

高等学校と大学のデータサイエンス教育の接続について

順天堂大学 数理・データ科学教育研究センター 大橋 真也
Shinya OHASHI, Juntendo University

1 はじめに

高等教育における数理・データサイエンス・AI教育に関しては、国の「AI戦略」に関するAI人材育成のために計画された教育の一つである。これに関連して、初等中等教育における小学校からのプログラミング教育や中学校の数学におけるデータ分析や技術のプログラミング、高等学校における数学や情報科目におけるプログラミングやデータサイエンス教育に関しても、今回の学習指導要領では、その連続性を意識して策定されている。本稿では、高等学校における数学や情報科目と大学における全学的な数理・データサイエンス・AI教育のリテラシーレベルとの連続性について、考察していく。

2 高等学校におけるデータサイエンス教育

本年度、高等学校で新しい新学習指導要領による新教育課程が開始された。科目編成が異なることや観点別評価の導入など、これまでに比べ大きな変化が期待される。小中学校においては、昨年までで新教育課程が開始されており、実際には移行措置により数年前から開始され、学年進行で進んでいる。

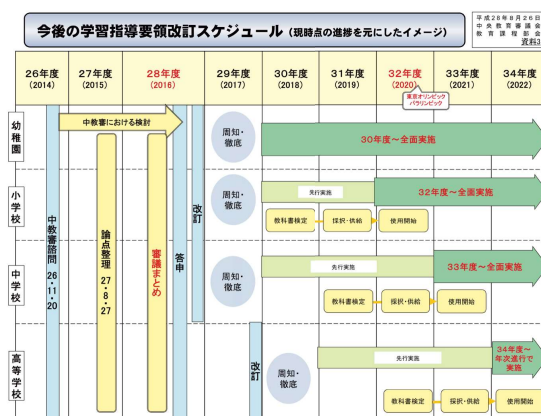


図1：学習指導要領のスケジュール (文部科学省説明資料より)

今回の学習指導要領においては、小学校における英語教育やプログラミング教育なども話題になっているが、とりわけ高等学校における「情報」科目に関してはかなりの話題になっている。

- ICTの遅れた日本において、先進的なICTを推進するような科目内容

- 一部の自治体の計画的な教員採用の不調による情報科の教員不足の問題
- 2025年度（令和7年度）大学入学共通テストにおける「情報Ⅰ」の新規導入
- 情報科目の大学入試への採用における受験産業の動き
- 選択科目「情報Ⅱ」について、東京都で5割、首都圏で4割以上の高等学校が採択

これら以外にもさまざまな点でインターネットや新聞、テレビなどのマスメディアで情報科に関しての話題が提供されており、高等学校学習指導要領の内容も今まで以上に社会に注目されている状況である。高等学校の情報科目においては、共通必修科目「情報Ⅰ」や共通選択科目「情報Ⅱ」において、プログラミングについてと同様にデータサイエンス教育も注目されている。

2.1 高等学校数学科におけるデータサイエンスの扱い

高等学校の新教育課程におけるデータサイエンス教育は情報科だけではない。前学習指導要領から、必修科目「数学Ⅰ」におけるデータの分析の内容の約半数が、中学校2年の学習内容に移動し、「数学Ⅰ」においては、実際のデータの分析について学べことを意識し、質的データと量的データ、欠損値・外れ値の扱い、仮説検定の考え方などが新たに導入されている。この中でも仮説検定の考え方においては、シミュレーションなどを通じて仮説検定の考え方を学ぶなどの教育実践も多く紹介されている。また前課程と同様に「数学A」では、確率が大きく取り扱われている。

また、「数学C」が新設され、そこでは表現方法としての行列が扱われるようになり、隣接行列や混同行列などが扱われる可能性がある。「数学B」における確率と統計分布においては、前課程と同様の確率分布や信頼区間について扱われるだけでなく、仮説検定やベイズ推定などについても扱われている。

これらの変化に関しては、現学習指導要領においては基本的な考え方である「カリキュラム・マネジメント」を考慮し、後述する情報科目との科目横断的な連携を意識して編成されている。

