

原著

高等学校「数学Ⅰ」教科書「データの分析」における 生成 AI 活用に関する研究

A Study on Use of Generative AI in Analysis of Data in High School ‘Mathematics I’ Textbook

市川隆司

大阪信愛学院大学 教育学部

要旨

今後の社会に必要とされるデータリテラシー獲得のため、学習指導要領にデータの活用やデータの分析に関する内容が追加された。2022年に公開された対話型生成 AI は、瞬く間に利用者が増加し、教育分野にも影響を及ぼし始めている。生成 AI の教育利用研究は始まったばかりであり、どのような有効活用が可能か様々な面から検討が必要である。本研究では、高等学校数学Ⅰ教科書の「データの分析」項目に生成 AI をいかに活用できるかを検討した。AI の回答から計算の出力には課題があるものの、考え方やプログラムコード生成には有用であることが確認できた。

キーワード：データサイエンス, AI, 中等教育, 数学, 教科書

1. はじめに

OECD が 2015 年から進めている Education 2030 プロジェクトでは、2030 年に望まれる社会のビジョンを実現する主体となる子どもたちのコンピテンシーが議論され、学びの中核的な基盤として必要とされる基礎能力にデータリテラシーが不可欠とされている¹⁾。

日本では 2016 年学習指導要領改訂に向けた中央教育審議会答申において、全ての学習の基盤として育まれ活用される資質・能力に統計的な分析に基づき判断する力が取り上げられた²⁾。その後 2017 年から順次改訂された現行学習指導要領で、小学校算数に「データの活用」、中学校数学に「データの活用」、高等学校数学に「データの分析」「統計的な推測」、情報に「データの活用」の内容が設定された。初等中等教育を通じてデータを取り扱う統計的な内容が追加され、教育内容の改善に歩を進めることになった³⁾。高等学校数学においては「統計的な推測」は数学 B の内容となり、数学Ⅰのみで高等学校数学の履修を終える生徒もいるため、「データの分析」の内容が高校生全体で共通した学習内容となる。

2022 年 11 月に公開された対話型生成 AI である OpenAI 社の ChatGPT は、わずか 2 ヶ月で利用者 1 億人を突破し、労働市場に大きな影響をもたらすことが指摘されている⁴⁾。その影響は、経済分野に留まらず、教育分野にも大きな影響を及ぼしつつある。

高等教育では、文部科学省から各高等教育機関に対して生成 AI の教学面での取扱いについて学生や教職員にガイドラインを示すことが求められた⁵⁾。また初等中等教育に対しては、文部科学省から生成 AI の利用に関する暫定的なガイドラインが示され、メリット、デメリットを十分理解し生成 AI 活用に関する適否を考慮して教育利用することが求められている⁶⁾。このガイドラインは、暫定的な内容とされており、今後の AI の進展や状況の変化によって改訂されることが予告されている。

藤村(2023)は、教育分野における生成 AI 利用研究は不十分であることを指摘しており⁷⁾、このような状況下で生成 AI の教育利用の研究は端緒を開いたところである。初等中等教育段階から生成 AI を活用することにより児童生徒が AI との対話から課題を認識し、生成 AI の有用性と限界を理解することで批判的思考力を高め

ていくことが考えられる。ひいては生成 AI が児童生徒に必要とされる汎用的リテラシーに関する資質能力を獲得していく上で役立つ可能性を有している。その中でデータリテラシーに関する学習も生成 AI 活用の要素を十分に備えている。

2. 研究の目的

中等教育においてデータの活用や分析に関する学習内容は、データサイエンスの素養を獲得する上で重要である。その学習内容に対して生成 AI をいかに活用できるかを高等学校数学 I 教科書の「データの分析」の項目について検討し、基礎的な知見を得ることを本研究の目的とする。

ChatGPT を始めとする対話型生成 AI は大規模言語モデル(LLM: Large Language Model)に基づくシステムである。LLM は、大規模なデータセットを用いた機械学習によって精度を高めた自然言語処理モデルである。

対話型生成 AI の回答には、あたかも人と対話しているかのように回答を出力するが、一方で回答内容には著作権侵害の恐れやハルシネーション（幻覚）と言われるウソが含まれる。このような生成 AI のデメリットを踏まえた上で教育利用を検討する必要があり、種々の質問に対する回答から基礎的情報を得ることは活用方法の検討につながるものと考えられる。

