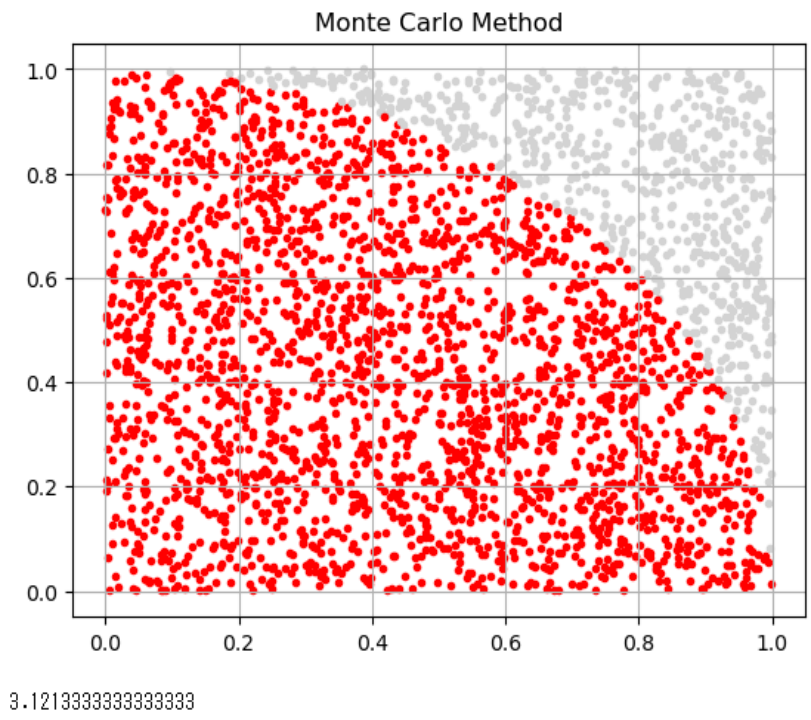


図 6 実行結果

また、解法その2として、全象限に点を打つことで求めることができる。さらに、この問題の理解

を深めるために、関連する内容の練習問題も用意してある。例えば、生命体の増加シミュレーション、ブラウン運動や株価の変動など、不確実な現象のシミュレーションとして用いられるモデルがある。練習問題として、ランダムウォークのシミュレーション問題を選択した。図5の(2)は、求めるプログラムコードおよび実行結果を表したグラフである。この学習ユニットでは、乱数を使ったシミュレーションとしてランダムウォークを取り上げ、複雑な物の動きについて、視覚的に確認できるようなシミュレーションについて学ぶことができる<sup>13)</sup>。

## 5. 結び

データサイエンスは、科学と産業の両方に革命をもたらす分野として台頭し、データ駆動型社会におけるデータサイエンス教育の重要性はますます高まっている。人材の育成は喫緊の課題であるが、なかなか難しい問題で、私立文系大学ではさらに難しい。その原因は文系学生の数学力と情報処理能力が欠けていることである。データ分析やデータ活用をキャリアにしない場合でも、論理的思考には数学、実世界に即した問題解決には情報の強さが必要で、データサイエンス教育の実施には、数学の基礎と情報の基礎の両方が不可欠である。文系学生がデータサイエンスを学習する際に直面するこれらの問題を少しでも克服する手助けになる願望から、本研究は、文系学生がデータサイエンスを学ぶ際に必要な中高の基礎数学知識と新教科「情報Ⅰ」の内容を Python プログラミングで復習できるオンデマンド型学習教材を開発した訳である。

本稿では、オンデマンド型コンテンツの構築において、中学校基礎数学と高等学校基礎数学の知識、および新しい必修教科「情報Ⅰ」の関連知識やスキルを知識モジュールとして選定し、それらに基づいた学習ユニットのコンテンツ作成について述べた。また、開発環境に関連して、システムプラットフォームの構成および学習ユニットの概略についても報告した。学習ユニットに関連する問題集は今後さらに充実する必要がある。また、教材コンテンツを授業支援としてどう使うか、学習者にどう提供するか、そして WebClass でコースウェアを如何に構築するかなどは、今後の課題である。

本研究の取組に対して、教員と学習者のそれぞれの視点から見ると、様々なメリットがあると考えられる。教員の視点からは、データサイエンス教育や人文・社会科学系学生のための Python プログラミング教育における学内資源が利用可能になる。また教員が学生に数学基礎力や IT リテラシーの補習に本プロジェクトで作成した教材(問題バンク)を生かすことができる。また、受講生は、データサイエンスを履修する前に数学への苦手意識を克服したり、IT スキルを向上させたりできる。これによって、データサイエンスの学習への準備が整う。また、高校の時に学んだ基礎数学を復習し、高校で必要な情報技術の基礎を身につけ、BYOD を活用した学習で IT 文房具として使いこなせるようになる。