2.2 高等学校情報科におけるデータサイエンスの扱い

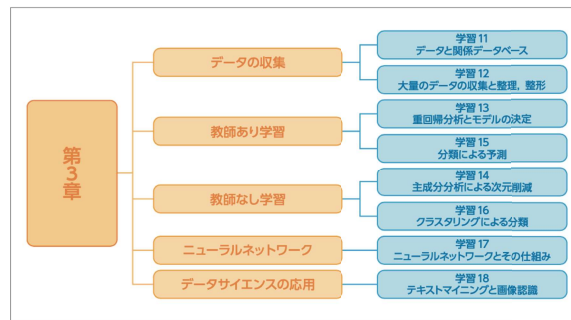
新学習指導要領の「情報Ⅰ」におけるデータ分析の扱いに関しては、昨年も報告したが、プログラミングとの連携も含め単回帰分析まで扱うこととなっている。数学科との違いは、数学科は統計学的なアプローチで進めていくのに対して、情報科はデータサイエンス的なアプローチで進めていくことであろう。昨年度報告した、教科書と内容の対応表を再掲する。

教科書	実データ 最新データ	尺度	異常値	データ クレンジング	類似相関	交絡因子	回帰直線	テキスト マイニング	取扱検定	ソフト
東書 新編 情報 I	○	○						○		なし
東書 情報 I Step Forward!	○	○			○	○	○	○	○	表
実教 高校情報 I Python	○	○	○				○	○	○	表,Py
実教 高校情報 I JavaScript	○	○	○				○	○	○	表,JS
実教 最新情報 I	○	○			○			○		表
実教 図説情報 I	○	○						○	○	表
数研 高等学校情報 I	○	○	○			○		○		表
数研 情報 I Next	○	○	○					○		表
日文 情報 I	○	○	○					○	○	表,Py
日文 情報 I 図解と実習	○	○		○				○		表,Web
第一 高等学校情報 I	○	○	○	○				○		表,Web

図 2：教科書とデータサイエンス項目との対応表

「情報 I」においては、大学入学共通テストサンプル問題や試作問題にも登場した散布図行列などを教科書では扱っておらず、単回帰分析においても散布図上に単回帰直線を引いたのみで予測などに用いることを記述している教科書がほとんどなく、学習指導要領を十分には反映されていないものが多くみられている。

「情報 II」においては、機械学習的な内容を学習し、それとプログラミングを合わせ、実際に機械学習の手法を使って探究を行う試みが想定されている。



図表2 データ分析と学習の分類

図 3：「情報 II」のデータサイエンス (文部科学省教員研修用教材より)

「情報 II」に関しては、前述したように東京都では5割以上、千葉、埼玉、神奈川県においては4割以上の高等学校が採択をしており、選択科目としてはかなり多い採択がある。ただしほとんど採択されていない県もあり、学校間格差は広がる傾向にあることが懸念されている。また、一部の大学において一般入学試験に「情報 II」を出題予定であるとの発表があり、今後採択する高等学校は増えていくとみられている。

3 大学における数理・データサイエンス・AI教育

高等教育におけるデータサイエンス教育は、国が進める「AI戦略」の一つとして、初等教育、高等教育、社会人教育との連続性を意識して、エキスパート人材の育成だけでなく、全学的な数理・データサイエンス・AI教育として、進められている。