3. 研究の方法

本研究においては、中等教育におけるデータリテラシー育成の観点から高等学校数学の教科書内容を生成 AI 適用の分析対象とした。令和 5 年発行の高等学校「数学 I」教科書 5 社 16 種類のうち、東京都立高等学校で最も採用校数の多い数研出版新編数学 I 第 5 章「データの分析」の内容（表 1）をもとに質問に対する生成 AI の回答結果を分析した⁸⁾⁹⁾。

表1 新編数学I 第5章「データの分析」の学習内容

節	項目	内容	設問データ	
1	データの整理	度数分布表		練習 1
		ヒストグラム		練習 2
2	データの代表値	平均値	例 1	練習 3
		最頻値	例 2	練習 4
		中央値	例 3	練習 5
3	データの散らばりと四分位数	範囲	例 4	練習 6
		四分位数	例 5, 例 6	練習 7
		箱ひげ図	例 7	練習 8
		ヒストグラムと箱ひげ図		
		外れ値		練習 9
4	分散と標準偏差	分散, 標準偏差	例 8, 例 9	練習 10
		分散と平均値の関係式	例 10	練習 11
5	2つの変量の間関係	散布図	考察	
		正の相関, 負の相関		練習 12
		相関係数	考察	練習 13
			例 11	練習 14
		相関関係と因果関係		
	質的データをとる2つの変量の間関係		練習 15	
6	仮説検定の考え方	仮説検定の考え方	例 12	練習 16

日本の学校教育における教科書は、文部科学省の教科書検定制度のもとで審査されており、どの教科書も審査に合格した内容である。従って学習指導要領の内容は網羅されており、構成が明確である。東京都は都道府県で18歳人口が全国最大であり、検定教科書の中で採択校数が最大の教科書を取り扱うことは、代表事例として生成AI活用の理解を広げる可能性があるものと考えられる。

比較対象とする生成AIは、2023年9月現在一般利用者が無料で使用でき、かつ日本語での対話が可能なChatGPT3.5（以下ChatGPT）、Bingチャット（以下Bing）、Google Bard（以下Bard）の3種類である。Bingでは会話のスタイルとして「より厳密に」を選択した。

Webブラウザには、ChatGPTとBardを使用する際はGoogle Chromeを、Bingを使用する際はMicrosoft Edgeを使用した。生成AIに回答させた内容をもとに概要をまとめた。

3.1. ChatGPTの概要¹⁰⁾

ChatGPTは、OpenAI社が開発したLLMの一つであり、GPT-3.5アーキテクチャをベースにしている。GPT-3.5は2021年9月までの学習データに基づいており、1750億個のパラメータで処理されている。有料のGPT-4は非公開ながら5000億個以上のパラメータを有するとされており、パラメータ数が多いほど精度は高まる。

3.2. Bingチャットの概要¹¹⁾

BingはMicrosoft社によって開発され、LLMとしてOpenAIのGPT-4を使用し、独自の技術Prometheusを組み合わせた検索エンジン型のAIである。対話から検索キーワードを拾い出し、ネット検索してその情報を活

用して回答を出力する。そのため事前に学習したデータに加えて最新データも検索から得るしくみになっている。また回答下部に情報源となる詳細情報に関するリンクが表示される。

3.3. Bard の概要¹²⁾

Bard は Google AI が開発した PaLM という LLM を使用し、テキストとコードの膨大なデータセットで強化学習された AI である。パラメータ数は日本語で 137 億個、英語で 1.56 兆個である。さらに Pathways といわれる機械学習のワークフレームを使用すると日本語パラメータ数は 540 億に拡張される。Bard の品質向上のため、回答を 2 つ並べて表示し選択できる機能がある。

教育利用を検討する場合に備えて利用規約を整理すると利用には、ChatGPT は 13 歳以上、かつ 18 歳未満の場合保護者の同意が必要である。また Bing は Microsoft アカウントの利用規約で成人とされており、未成年の場合は利用に親権者の同意が必要である。Bard の利用は 18 歳以上とされている。

4. プロンプトの実行と結果

高校生が利用することを想定して対象教科書記載のデータ、表、練習問題に基づき各生成 AI のプロンプトに質問を入力し、回答を実行したところ以下の結果出力を得た。練習問題の設問は、教科書の記載に従いながら、設問意図が変わらないよう命令口調をていねいな口調に編集して生成 AI のプロンプトに入力した。表のデータはプロンプトにテキストとして貼り付けた。また入力した質問に対して対話による応答であることから、解答ではなく、出力は回答と記している。

