

2024年度 統計データ分析コンペティション

## 審査員奨励賞 [高校生の部]

学力と外見への投資に関する回帰モデルを用いた分析

倉本 佳詩野 (宮崎県立五ヶ瀬中等教育学校)

# 学力と外見への投資に関する回帰モデルを用いた分析

倉本佳詩野

宮崎県立五ヶ瀬中等教育学校

## 1. 研究のテーマと目的

### 1.1 背景

学習への影響要因については様々な研究がなされているが、筆者の個人的関心である「ファッション・美容」の分野との関連を研究した例は多く見られない。一方、筆者の周りの素朴な会話からは、基本的に「勉強と外見」に関して負の相関があるはずだという論調のものを多く耳にする印象がある。しかし、「外見や身なりを整えることは集中力を高め、結果的に学力を高め得る」という見方にも、一定の妥当性を感じる。そこで、これらを統計的に分析することにより、学習に影響する要因の発見に繋がるのではないかと考えた。また、学習能力に影響を与える他の項目項目を同時に回帰分析の説明変数とすることで、先行研究と本研究の効果量を相対的に考察できるのではないかと考えた。

### 1.2 目的

本研究では外見に対する投資が学習能力に影響するのか回帰モデルを用いて探索的分析をおこなう。学習における外見への投資を科学的に明らかにし、外見が学習に与える新たな知見を得ることを目指す。そして、自己投資以外に学習に影響する可能性のあるものを分析し、より学習能力を高める知見の発見を目指す。

## 2. 研究方法

①データの収集:SSDSE(教育用標準データセット)などのオープンデータをもとにファッションや化粧品の購入金額と他の学力に影響する可能性のあるデータを選ぶ。そして、本研究での外見への投資は、「被服履物購入率」「化粧品購入率」のデータを用いて分析することにする。また、他の学力に影響する要因は先行研究をもとに特定する。

②モデルの変数選択:目的変数を学力、説明変数を外見への投資と学習に影響する要因として、回帰モデルを使用し、ステップワイズ変数増減法を適用し、最終的に分析対象とするモデルを選ぶ。

③最終モデルの分析:②の分析で選ばれたモデルの当てはめ値と観測値の一致度を可視化し、モデル性能を評価する。

## 3. データセットの加工

### 3.1 外見の投資以外の学習に影響すると考えられる要因の選定

①読書:アンドリュー(2013)<sup>(1)</sup>の研究より、「朝の読書や休憩時間での自由読書は子どもたちを落ち着かせ学習にふさわしい姿勢を促進することにつながる」と論じている。よって、読書は学力の向上に影響するのではないかと考えた。今回の研究では、全国の高校の朝読書率をデータとして使用した。

②収入：耳塚(2014)<sup>(2)</sup>の研究により、世帯収入が高いほど子どもの学力が高いと論じている。よって、世帯収入は学力の向上に影響するのではないかと考えた。

③スマートフォン利用：増田・成田(2018)<sup>(3)</sup>の研究より、「スマホやゲーム利用が1時間未満までは、勉強時間と成績は正比例しているが、1時間を超すと顕著に成績が低下する」と論じている。よって、踏まえて、スマホやゲームの利用率が高ければ、学力も低いのではないかと考えた。今回の研究では、スマホの利用率のみを使用した。

④教育費・教養娯楽費：耳塚(2014)<sup>(2)</sup>の研究により、「学校外教育支出と学力との関係は強く、学校外支出が多い家庭ほど子どもの学力も高い。世帯収入が高くなるにつれ学校外教育支出も多くなる傾向がある」と論じている。よって、教育支出と学校外教育支出は学力の向上に影響があるのではないかと考えた。今回の研究では、学校外教育支出は教養娯楽費とする。

⑤運動：生駒(2011)<sup>(4)</sup>の研究により、「体力が学力に対して正の統計的予測力を持つことが示された」と論じている。よって、体力は学力向上に影響するのではないかと考えた。今回の研究では、中学生の男女の体力テストの合計点の平均を使用した。

⑥ストレス・幸福感：ストレスと幸福感の2項目は、先行研究を筆者が調べた限り見つからなかった。しかし、ストレスと幸福感は学力に影響すると考えられるため、学力に影響する要因として使用する。

### 3.2 使用するデータとデータの加工

本分析では、「学力」を目的変数とし、下表で選定したデータを目的変数とした。SSDSEからいくつかデータを選定し、SSDSEで得ることができなかったデータについては、外部のオープンデータを使用した。Education Expensesのみ正規分布に近づける目的で対数変換を行った。また、回帰分析前に全データを標準化した。