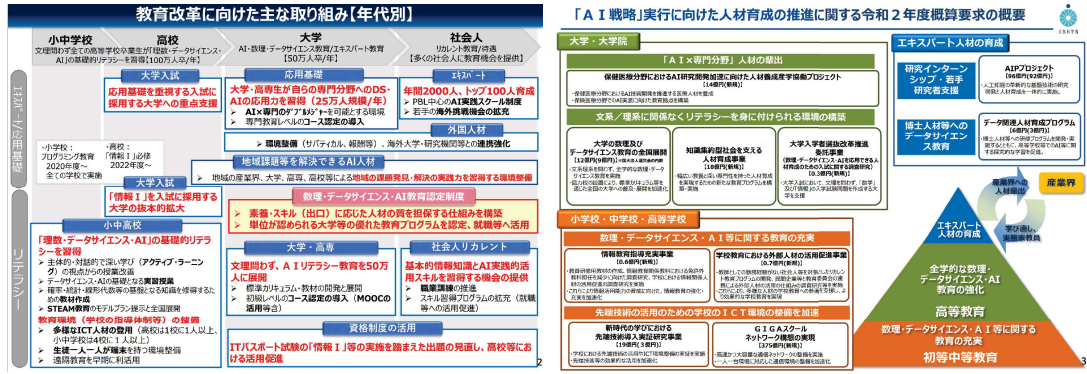


図 4：国の「AI 戦略」におけるデータサイエンス教育（報道資料より）

このような「IT 戦略」の一環として、プログラミング教育と同様に初等教育と一連の連続性も意識して作られた教育である。これは、初等中等教育においては、「新学習指導要領の下で、すべての高等学校卒業生（約 100 万人／年）に『理数・データサイエンス・AI 教育』に関する基礎的なリテラシーを習得させるとともに、問題発見・解決学習を通じた創造性を涵養する。」という大目標を掲げている。高等教育における大目標は、「入口（入試）から出口（就職）までの数理・データサイエンス・AI 教育の促進につながるシステム構築により、文理を問わず、すべての大学・高専生（約 50 万人／年）が AI リテラシー教育を習得するとともに、一定規模の大学・高専生（約 25 万人／年）が自らの専門分野への数理・データサイエンス・AI の応用基礎力を習得」などとしている。

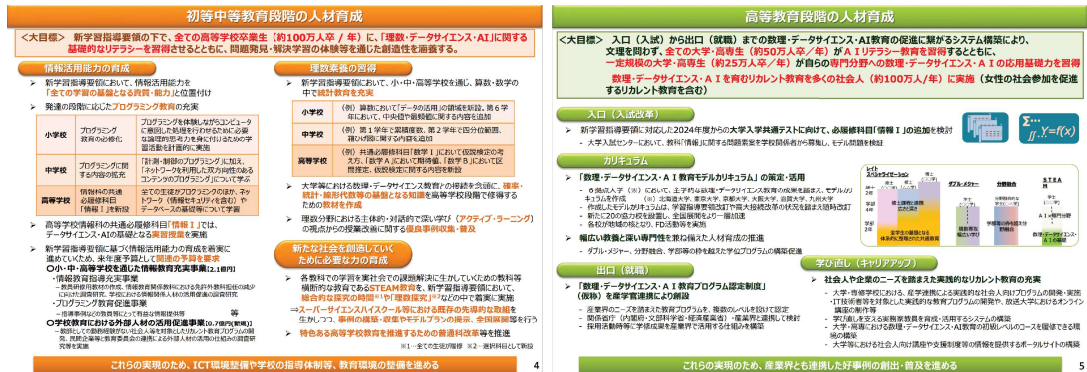


図 5：初等中等教育と高等教育におけるデータサイエンス教育（報道資料より）

3.1 数理・データサイエンス・AI 教育認定制度

高等教育における数理・データサイエンス・AI 教育については、国による認定制度がすでに開始されており、拠点となる大学を中心に数理・データサイエンス・AI 教育強化拠点コンソーシアムが作られ、「リテラシーレベル」と「応用レベル」の2つの分野に関する基本カリキュラムが公開されている。