4.1. データの整理

4.1.1. 度数分布表

東京の 2018 年 4 月の日ごとの最高気温のデータおよびそのデータに基づいて作成された表 2 の度数分布表を用いて以下の設問(1)から(3)をプロンプトに入力し回答を出力させた。

- (1) 度数が 9 である階級の階級値は求めてください。

ChatGPT, Bing, Bard のいずれも、階級は 24 以上 27 未満で、階級値は 25.5 と適切な回答を出力した。

- (2) 最高気温が、21°C 以上の日は何日ありますか。

ChatGPT は、各データが 21°C 以上であるか未満であるかを判断して正しい回答 20 日を出力した。Bing は、21°C 以上のデータを順番に出力し正しい日数を出力した。Bard の回答は不正確な結果であったが、日数を求めるための Python のコードが出力された。このコードを Web ブラウザ上で Python を実行できる Google Colaboratory で実行したところ、正しい値が出力された。

- (3) 最高気温が低い方から数えて 12 番目の日のデータを教えてください。

ChatGPT は、(1)の質問に続いて(3)を尋ねたところ、誤った結果を出力した。そこで最高気温を昇順に並べ替えるよう指示するとソートされたデータが出力された。続いて再度 12 番目のデータを尋ねると正しい値 21.8°C を回答した。Bing に関しても ChatGPT と同様の結果となった。Bard の出力結果は不正確であった。さらに最高気温を昇順に並べ替えるよう指示をしたが、並べ替えることができなかった。

表2 最高気温の度数分布表

階級(°C)	度数(日)
15 以上 18 未満	4
18~21	6
21~24	10
24~27	9
27~30	1

4.1.2. ヒストグラム

データ数が20のハンドボール投げの記録をもとに次の質問をした。

- (1) 11m 以上 13m 未満を階級の1つとして、どの階級の幅も2mである度数分布表を作ってください。

対象となるデータを用いた度数分布表の作成において、いずれの生成AIも階級は正しく認識したが、出力された度数が正確ではなかった。教科書ではヒストグラム作成の設問があるが、無料で使用できる生成AI単独ではグラフの作成ができない。そこで正しい度数を含む度数分布表からヒストグラムを作成するPythonコードの出力を回答させた。ChatGPTとBingは出力されたコードを実行すると適切な図1のヒストグラムを作成したが、Bardは適切なコードが出力できなかった。

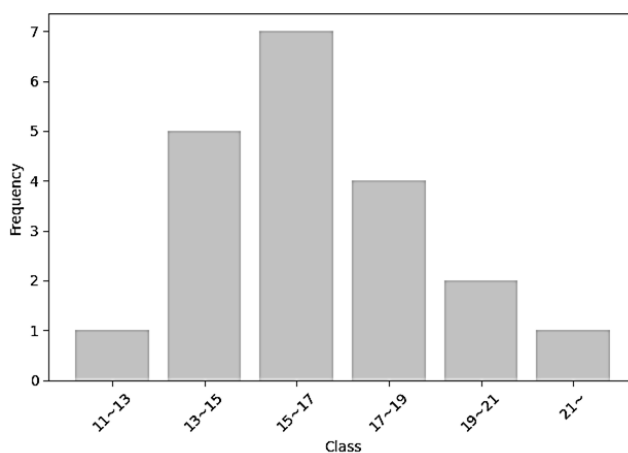


図1 ある高校の1年生女子20人の、ハンドボール投げのヒストグラム

4.2. 代表値

4.2.1. 平均値

すべて2桁の数値でデータ数が5つのデータセットにおいては、すべての生成AIで正しい値の回答を得ることができた。

次にすべて3桁の数値でデータ数が8つのデータセットになるとChatGPTでは平均値を計算する際に、その過程で合計を間違えて正しい平均値が出力されなかった。合計の誤りを指摘して、再度質問したが正しい結果は得られなかった。次に段階を経て回答を得ようとまず合計を出力させたが、この段階で正しい合計が得られなかった。そこで再度誤った合計をプロンプトに入力して正しいかどうか尋ねたところ、誤っていることを認識し、正しい合計が回答された。