変数名	年度	出典
Education Expenses	2021	SSDSE-A
Culture and Entertainment Expenses	2021	SSDSE-B
Clothing and Footwear	2021	SSDSE-B
Income	2019	SSDSE-E
Cosmetics	2023	e-Stat
Smartphone Usa	2019	総務省
Physical Fitness	2023	文部科学省
Reading	2024	株式会社トーハン
Stress	2021	一般社団法人日本リハビリ協会
Happiness	2023	一般社団法人スマートシティ・インスティテュート
Academic ability	2022	データネット実行委員会

図1.分析に使用する説明変数と目的変数

## 4. データ分析

### 4.1 モデルの変数選択

応答変数の誤差構造には正規分布を仮定する。説明変数は、フルモデルおよびヌルモデルを起点としたステップワイズ変数増減法を2通り試し、最終的に解釈するモデルを決める。

#### 4.1.1 フルモデルを起点とした変数増減法

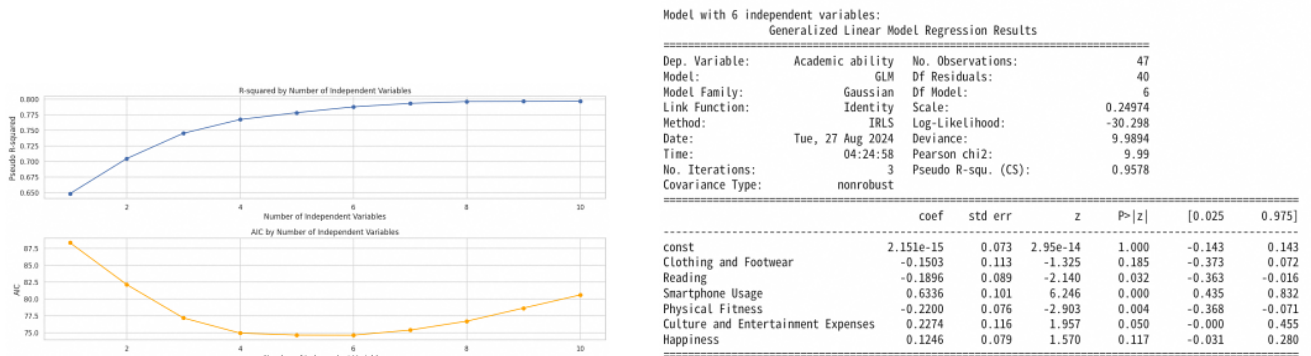


図2.フルモデルの分析結果

AICの観点では、4~6変数モデルが最適だと示されており、4~6変数モデルが構成しているデータ項目を見てみると、6変数モデルに、Clothing and Footwear が入っていた。

#### 4.1.2 ヌルモデルを起点とした変数増減法

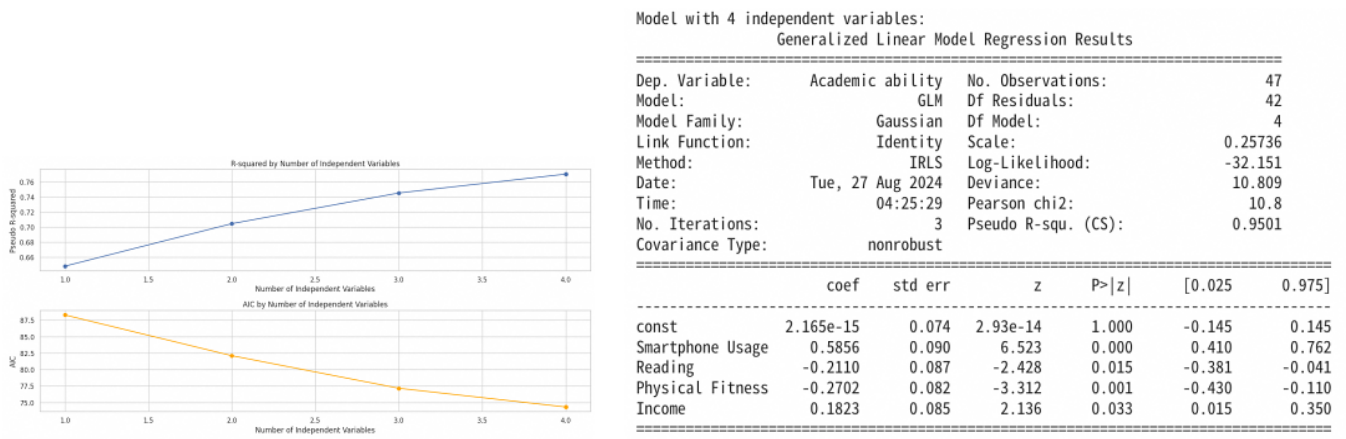


図3.ヌルモデルの分析結果

AICの観点では、Smartphone Usage、Reading、Physical Fitness、Income の4変数モデルが最適だと示されている。

ステップワイズ変数減少法とステップワイズ変数増減法により6、4変数モデルが選出された。しかし、後の分析でRMSEを算出したところ、6変数モデルは0.46、4変数モデルは0.48とモデル性能にほとんど差はなかった。よって、過度に複雑なモデルを採用すべきではないという考えから、4変数モデルをベストモデルとした。