図 6：数理・データサイエンス・AI 教育強化拠点コンソーシアム

この中でも「リテラシーレベル」は文理を問わず、すべての学生が学ぶべき基礎的な内容であり、各大学が全学的な取り組みを始め、本年度も多くの大学が認定をとっている。

数理・データサイエンス・AI教育プログラム（リテラシーレベル）の要件		
<5つの審査項目と、モデルカリキュラムの対応箇所>		
項目	審査項目	モデルカリキュラム対応箇所
項目①	● 数理・データサイエンス・AIは、現在進行中の社会変化（第4次産業革命、Society 5.0、データ駆動型社会等）に深く関与しているものであること、また、それが自らの生活と密接に関わっているものであること。	導入 1-1. 社会で起きている変化 1-6. データ・AI利用の最新動向
項目②	● 数理・データサイエンス・AIが対象とする社会で活用されているデータや「データの活用領域」は非常に広範囲で、日常生活や社会の課題を解決する有用なツールになり得ること。	導入 1-2. 社会で活用されているデータ 1-3. データ・AIの活用領域
項目③	● 様々なデータ利用の現場におけるデータ活用事例が示され、数理・データサイエンス・AIは様々な適用領域（交通、製造、金融、サービス、インフラ、公共、ヘルスケア等）の知見と組み合わせることで価値を創出するものであること。	導入 1-4. データ・AI利用のための技術 1-5. データ・AI利用の現場
項目④	● ただし数理・データサイエンス・AIは万能ではなく、その活用に応じた様々な留意事項（ELSI、個人情報、データ倫理、AI社会原則等）を考慮することが重要であること。	心得 3-1. データ・AI活用における留意事項 3-2. データを守る上での留意事項
項目⑤	● 実データ・実課題（学術データ等を含む）を用いた演習など、社会での実例を題材として、「データを読む、説明する、扱う」といった数理・データサイエンス・AIの基本的な活用法に関すること。	基礎 2-1. データを読む 2-2. データを説明する 2-3. データを扱う

図 7：数理・データサイエンス・AI 教育リテラシーレベル

カリキュラムの内容としては、AIを用いた社会的・技術的な内容や、研究倫理などを含めた倫理観の育成、プログラミングに関しての基本的なサイクル、統計学の初歩など多岐にわたる。従来の大学の科目で喩えれば、「情報科学＋情報倫理＋統計学の初歩」といった内容と考えられる。

また、分野が基礎的であるにもかかわらず、広いためか、特定の学会によらない、大学間の実践研究の発表なども各所で頻繁で行われている。

第2回データサイエンス・AI授業実践研究ワークショップ
オンライン開催



開催趣旨

数理・データサイエンス・AIの授業実践の強化が求められており、多くの学系と多様なレベルの学業を擁する私立大学では、モデルカリキュラムの大幅な刷新も、各大学の担当教員で教育プログラムの内容や教材開発に向けて、具体的な授業設計の検討が求められておられるかと想います。そこで、当園は本ワークショップにおいて、リテラシーレベルの授業実践を行うこととしており、データサイエンス・AIを体感する授業構成、学生に興味・関心を持たせる授業の学術内容と工夫などの観点から、関係教員の方々と意見交換を行い、理解の促進を目指すこととしております。

1. 講演 17:00~18:00

「文系学生向けデータサイエンス授業の実践事例紹介」

辻 智氏 (大阪公立大学研究推進機構特任教授、成城大学非常勤講師)

モデルカリキュラムに基づいた実践コンテンツの例

- (1) データビジュアライゼーションの事例
世界のオープンデータ COVID-19 比較
- (2) テキストマイニングの事例
東京オリンピックに関する報道など
- (3) 画像処理の事例
ヒストグラムの意外な使い方など

2. 意見交換 18:00~18:30

・講師への質問、授業での課題や今後の取組み計画などを議論いただきます。

公益社団法人私立大学情報教育協会 情報教育研究委員会データサイエンス教育分科会

図 8: データサイエンス・AI 授業実践研究ワークショップ

4 データサイエンス教育における高大接続

本学においては、数理・データサイエンス・AI教育に関する科目として、「データサイエンス基礎」「データサイエンス実践」の2つの一般教養科目を作成し、数理・データサイエンス・AI教育強化拠点コンソーシアムが提供するリテラシーレベルの内容に準拠した内容で、構成をした。

科目の構成にあたり、国が初等中等教育との連続性を意識した科目でもあるので、連続性を意識して、高等学校の前課程の「情報の科学」「社会と情報」や新課程の「情報I」「情報II」との連携も意識し、色分けした項目対応表を作成した。