Bingは正しい計算結果を出力できなかった。合計の誤りを指摘して、再度質問したところ正しい値が回答された。

Bardは、結果の表示が不正確であった。3回合計の誤りを指摘して回答を求めたが、いずれの場合も正しい

値が出力されることはなかった。そこで平均値を求める Python コードの出力を回答させたところ、出力されたコードから正しい値が得られた。

単純な計算の回答精度を確認するため、1桁の数値でデータ数10のデータセットをランダムに作成して合計を回答させた。ChatGPTとBingはすべて正しい合計を出力した。Bardは10セット中正しい合計を出力したのは1セットのみであった。次に2桁の数値でデータ数10のデータセットをランダムに作成して合計を回答させた。ChatGPTとBardは10セット中正しい合計を出力したのは1セットのみであった。Bingは10セット中9セットにおいて正しい合計を出力した。さらに3桁の数値でデータ数10のデータセットをランダムに作成して合計を回答させた。ChatGPTは10セット中正しい合計を出力したのは1セットのみであった。Bingは2桁と同様に10セット中9セットにおいて正しい合計を出力した。Bardは10セット中正しい合計をすべて出力しなかった。

4.2.2. 最頻値

表3をもとに最頻値を各生成AIに回答させた。

ChatGPT, Bing, Bardのすべてで、正しい値26.0の回答が出力された。

表3 成人男子100人の靴のサイズ

サイズ(cm)	24.0	24.5	25.0	25.5	26.0	26.5	27.0
人数	3	11	18	23	32	10	3

次に前述表2の度数分布表を用いて最頻値を求めた出力結果では、Bardのみ正しい結果が得られなかった。

4.2.3. 中央値

表4にあるデータ数の異なるある商品のA市とB市それぞれの価格の中央値を回答させた。

ChatGPTおよびBingはデータ数が偶数および奇数いずれの場合も正しい結果を出力した。Bardは、どちらも正しい結果が出力されなかった。また別のすべて2桁の数値でデータ数8のデータセットから中央値を求めた。ChatGPTとBingは正しい値が出力されたが、Bardは正しい値が出力できなかった。

表4 ある商品の価格(円)

A市	260	280	280	300	270	
B市	280	280	260	100	280	270

4.3. データの散らばりと四分位数

4.3.1. 範囲

データの範囲を求める場合、データセット中の最大値から最小値を引く計算となる。月ごとの降水日数のデータ数12のデータセットからデータの範囲を回答させた。ChatGPT, Bing, Bardのいずれも、正しい結果が出力された。Bardは、値に加えてPythonのコードも出力された。このコードを実行したところ、正しい結果が得られた。

次にデータ数が同じ12の異なるデータセットと範囲を比較する質問を行った。2つのデータセットで範囲を比較した場合、ChatGPT, Bing, Bardのいずれも、散らばりの度合いを比較して適切に比較することができた。加えてBingとBardでは範囲だけで散らばりを評価することは不十分であり、分散や標準偏差を求めて正確さを期すよう補足説明が出力された。

4.3.2. 四分位数

10人の生徒に100点満点のテストを行った結果を、値の大きさの順に並べたデータで四分位数について回

答を求めた。

1 群のデータセットについて四分位数を求めたところ、ChatGPT で第2 四分位数（中央値）は正しい結果を出力したが、第1 四分位数および第3 四分位数は正しい結果を出力しなかった。Bing の出力も ChatGPT と同様の結果を出力した。Bard は第1 四分位数、第2 四分位数、第3 四分位数のすべてで正しい結果が出力されなかった。「データ数は 10 なので、中央値は 5 番目の値」と中央値に関する定義も誤って回答された。

表5 のデータ数が同じ異なる2 群のデータセットに関する四分位数は、2 群のデータセットを同時にプロンプトに入力した場合、ChatGPT では、第1 四分位数は正しい値であったが、それ以外は誤っていた。Bing は第1 四分位数、第2 四分位数、第3 四分位数とも正しい結果を出力した。そのため四分位範囲についても正しい値となり、データの散らばりの度合いについて正しい判断を行った。Bard は、第2 四分位数を出力せず、第1 四分位数および第2 四分位数とも正しい結果を出力しなかった。

この2 群のデータセットを1 群ずつ対話形式で四分位数を求めた。Bing では、正しい値が出力され、次に各データセットの四分位範囲を求めて散らばりの度合いについて尋ねたところ、正しい結果が出力された。ChatGPT および Bard では最初のデータセットで四分位数に正しい結果が出力されなかった。