#### 4.2 ベストモデルの分析

ステップワイズ変数減少法とステップワイズ変数増減法を適用し、ベストモデルとして選定した、4変数モデルのSmartphone Usage、Reading、Physical Fitness、Incomeの分析を行う。当てはめ値と観測値の一致度と、効果量を可視化する。

##### 4.2.1 4変数モデルの当てはめ値と観測値の一致度可視化

実際のデータと、モデルの予測が一致しているかを視覚的に確認するために図を作成した。

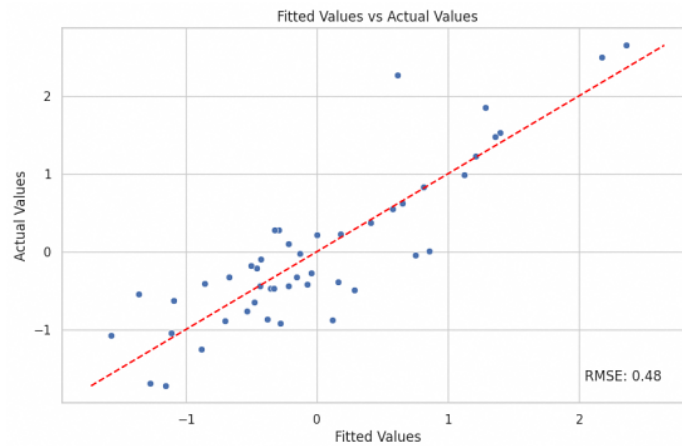


図4.当てはめ値と観測値の一致度可視化

##### 4.2.3 4変数モデルの偏相関係数

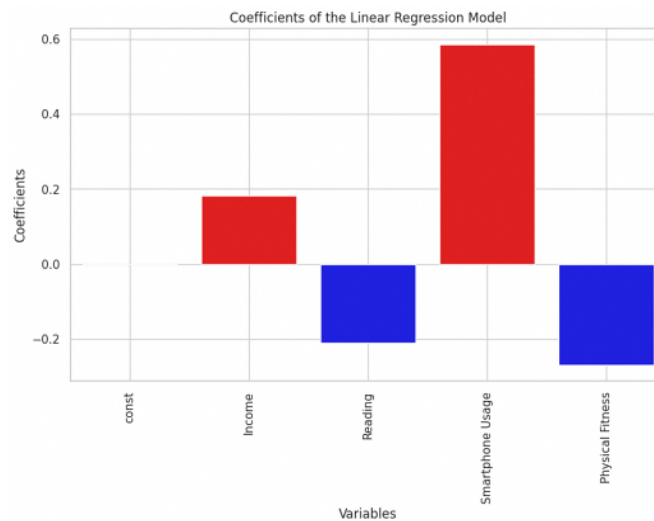


図5.偏相関係数結果

4変数モデルでは、Reading、Physical Fitnessは負の効果を示し、Smartphone Usage、Incomeは正の効果を示した。

## 5 結果の解釈

### 5.3.1 モデルの効果量の解釈

ファッションや化粧品などの購入率は統計的有意性がなかったため、今回の研究の「整った服や身なりになることで、モチベーションが上がり、学力の向上につながっているのではないか」という仮説は立証されなかった。しかし、選出したデータの妥当性に根本的な原因がある可能性は十分考えられる。すなわち、今回用いたファッションや化粧品の購入金額のデータは、主におしゃれを目的とした投資のデータであり、外見や身なりを整える服飾品との区別はできない。

Smartphone Usage、Physical Fitness、Reading、Incomeについて、効果量の大きい順に解釈を行う。  
Smartphone Usage: スマホでのインターネット利用率は正の効果を示された。参考文献ではスマホの過度な使用率は学力に悪影響を及ぼすと論じていたが、今回は都道府県単位のスマホ利用時間データを探ることができなかつたため、スマホの利用時間と学力の影響を示すことはできなかつた。スマホ利用率と学力の部分で考えられる解釈では、スマホを学習ツールとして使用することは学習効果を高める可能性がある。

