

白旗論文に対するコメント

東京大学 繁樹 算男

1. 問題

大学入試データの解析は、入学者の選抜方針決定に関連するので、慎重になされなければならないのはもちろんである。そして、大学入試データは、無作為サンプリングや無作為割り当てによって得られた「行儀のよい」データであるとは限らない。したがって、確立された方法を機械的に応用すると誤解を招くことがあるし、分析の目的にかなうように新しい方法を工夫することが必要なときもある。

たとえば、入試選抜資料の妥当性を確かめるために、大学での成績との相関を取る場合、その相関を計算する本来の対象（母集団）を定めておく必要がある。その母集団は、各種の選抜資料のそれぞれの妥当性を比較する際には、同一でなければならない。仮に、大学の入学試験成績が直接合否に反映し、高校の内申書が間接的にしか大学の合否に影響しないとしよう。このとき、受験者全体を母集団とした場合、入学試験成績と高校の内申書とが、同等の妥当性を持つならば、実際にデータが得られる集団（大学入学者、母集団の一部）から計算される相関係数は、高校内申書の相関の方が高くなる。このような統計的性質を知らないと、計算された相関係数の基づき、誤った方針をたてる可能性がある。この誤解は、今ではかなり知られるようになったが、大学入試関係の会議でもこのような事実を知らず、ある特定の選抜資料を推薦する根拠にする議論がある。

本論文で扱うのは、入学試験の各項目の性能を把握するための設問別解答率分析図という分析法の持つ欠陥の指摘と、その対策法についての論文である。この方法は、総得点によって受験者をいくつかのグループに分類し（たとえば、

最上位群、上位群、中位群、下位群、最下位群）、その各々のグループにおける正答率を表示する方法である。総得点の値が小さいと、得点のそれぞれを条件とする正答率を表示してもよい。この方法は、各試験項目がどの程度よいか、試験項目の性能を評価するためのよい工夫であると思われる。しかし、この方法には次のような問題が内在している。総得点には、その性能を評価しようとする項目の得点も当然含まれており、そのために、本当の関係が見えにくくなっている。この分析図に基づいて、項目の良し悪しを判断することもあり、この欠点を是正することは重要な意味を持つ。

2. 提案されている方法の有効性

本論文では、そのバイアスを除去し、本来の関係を修復するための方法を提案している。すなわち、総得点を所与とする条件付の解答率の代わりに、当該の項目の得点を除いて、総得点を計算し、その得点を所与とする正答率の方法である。

着眼点が良く、意味的にも分かりやすい。入試関係のデータ分析には、このような鋭い吟味が必要であり、統計学の基本的な道具立てによって、たいていの問題点は解決できることの好例である。

ただし、従来、入試センター試験の分析に使われる場合には、項目の数が本論文のシミュレーション例の5項目などよりはるかに多く、解答率分析図の弊害もそう大きくはなかったと思われる。項目数が多い場合の実際に近いテストの分析図のシミュレーションもチェックしたいところである。また、総得点を所与とした分析図は、“真の”関係は、隠されるかもしれないが、相対的な意味での項目の性能評価には、項

目数が少ない場合でも使うことができると思われる。

3. モデル準拠による項目の分析

バイアスのかからない項目の分析を可能にするもう一つの方法は、テスト得点の発生機序をモデル化し、真の得点 θ を所与とする正答率の値を推定する方法である。(テストの得点を解析するためのテスト理論では、項目特性曲線と呼ばれる。)この方法の利点は、 θ は、解析に使われる項目や被験者から独立した解釈が可能である点である。(したがって、テスト得点の総得点からも独立である。)欠点は、この方法は、モデルに依拠しているのだから、モデルが妥当ではないときには、結論も妥当ではない。しかし、通常のテスト得点のモデルは、それほど縛りの強いものではない。本論文では、シミュレーションにおいて、モデルを使っている。一見、モデルを使っていないように見えるかもしれないが、たとえば、 θ の変域が、0から1の間で、条件付正答率が θ であるモデルは、 $\eta = \log|\theta / (1 - \theta)|$ という変換(ロジスティック変換)を施し、

変域を $-\infty$ から、 $+\infty$ までにすると、条件付正答率はロジスティック関数となる。モデルといってもこの程度であり、パラメータの値によって十分に柔軟にいろいろな形状を表現できる。モデルがデータに適合するかどうかは常に意識していなければならないが、モデルが妥当であることを確かめた上で、モデルを使った分析が、使われた項目や被験者を超えて一般化することができる利点を行かすことに意味があると考えられる。モデルが正しい場合には、テストの等化や、テスト編成の最適化等のために、理論的な指針を提供できる。

もっとも、人工的な概念 θ の導入は、専門家以外の関係者への説得性を弱めるかも知れず、そのような意味では、本論文のように、データだけの項目の特徴を明らかにする努力は貴重である。モデル依存の推論をすると同時に、データの記述的な方法とを両者を行い、二つのアプローチによる結果を相互に参照することがモデル依存の推論や指針を裏付けることに有益であろう。