表5 2 群のデータセット

A	21	29	32	36	38	40	49	53	55	68	80
B	25	31	39	42	45	46	50	53	54	65	80

4.3.3. 箱ひげ図

表6 の3 群のデータセットに関して箱ひげ図を並べて描く設問を考える。無料の生成AI で直接箱ひげ図を描くことはできない。そこで ChatGPT, Bing, Bard のそれぞれ箱ひげ図を描く Python コードを出力させ、Google Colaboratory でそのスクリプトを実行させたところ、すべての生成AI で図2 の箱ひげ図を描くことができた。

ChatGPT および Bing では、3 群のデータセットに関する散らばり度合いの比較に標準偏差を使用するよう回答が Python コードとともに出力された。出力されたスクリプトを実行したところ、正しい値が得られた。Bard においても分散や標準偏差を用いることでデータの散らばりの度合いがより正確に判断できると回答された。

表6 生徒10 人のテストの得点（点）

数学	68	35	86	63	30	91	50	63	46	58
英語	75	65	90	78	52	88	70	75	59	82
国語	63	60	73	75	58	79	68	70	66	80

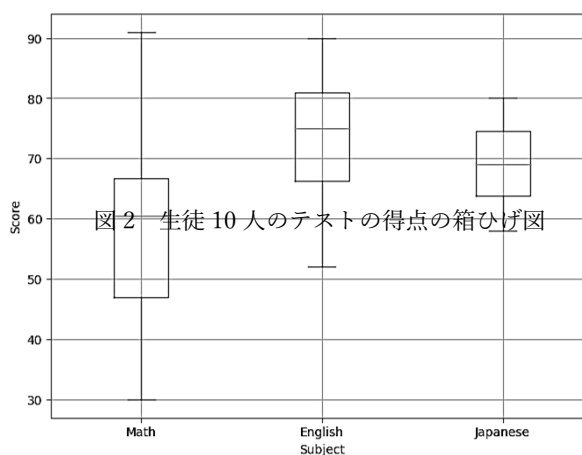


図2 生徒10 人のテストの得点の箱ひげ図

4.3.4. 外れ値

箱ひげ図から最大値および最小値について外れ値かどうかを判断させる設問がある。教科書に記載された箱ひげ図から中央値 48, 第 1 四分位数 40, 第 3 四分位数 60 を与えてプロンプトに入力し, 外れ値を判断させた。

ChatGPT では, 四分位範囲から外れ値の範囲を求めて, 最大値 92 が外れ値であり, 最小値 23 は外れ値ではないと正しい判断を回答した。Bing も ChatGPT と同様の判断を回答した。Bard は, 外れ値の範囲を判断できず, 曖昧な判定を回答した。

4.4. 分散と標準偏差

10 点満点の漢字テストでデータ数 10 のデータセットに対して分散と標準偏差の値を求めたところ, ChatGPT および Bing では, 正しい値が回答された。Bard では正しい値が出力されなかった。

データ数が同じ 10 の異なるデータセットに対して同様に分散と標準偏差の値を求めたところ, ChatGPT および Bing では正しい値が回答されたが, Bard では正しい値は出力されなかった。

さらにデータ数が同じ 10 の異なるデータセットに対して同様に分散と標準偏差の値を求めたところ, ChatGPT および Bing では正しい値が回答されたが, Bard では正しい値は出力されなかった。

4.5. 散布図

表 7 のデータセットから散布図を描く Python コードを出力させた。出力したコードを実行したところ, ChatGPT, Bing, Bard のいずれも適切な散布図を描いた。

表 7 陸上部員男子 20 人の身長 x(cm) と体重 y(kg)

	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩
x	168.4	164.5	171.2	173.0	162.3	170.8	172.5	164.2	169.0	168.6
y	59.0	58.4	60.5	65.6	54.2	61.4	62.2	56.1	63.7	57.3
	⑪	⑫	⑬	⑭	⑮	⑯	⑰	⑱	⑲	⑳
x	172.6	166.0	173.7	176.4	178.5	167.5	177.8	174.6	172.3	173.5
y	64.1	56.5	68.3	68.2	69.6	61.2	66.4	66.7	60.9	70.8