Physical Fitness: 体力は負の効果を示された。今回使用したデータは高校生の体力のデータが見つからなかつたため、中学生の体力のデータを使用した。高校生は受験勉強などで運動に割く時間が減る可能性があることから、参考文献と異なつた結果となつた可能性がある。

Reading: 朝読書実施率は負の効果を示された。先行研究では読書は学習に対しての姿勢に影響があると論じられていたが、今回の研究では学習姿勢と読書実施率の分析は行えていない。そして高校の朝読書などは学校でルール化され、時間を限定されているもののため、制限された読書は学力に負の効果をもたらすのかもしれない。

Income: 世帯収入は正の効果を示された。これは参考文献を支持する結果となつた。世帯収入が高い家庭の方が教材費用や塾代など学習に関係する費用にお金をかけるからだろうと考える。

### 5.3.2 本研究をふまえた提案

今回の研究では、スマホの使用率が学力に正の影響を与える結果が得られた。しかし、増田・成田(2018)<sup>③</sup>の研究では、「スマホやゲームの利用が1時間未満では勉強時間と成績は正比例しているが、1時間を超えると成績が顕著に低下する」と論じている。これに基づくと、スマホの利用時間や学習以外での利用方法によっては、学力を低下させる可能性があることが分かる。今回使用したスマホの利用率データには、ゲームの利用や学習アプリの使用といった項目の区別がなく、スマホの利用内容が複数含まれているため、今回の研究結果から「スマホの使用が学力の向上に影響を与える」とは言い切れないだろう。しかし、先行研究と今回の分析結果が異なっていることから、スマホ利用の全てが学力に対して負の影響を与えるわけではないと考えられるだろう。

### 5.3.3 今後の展望

今回の分析では外見への投資と学力には統計的有意性がなかった。しかし、外見への投資によってモチベーションが上がり、学力の向上に繋がることは分析されていない。そのため今後の分析では外見の投資によるモチベーションから学力の向上についての分析を行いたい。学力に影響する要因の中で参考文献と一部違つた結果が出た。今後の研究では参考文献の研究環境と可能な限り近い条件で再分析を行う必要がある。

## 参考文献

(1)デュアーアンドリュウ. (2014). 読書が子どもの発達に及ぼす影響. 東海学院大学紀要, 7, 261-267.

(2)耳塚寛明. (2014). 文部科学省委託研究「平成 25 年度全国学力・学習状況調査(きめ細かい調査)の結果を活用した学力に影響を与える要因分析に関する調査研究」.

(3)増田修治, 成田弘子. (2018). 子どもの発達・学力と健康に視点を置いたメディアリテラシー教育指導法の研究. 研究年報, 23, 43.

(4)生駒忍. "体力は経済力とは無関係に学力と相関する." チャイルドサイエンス: 子ども学 7 (2011):54-57.

家計調査 / 家計収支編 二人以上の世帯 詳細結果表

[https://www.e-stat.go.jp/stat-search/files?stat\\_infid=000040140558](https://www.e-stat.go.jp/stat-search/files?stat_infid=000040140558) 2024年8月10日

「朝の読書」全国都道府県別実施校数

[https://www.tohan.jp/wp/wp-content/themes/tohan/pdf/asadoku\\_school.pdf](https://www.tohan.jp/wp/wp-content/themes/tohan/pdf/asadoku_school.pdf) 2024年8月10日

総務省 | 令和2年版 情報通信白書 | インターネットの利用状況

<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r02/html/nd252120.html> 2024年8月6日

都道府県別集計(小学校児童・中学校生徒に対する調査結果)

[https://www.mext.go.jp/sports/content/20231218-spt\\_sseisaku02-000032954\\_208.pdf](https://www.mext.go.jp/sports/content/20231218-spt_sseisaku02-000032954_208.pdf) 2024年8月10日

ストレスオフ県ランキング2021発表！ コロナ禍でもストレスオフだった第1位は「茨城県」 | 一般社団法人日本リカバリー協会

<https://www.recovery.or.jp/research/1094/> 2024年8月17日

地域幸福度(Well-Being)指標 令和5年度全国調査結果

[https://www.digital.go.jp/assets/contents/node/basic\\_page/field\\_ref\\_resources/1eb514b0-5e04-4b40-9cac-e142439bad12/07cac8c5/20230719\\_meeting\\_digital-garden-city-nation-wellbeing\\_agenda\\_outline\\_02.pdf](https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/1eb514b0-5e04-4b40-9cac-e142439bad12/07cac8c5/20230719_meeting_digital-garden-city-nation-wellbeing_agenda_outline_02.pdf) 2024年8月10日

データネット概況<https://dn-sundai.benesse.ne.jp/dn/dn2022/doukou/dnbook/dl/2022-dnbook-gaikyo.pdf>  
2024年7月21日