表 1: リテラシーレベルと高等学校情報科の項目比較

項目	単元番号	タイトル	赤: 情報IIまで, 緑: 情報Iまで, 紫: 高校旧課程まで
項目①	1-1	社会の変化	<ul style="list-style-type: none"> ・ビッグデータ, IoT, AI, ロボット ・データ量の増加, 計算機の処理性能の向上, AIの非連続的進化 ・第4次産業革命, Society 5.0, データ駆動型社会 ・複数技術を組み合わせたAIサービス ・人間の知的活動とAIの関係性 ・データを起点としたものの見方, 人間の知的活動を起点としたものの見方
	1-6	データ・AI活用の最近の状況	<ul style="list-style-type: none"> ・AI等を活用した新しいビジネスモデル(シェアリングエコノミー, 商品のレコメンデーションなど) ・AI最新技術の活用例(深層生成モデル, 敵対的生成ネットワーク, 強化学習, 転移学習など)
項目②	1-2	社会におけるデータの活用	<ul style="list-style-type: none"> ・調査データ, 実験データ, 人の行動ログデータ, 機械の稼働ログデータなど ・1次データ, 2次データ, データのメタ化 ・構造化データ, 非構造化データ(文章, 画像/動画, 音声/音楽など) ・データ作成(ビッグデータとアナレーション) ・データのオープン化(オープンデータ)
	1-3	データ・AIの活用領域	<ul style="list-style-type: none"> ・データ・AI活用領域の広がり(生産, 消費, 文化活動など) ・研究開発, 調達, 製造, 物流, 販売, マーケティング, サービスなど ・仮説検証, 知識発見, 原因究明, 計画策定, 判断支援, 活動代替, 新規生成など
項目③	1-4	データ・AIの技術	<ul style="list-style-type: none"> ・データ解析: 予測, グルーピング, パターン発見, 最適化, シミュレーション・データ同化など ・データ可視化: 複合グラフ, 2軸グラフ, 多次元の可視化, 関係性の可視化, 地図上の可視化, 挙動・軌跡の可視化, リアルタイム可視化など ・非構造化データ処理: 自然言語処理, 画像/動画処理, 音声/音楽処理など ・特化型AIと汎用AI, 今のAIで出来ることと出来ないこと, AIとビッグデータ ・認識技術, ルールベース, 自動化技術
	1-5	データ・AIの利活用の現場	<ul style="list-style-type: none"> ・データサイエンスのサイクル(CRISP-DM, 課題抽出と定式化, データの取得・管理・加工, 探索的データ解析, データ解析と推論, 結果の共有・伝達, 課題解決に向けた提案) ・流通, 製造, 金融, サービス, インフラ, 公共, ヘルスケア等におけるデータ・AI活用事例紹介

項目	単元番号	タイトル	赤：情報Ⅱまで、緑：情報Ⅰまで、紫：高校旧課程まで
項目④	3-1	データ・AIを活用する上での留意事項	<ul style="list-style-type: none"> ・ ELSI (Ethical, Legal and Social Issues) ・ 個人情報保護, EU一般データ保護規則(GDPR), 忘れられる権利, オプトアウト ・ データ倫理: データのねつ造, 改ざん, 盗用, プライバシー保護 ・ AI社会原則 (公平性, 説明責任, 透明性, 人間中心の判断) ・ データバイアス, アルゴリズムバイアス ・ AIサービスの責任論 ・ データ・AI活用における負の事例紹介
	3-2	データを守るための留意事項	<ul style="list-style-type: none"> ・ 情報セキュリティ: 機密性, 完全性, 可用性 ・ 匿名加工情報, 暗号化, パスワード, 悪意ある情報搾取 ・ 情報漏洩等によるセキュリティ事故の事例紹介
項目⑤	2-1	データを読む	<ul style="list-style-type: none"> ・ データの種類 (量的変数, 質的変数) ・ データの分布(ヒストグラム)と代表値 (平均値, 中央値, 最頻値) ・ 代表値の性質の違い (実社会では平均値=最頻値でないことが多い) ・ データのばらつき (分散, 標準偏差, 偏差値) ・ 観測データに含まれる誤差の扱い ・ 打ち切りや脱落を含むデータ, 層別の必要なデータ ・ 相関と因果 (相関係数, 擬似相関, 交絡) ・ 母集団と標本抽出 (国勢調査, アンケート調査, 全数調査, 単純無作為抽出, 層別抽出, 多段抽出) ・ クロス集計表, 分割表, 相関係数行列, 散布図行列 ・ 統計情報の正しい理解 (誇張表現に惑わされない)
項目⑤	2-2	データで説明する	<ul style="list-style-type: none"> ・ データ表現 (棒グラフ, 折線グラフ, 散布図, ヒートマップ) ・ データの図表表現 (チャート化) ・ データの比較 (条件をそろえた比較, 処理の前後での比較, A/Bテスト) ・ 不適切なグラフ表現 (チャートジャンク, 不必要な視覚的要素) ・ 優れた可視化事例の紹介 (可視化することによって新たな気づきがあった事例など)
項目⑤	2-3	データを扱う	<ul style="list-style-type: none"> ・ データの集計 (和, 平均) ・ データの並び替え, ランキング ・ データ解析ツール (スプレッドシート) ・ 表形式のデータ (csv)
オプション			大学の統計学や機械学習として扱われる内容, プログラミング等