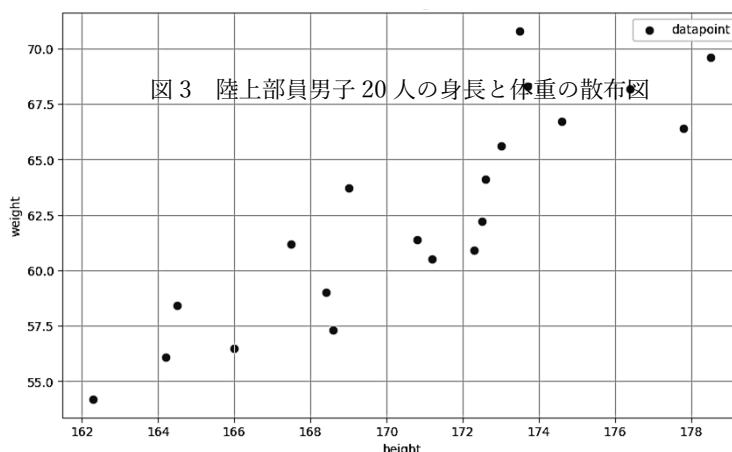


図 3 陸上部員男子 20 人の身長と体重の散布図

4.6. 2つの変量の間関係

4.6.1. 相関関係

表 8 から緯度と平均気温について相関関係を問うたところ, ChatGPT, Bing, Bard のいずれも関係を把握す

る適切な相関係数を出力することはなかった。そこですべての生成 AI に相関係数を求める Python コードを出力するように求めた。

出力された Python コードを実行したところ、すべての生成 AI で正しい相関係数が出力された。出力された相関係数から相関関係について尋ねたところ、すべての生成 AI で正しく負の相関と回答された。

表 8 各地点の緯度 x (度) と 2018 年 4 月の平均気温 y ($^{\circ}\text{C}$)

地点	札幌	青森	仙台	東京	長野	大阪	高知	鹿児島
x	43.1	40.8	38.3	35.7	36.7	34.7	33.6	31.6
y	8.2	9.6	12.5	17	13.1	16.9	17.3	18.5

4.6.2. 相関係数

x の標準偏差 4.40, y の標準偏差 4.71 と x と y の共分散 18.22 から, x と y の相関係数を求めるよう質問したところ, ChatGPT では正しい計算式を回答するが, 正しい計算結果は出力されなかった。Bing においても同様であった。Bard は計算式が正しくなく, 正しい計算結果を出力できなかった。

表 9 について太さ x と高さ y の相関係数を求めるよう質問したところ, ChatGPT, Bing, Bard のいずれも正しい相関係数を出力できなかった。

次に表 10 についてテスト A とテスト B の相関係数を求めるよう質問したところ, ChatGPT および Bing では相関係数を求める計算式は出力するが, 計算過程の数値を誤り正しい結果を出力できなかった。Bard は相関係数を求める計算式は出力せず, 正しい数値も出力しなかった。相関係数を求める Python コードを出力させたところ, 3 つの生成 AI すべてで正しい値が得られるコードが出力された。

表 9 同じ種類の木の根もとの太さ x (cm)と高さ y (m)

	①	②	③	④	⑤
x	21	27	29	23	30
y	13	20	19	17	21

表 10 2 種類のテストの得点結果

生徒番号	①	②	③	④	⑤	⑥
テスト A	5	7	5	4	3	6
テスト B	4	1	3	5	9	2

4.6.3. 質的データをとる 2 つの変量の間関係

表 11 に示すクロス集計されたデータから教材 A と B の種類が合否に影響したかを尋ねた。ChatGPT はクロス集計表を適切に読み取れず正しい判断が出力できなかった。Bing は, クロス集計表を正しく読み取り, 教材 A を使用した場合の合格率は 64%, 教材 B を使用した場合の合格率は 79% と回答され, 2 つの教材がそれぞれ合否に及ぼす影響を出力した。Bard はクロス集計表をスプレッドシートに変換できる形式に読み取ったが, 表内の人数を正しく読み取れず合否の割合は正しく出力しなかった。

表 11 教材 A、B の使用の有無による合否の人数

	合	否
A 有 B 有	6	1
A 有 B 無	3	4
A 無 B 有	31	9
A 無 B 無	11	35

4.7. 仮説検定

30 人に対して A と B の 2 種類のボールペンの書きやすさを尋ねて、70%にあたる 21 人が B を選んだ場合の有意差について回答させた。

ChatGPT は、仮説検定の説明から二項分布の考え方を出力するが、計算した値は出力しなかった。次に確率を求める計算をするよう質問しても正しい値は出力されなかった。Bing は、明確な根拠を示さず B の方が書きやすいと出力した。次に仮説検定を行うよう質問すると帰無仮説と対立仮説の考え方を回答して Python コードを出力した。このコードを Google Colaboratory で実行すると正しい値が表示された。Bard は、70%という値から B の方が書きやすいと回答を返し、特に検証する方法は回答しなかった。

