

医学教育における探索的データ分析の導入

中山 伸朗（浜松医科大学）

1. 本発表の目的と課題

日本国内の医学部、医科大学の多くは、この数年、世界医学連盟（WFME）の基準に適合することを目的に、IR部門を設置しカリキュラム評価の一端を担わせた。しかし、各大学のIR担当者から事情を聞く限りで、教員と事務職員の数名で構成されることが多いようで、教員は病棟や外来での業務を担っていたり、事務職員はデータや統計学等に必ずしも明るくなかったりと、基盤はいまだ脆弱だと推察される。

加えて、基盤を脆弱にさせている要因として考えられるのが、医学教育、臨床医養成教育における特殊な知識とその量である。例えば医学教育モデル・コア・カリキュラムや、初期臨床研修、専門医研修などの知識がなければ問題点だと気づけないことも多い。これはデータサイエンスでいう、いわゆる「ドメイン知識」に相当するもので、これが欠けると推論を立てることも難しくなってしまう。

この状況から、筆者は「経験から仮説を得て分析を進める」というアプローチから、「データから特徴や仮説を見だし、問題の糸口を探る」という手法を模索してきた。医局や事務局による定期的な人事異動が組み込まれている限り、ドメイン知識とデータ分析の両方を持つことは難しい。データから特徴を見だし、その気づきを現場の教職員に投げかける方が効果的だと考えたためである。

試行錯誤の結果、現在、筆者は探索的データ分析のツールである Exploratory を用いて仕事を進めるようになった。今回はそれを用いた仕事の進め方を紹介したい。なお、本稿は研究成果の発表ではなく、あくまで事例紹介であることにご留意願いたい。

2. Exploratory の機能と使用方法。

Exploratory とは統計に用いられる R 言語をエンジンとしたソフトウェアであり、日本語では“探索的”と訳される。基本的に GUI を使うため、jamovi や Tableau に似た部分もあるが、R と関連パッケージを組み込んでいるため、分析という面で秀でている。以下、このツールを使用した分析のプロセスを、1) 入学者選抜、2) 学業成績、3) 医師国家試験の3点を取り上げ紹介したい。

1) 入学者選抜

データをインポートすると、まず全体の分布が表示される（図1）。ここでは変数同士での相関や95%信頼区間などをまとめて表示させることも出来るため、自ずと「全体の傾向を見てから個々の分析へ」という流れになる。医学部では、性別や出身地などのデモグラフィックデータが重要な変数となりうることから注意する。この時点で何か特徴や傾向が見られれば、平均の差の検定や多重比較を行い、有意差を含め確認する。