この対応表を見ると、高等学校の旧課程で学んだ現在の大学1年生や2024年度入試までで入ってくる大学生にとっては、項目⑤の一部を除いて、新出の内容が多いために詳細な資料が必要であることがわかるが、2025年度から入学してくる学生に関しては、1-6,3-1,3-2などの単元以外に関してはある程度の内容を学んできてくることがわかる。まさに、高等学校の「情報2.5」と言えるような内容であるといってもよいだろう。

この項目対応表を元に、現状ではすべての項目を網羅できるような内容で教材やオンデマンドビデオを作成している。また、本学が医療系大学である特徴を活かし、素材に関しては、意識して医療系の題材を多く扱っていることを特徴としている。本学における数理・データサイエンス・AI教育のリテラシーレベルの教材の作成はすでに終わり、学生や教員の評価を受けて修正をする段階になっている。現時点では好評であり、学生の評価も高く、他の科目でのAI活用を考える上で役に立っているようである。本学に関しては、来年度に認定を受けるように準備中である。

高等学校の新課程における必修教科目の「情報Ⅰ」や首都圏で高い採択率となっている「情報Ⅱ」を学んできた生徒が大学に入学する2025年には、このカリキュラムは大幅に見直す必要があると考えている。数理・データサイエンス・AI教育強化拠点コンソーシアムにおいても数年内に新しいリテラシーレベルのカリキュラムを作成する予定であることが報告されている。

5 今後の課題

今後、数理・データサイエンス・AI教育に関してのシステムを系統的に構築し、オンラインテストなども実施できるように整備中である。また、他学部や他の担当教員、

学生が使えるような教材の作成も進行している。他学部その他の科目との連携などに関しても本センターにおいて検討を行っており、企業との連携や関連イベントの開催なども検討中である。

特殊な学部が多く、キャンパスがそれぞれ離れている点から、縦割りでの科目構成が進行していた中で、学部横断的な一般教養科目として「データサイエンス基礎」を導入していくことに、様々な問題点はあるが、これがきっかけとなり学生の研究や他職種連携などへの応用など幅広い活用について、教員も学生も検討を進めている。

6 おわりに

本学のデータサイエンス教育は今年始まったばかりであり、これから認定を申請する段階である。そのためまだまだ改良が多くみられる内容ではあるが、他大学の実践も参考に学生にとって、有意な授業科目を作成していきたいと考えている。また、応用レベルに関しても来年度開設される「健康データサイエンス学部」との連携を考え、進めていきたいと考えている。

参考文献

- [1] 文部科学省：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, 2018.
- [2] 文部科学省：高等学校学習指導要領解説(情報編)，
http://www.mext.go.jp/component/a_menu/education/micro_detail/___icsFiles/afieldfile/2018/07/19/1407085_16_01.pdf, 2018.
- [3] 文部科学省：高等学校情報科「情報I」教員研修用教材，
https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1416756.htm, 2019.
- [4] 文部科学省：高等学校情報科「情報II」教員研修用教材，
https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/mext_00742.html, 2020.
- [5] 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)，
https://www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/00002.htm, 2021.
- [6] 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/>.