5. 考察

5.1. 生成 AI の出力結果

計算に関して生成 AI が出力する結果は、表計算ソフトで計算結果を表示させているしくみと異なり、事前学習に基づいた情報による回答である。そのため同じ質問をプロンプトに入力しても計算結果が必ずしも毎回同じになるとは限らない。以下に今回の出力結果を考察した。

表 12 に示すグラフ作成を除くデータの分析に関して出力結果を全般的に検討すると、現状では ChatGPT および Bing が Bard より正しい結果を回答する割合が優る。さらに ChatGPT と Bing を比較すると、Bing の方が正しい出力結果の可能性が高まる。これは Bing のシステムが OpenAI 社の GPT-4 をベースとして Microsoft の検索エンジンと組み合わせられて使用されているためと推測される。

代表値に対する出力結果から考えると、データの比較というよりも数字という文字の比較は認識しやすいと考えられる。数値の大小ではなく、文字の出現や順番が手がかりとなっていることが推測できる。合計の計算を繰り返し試行したが、データが 1 桁の計算では正しい値を回答する可能性が高い一方でデータが 2 桁の以上になると Bing 以外は正しい値を回答しなかった。計算自体は単純でもデータの内容によって出力結果に影響を受ける。計算の出力結果は、西田(2023)が指摘するように生成 AI は必ずしも正しい計算結果を出力しない場合があることを確認できた¹³⁾。

2 群のデータセットを比較する場合対話形式で 1 群ずつ結果を出力させ、次のステップで比較する手順を用いると出力結果の精度があがる可能性が高い状況が見られた。

統計的な内容を検討すると相関係数は生成 AI への質問で計算式や計算過程を出力する場合がある。その場合過程のどこで誤りが発生しているのかをチェックすることによって求めたい値に対する計算方法の理解が深まる。

生成 AI から出力された Python コードを実行してヒストグラム、箱ひげ図、散布図を適切に作成できたが、生成 AI 単独でグラフ作成は行えない。有料の ChatGPT Plus を使用し、かつプラグインを導入すれば可能であることから今後の機能向上が望まれる。

表 12 処理結果の出力数値まとめ

処理項目	ChatGPT	Bing Chat	Bard
度数分布表の階級値	○	○	○
度数分布表で用いたデータの抽出	○	○	×
度数分布表で用いたデータの順位	×	×	×
平均値	△	△	△
最頻値	○	○	△
中央値	○	○	△
データの範囲	○	○	○
四分位数	△	△	×
外れ値	○	○	×
分散, 標準偏差	○	○	×
相関係数	×	×	×
質的データをとる2つの変量の間関係	×	○	×

○：出力結果が正確 ×：出力結果が不正確 △：出力結果が正確な場合と不正確の場合がある

5.2. 生成AIの活用方法

数学Ⅰの「データの分析」に関して、計算を含む質問に対する生成AIの出力結果を学習の過程で提示し、誤った出力結果を学習者に確認させて、点検することで学習内容の理解を深めることに活用できる。

計算式を理解する上では、生成AIには順を追って詳細に説明を出力する利点がある。計算過程を提示してくれる場合は、学習内容の理解が特に進むと考えられる。

計算を回答させた場合誤った値が出力される可能性があるのに対して、ChatGPTやBingで出力されたPythonコードは正しい結果を導く回答である可能性が高く、プログラミングのソースコード生成として有用性を評価できる。学習指導要領では「データの分析」でコンピュータなどの情報機器を適切に活用する記載はあるが³⁾、数学Ⅰの教科書ではプログラミングを直接扱う内容はない。

統計的な内容は、生徒の実態に応じて数学科と情報科相互に関連を図る工夫の必要性が学習指導要領で示されている³⁾。数学Ⅰの「データの分析」と情報Ⅰの「データの活用」は同学年で履修することが多く、どちらも教科書後半の内容である。その点からカリキュラム・マネジメントによって生成AIの出力結果を有効活用した教科の連携が考えられる。

「情報Ⅰ」ではプログラミングが取り扱われ、Pythonの活用を記載する教科書がある。データの分析・活用の内容と関連付けてPythonを活用した学習を進めることは、データの分析や活用とプログラミングの相互に学習効果を発揮する可能性が見込まれる。