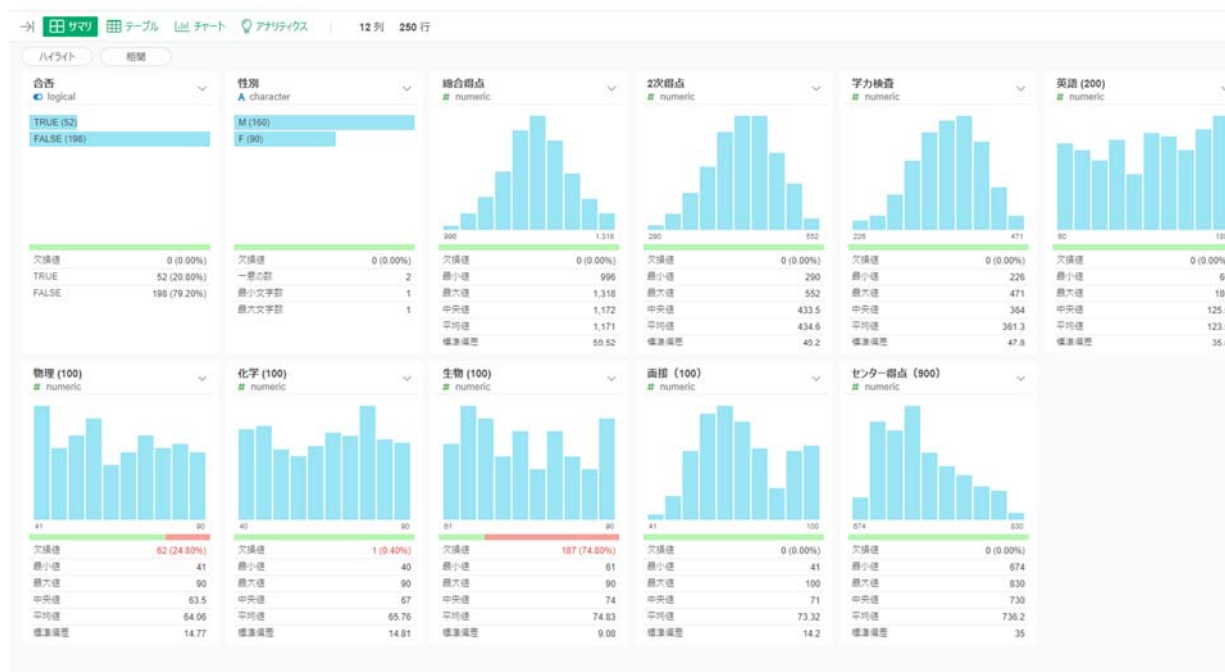


図1 Exploratoryのサマリビュー（数値はダミー生成）

次に過去のデータと比較し、大きな変動がないか確認をする。もし、傾向に変化が見られた場合、現場の教職員にデータから得た気づきを伝える。例えば、大きな得点差が見られれば「センター試験の難易度はどうだったのか?」、志願者数で増減が見られたら「競合他大学で選抜方法を変えたところはあったか?」などの質問を現場の教職員に投げかける。そこで何かしらのフィードバックを得られたら、追加の分析や事実確認を経て、会議等での報告に加える。

2) 学業成績

医学部の学業成績において、最も重要といえるものの一つは臨床実習開始前の共用試験 CBT である。共用試験 CBT について触れると、これは項目反応理論 (Item Response Theory, IRT) に基づいて実施される Computer Based Testing で、基礎・臨床・社会医学の知識が身についているか測るものである。この試験は全国実施結果の過去データから、問題の難易度を推定し、基準集団の分布を用いて評価を行うところに特徴がある。具体的には、2012年から2014年にかけて実施した過去データから基準集団を推定して、平均値 500 (SD100)、つまり標準正規分布に従うように難易度をセットした問題が出題される。出されるスコアは分布の中での位置を意味し、スコアの実質的な見方は学力偏差値と同様になる。しかし、成績基準集団は固定されているため、学生の学力が上がれば、分布のカーブは右側にシフトする。過去の学生と同じ基準で能力を測定できるため、この傾向を見ることは極めて重要である。

この共用試験 CBT の分析は、密度曲線を用いてスコアの傾向を見ることである。Exploratory では、チャートから密度曲線を作成することができる (図 2)。このグラフは Microsoft Excel でも、関数を用いて作成することが出来るが、歪みのないきれいな左右対

称の正規分布として表されるため結果の読み取りには注意を伴う。グラフから平均値が落ち込みや、分布に二極化が見られれば、関係者へフィードバックすべきである。

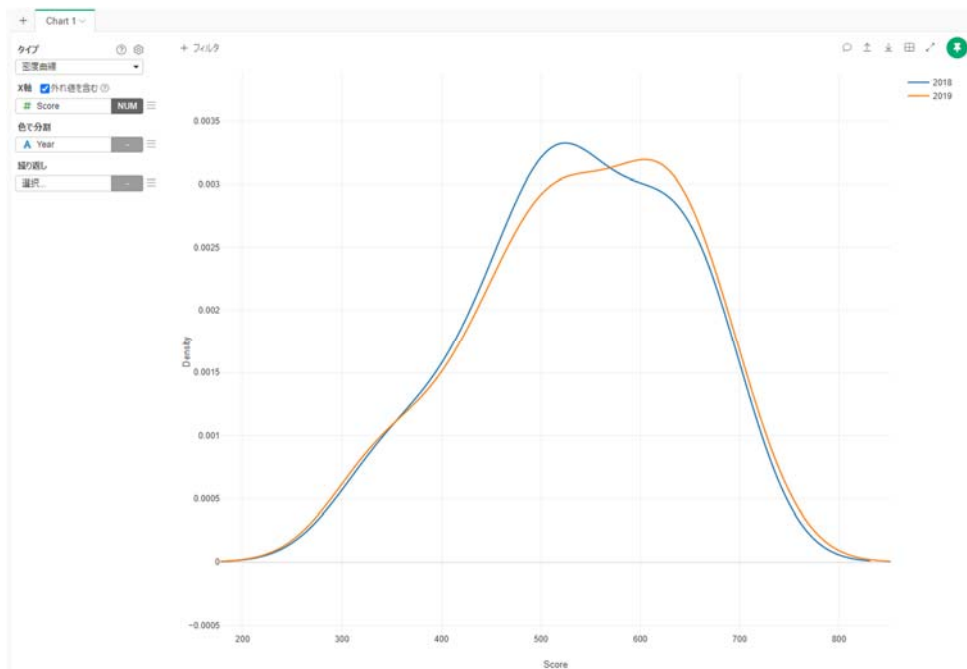


図2 Exploratory のチャートビュー（数値はダミー生成）

3) 医師国家試験

最後に、学修成果の指標ともいえる医師国家試験の分析について述べたい。先行研究というより、すでに通説になっていると考えているが、医師国家試験の成績と共用試験 CBT スコアには強い相関があると報告されている。国家試験の自己採点結果と共用試験 CBT の2種類の連続値データがあれば、回帰分析により不合格者の予測を試みることができるが、実際には合否以上の情報を得ることは難しい。そのため共用試験 CBT 以外で合否に相関する変数を見いだすことができれば、それと組み合わせ、分析モデルの作成が期待できる。