本研究プロジェクトは、データサイエンス教育カリキュラムモデルの開発という、開拓中の分野における先駆的な研究であり、全国の類似する小規模な文系私立大学でのデータサイエンス教育カリキュラムの実装に参考となることが期待できる。また、開発されたモデルやオンデマンド教材を用いて、数学基礎や IT 基礎のリメディアル教育を行うことで、データサイエンスのマインドセットを養うこ

ととなる。さらに、学生の実態に適したコンテンツの提供やデータサイエンス教育質保証への貢献も可能となる。

## 謝 辞

本研究は、「2022年度 関西国際大学教育総合研究所」研究プロジェクト助成を受けたものである。

## 6. 参考文献

- 1) 数理・DS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および学修目標」-第1次報告(リテラシーレベル)」, (2019年11月)
- 2) 松尾由美, 玉田和恵, 「文系・私立大学におけるデータサイエンス教育の課題」, 『江戸川大学紀要』第31号(2021年3月) pp. 249-255
- 3) ACM Data Science Task Force, Task Force Final Report, 「Computing Competencies for Undergraduate Data Science Curricula」, <http://dstf.acm.org/>, (January 2021)
- 4) EDISON Data Science Framework: Part 2. Data Science Body of Knowledge (DS-BoK) Release 2, (July 2017)
- 5) 情報処理学会, データサイエンス教育委員会, 「データサイエンス・カリキュラム標準(専門教育レベル)」(2021年4月)
- 6) 章志華, 「新しい情報技術の発達とデータサイエンス教育の必要性」, 『山陽女子短期大学紀要』第41号(2020年3月) pp. 1-20
- 7) 章志華, 「デジタル変革時代で求める人材像と文系大学データサイエンス教育」, 第84回情報処理学会全国大会, (2022年3月)
- 8) Zhihua ZHANG, T. Yamamoto and K. Nakajima, 「Development of Education Curriculum in the Data Science Area for a Liberal Arts University」, Proceedings of IFIP WCCE 2022, (August 2022)
- 9) 章志華, 「文系私大生のデータサイエンス教育におけるオンデマンド型授業の実践」第85回情報処理学会全国大会, (2023年3月)
- 10) 章志華, 「DX人材の育成と文系大学におけるデータサイエンス教育プログラムに関する研究」, 『関西国際大学教育総合研究叢書』第15号(2022年3月) pp. 163-175
- 11) National Academies of Sciences, Engineering, and Medicine 2018, 「Data Science for Undergraduates: Opportunities and Options」, Washington, DC: The National Academies Press, (2018)
- 12) 文部科学省, 高等学校情報科「情報I」教員研修用教材(本編), 第3章コンピュータとプログラミング, (2021年3月)
- 13) 文部科学省, 高等学校情報「情報」実践事例集, 情報I(3)「プログラミングをはじめよう」, (2021年3月)
- 14) 谷尻かおり, 文系プログラマーのためのPythonで学び直す高校数学, 日経BP社, (2019年)
- 15) 谷尻かおり, 文系でも必ずわかる中学数学×Python, 日経BP社, (2020年)

- 16) 萩谷昌己, 図説情報 I (高等学校情報科用), 実教出版, (2022 年)
- 17) 実教出版編集部, ベストフィット情報 I (中学の内容から大学入試共通テストまで), 実教出版, (2022 年)
- 18) 佐藤諭史, 教科書ガイド 実教出版 情報 I 「最新」 「Python」 「JavaScript」, 文研出版, (2022 年)
- 19) Curriculum Guidelines for Undergraduate Programs in Statistical Science (2014),  
[https://www.amstat.org/asa/files/pdfs/EDU\\_guidelines2014\\_11\\_15.pdf](https://www.amstat.org/asa/files/pdfs/EDU_guidelines2014_11_15.pdf).
- 20) National Academies of Sciences, Engineering, and Medicine 2018. Data Science for Undergraduates: Opportunities and Options. Washington, DC: The National Academies Press, (2018).

## Abstract

Basic mathematics and information knowledge are essential for learning data science. To utilize data, knowledge of statistics is necessary, and mathematical techniques such as linear algebra and calculus are required to understand machine learning algorithms. In this paper, we discuss the development of an on-demand learning material that allows students of humanities to review basic knowledge with Python programming, considering how to relate the basic mathematics knowledge learned in junior and senior high school and the new subject “Information I” that started in 2022 to data science education for students of humanities. Such contents are showcased in the paper.