6. まとめ

高校生が生成AIを利用するには、各利用規約を遵守して活用する必要があり、授業では教師の指導の下に利用することが重要である。そのため教師にも一定のAIリテラシーが必要となる⁶⁾。

高等学校数学Ⅰの内容について「データの分析」に関して現状の生成AI出力結果を検討する限り、必ずしも正しい結果を出力できない。しかし目的とする処理内容に応じて生成AIを活用することにより、学習課題の理解や解決につながる可能性があることが示唆された。

プロンプトから生成内容を適切に指示できれば、目的に沿う回答結果を得られる。現状では英語の方が日本語よりもパラメータ数が多いため、日本語で質問するよりも確かな回答が得られる場合が多い。そのためには英語によるプロンプトの入力が必要となる。その点プログラムコード生成は言語による影響を受けにくい。

現状ではデータの分析に表計算ソフトの操作を理解して活用することは避けられないが、生成 AI を併用することによって学習内容の理解が深まることを期待する。

なお、本論文に関して、開示すべき利益相反関連事項はない。

参考文献

- 1) OECD. THE FUTURE OF EDUCATION AND SKILLS Education 2030. p.4, 2018
<https://www.oecd.org/education/2030-project/contact/> (参照日 2023年9月1日)
- 2) 中央教育審議会. 幼稚園, 小学校, 中学校, 高等学校および特別支援学校の学習指導要領等の改善および必要な方策等について (答申). p. 35, 2016
https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/_icsFiles/afieldfile/2017/01/10/1380902_0.pdf (参照日 2023年9月2日)
- 3) 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. p.6, p.11, p.44
https://www.mext.go.jp/content/20230217-mxt_kyoiku02-100002620_05.pdf (参照日 2023年9月2日)
- 4) Tyna Eloundou, Sam Manning, Pamela Mishkin, et al. GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. arXiv, 2023
<https://arxiv.org/pdf/2303.10130> (参照日 2023年9月7日)
- 5) 文部科学省. 大学・高専における生成 AI の教学面の取扱いについて. 2023.
https://www.mext.go.jp/content/20230714-mxt_senmon01-000030762_1.pdf (参照日 2023年9月2日)
- 6) 文部科学省. 初等中等教育段階における生成 AI の利用に関する暫定的なガイドライン. 2023.
https://www.mext.go.jp/content/20230710-mxt_shuukyoku02-000030823_003.pdf (参照日 2023年9月2日)
- 7) 藤村裕一. 生成 AI の教育利用に関する研究. 日本教育工学会研究報告集 2023 (2), pp.75-82, 2023
- 8) 東京都教育委員会. 令和5年度使用都立高等学校及び都立中等教育学校 (後期課程) 用教科書教科別採択結果 (教科書別学校数). p.8, 2022
https://www.kyoiku.metro.tokyo.lg.jp/school/textbook/adoption_policy_other/adoption_result/files/result_2023_h/05saitaku.pdf (参照日 2023年9月5日)
- 9) 阿原一志, 市原一裕, 岡部恒治, 他. 新編数学 I. 数研出版, pp.170-198, 2023.
- 10) OpenAI. Introducing ChatGPT. <https://openai.com/blog/chatgpt> (参照日 2023年12月21日)
- 11) Microsoft. The new Bing: Our approach to Responsible AI. <https://support.microsoft.com/en-us/topic/the-new-bing-our-approach-to-responsible-ai-45b5eae8-7466-43e1-ae98-b48f8ff8fd44#:~:text=Introduction%20In%20February%202023%2C%20Microsoft%20launched%20the%20new,jokes%2C%20stories%2C%20and%2C%20with%20Bing%20Image%20Creator%2C%20images.> (参照日 2023年12月21日)
- 12) Google. An overview of Bard: an early experiment with generative AI.
<https://ai.google/static/documents/google-about-bard.pdf> (参照日 2023年12月21日)
- 13) 西田宗千佳. 生成 AI の核心: 「新しい知」といかに向き合うか. p.83, NHK 出版, 2023.

公開 2024年4月1日

<連絡先>

市川隆司

宛先 〒536-8585 大阪府大阪市城東区古市2丁目7番30号

電話番号 06-6939-4391 (代表)

E-mail t-ichi@osaka-shinai.ac.jp