このプロセスでは機械学習のアルゴリズムであるランダムフォレストを使用する。アルゴリズムの詳細は割愛するが、簡単に言えば大量に生成した決定木を用いて変数の重要度を見るものである。説明変数には量的変数、質的変数の両方が使用できるため、ある学年や時期の GPA や、試験の点数、入試区分などのデータを用いることが可能である。

図3は Exploratory のランダムフォレストを使って、変数重要度を調べたものである。この機能であるが、2020年9月現在配布されている Version6.1 ではボルータという手法が実装されており、変数重要度を仮説検定の結果とともに箱ひげ図で表すことができるようになっている。医師国家試験の分析では、まず目的変数に合否（論理値）を説明変数に各種データを取り込み、関連の高い変数を絞り込む。そして、説明変数に、変数重要度が高いと示されたデータと共用試験 CBT のスコアを、目的変数に医師国家試験の合否を設定すると、二項ロジスティック回帰分析を用いて、不合格者を予測するためのモデルを作成することができる。

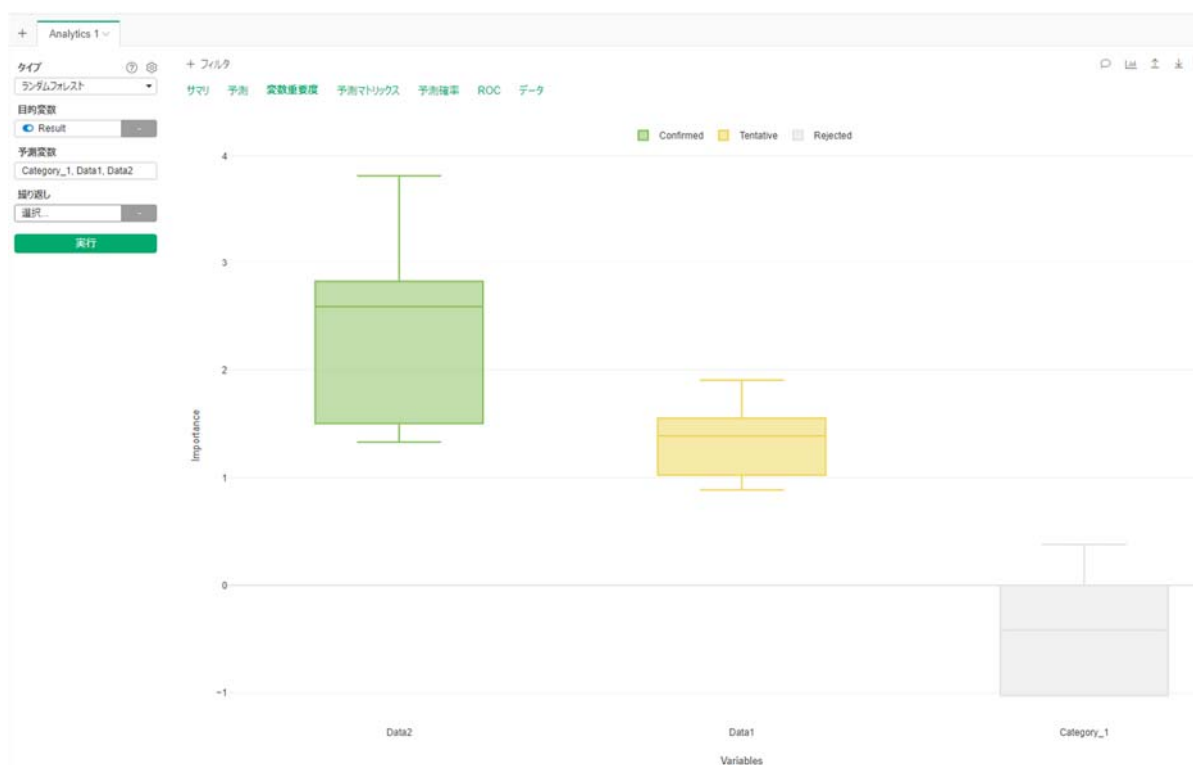


図3 Exploratory のアナリティクスビュー（数値はダミー生成）

3. まとめ

以上、簡単ではあるが具体例を3つ取り上げ、Exploratoryを用いた仕事の進め方を紹介した。現在の大学教育では、教学マネジメント改革、高大接続改革、何より新型コロナウイルスへの対応など、IRは現状を把握するだけで手一杯なのではないだろうか。また、データ分析についても、過去との繋がりが途切れる、つまりパラメータをゼロベースで考えざるを得ないという状況は、精神的にも厳しいものと容易に予想される。「データから特徴や仮説を得る」、探索的データ分析というアプローチは今後のIRにおいて一層重要になっていくものと考えられる。

【参考文献】

- [1] Exploratory <https://exploratory.io/>
- [2] 2019年（令和元年）度第17版臨床実習開始前の「共用試験」P21
- [3] 中谷晴昭, 日本の医学教育の現状と医師国家試験, 医学教育 2015, 46 (1): pp. 14-17
- [4] どの変数が最も関係しているのか調べるときに便利なボルータの紹介
<https://exploratory.io/note/GMq1Qom5tS/mAb9gDB1XX>