

統計検定 CBT の試験問題評価とその課題

桜井 裕仁 (大学入試センター)

○ 林 篤裕 (名古屋工業大学)

本多 正幸 (千葉大学)

1. 背景

一般社団法人 日本統計学会が認定し、一般財団法人 統計質保証推進協会が実施している統計検定は、2023 年 2 月現在、1 級からデータサイエンスエキスパートまで、10 種類の検定試験が取り揃えられており(日本統計学会, 2023)、この中で 1 級以外はすべて CBT 方式で実施されている。

統計検定の作成・運営に関わっている各級の分科会では、それぞれで検討を行っている以外に、合同で定期的に評価・検討分析に関する会合を開催し分科会相互での情報交換も行っている。これら各級の分科会とは独立に CBT 問題評価分科会が設置されており、CBT での運用に関する検討や支援が行われてきた。本発表では CBT 問題評価分科会の活動について紹介し、今後の課題を述べる。

2. 現在整備中の作業内容

広辞苑第七版(岩波書店, 2018)によると、資格試験とは「資格付与の際、それにふさわしい力量を持っているかどうかを調べる試験」とあり、資格試験全般に言えることであるが、ふさわしい力量を有しているかどうかを判定するためには、事前に定められた出題範囲の中から分野を偏りなく出題する必要がある。例えば、2 級の出題範囲は統計検定の Web ページ(2018)や公式問題集(日本統計学会, 2023)に掲載されており、これに沿って出題される。また、受検者の勉強度合い(能力)も広範に亘るため、易しい基礎的な試験問題(以下、設問)から難しいものまでいろいろな難易度の設問を準備する必要もある。

出題される設問は分科会で幾多の検討を慎重に重ねて完成される。試験を実施した結果として得られる受検者の解答状況から、設問を評価する指標は幾つか存在し、代表的なものとして「難易度」と「識別力」が挙げられる。設問の難易度は、一般的には受検者全員を対象とした正答率(平均点)で表すことが多いが、これら 2 つの指標は、受検者の勉強度合いによって正答率が変動することに注目して評価している。

また、設問の評価は視覚的に行われることがあり、例えば、設問解答率分析図と呼ばれる折れ線グラフは次のように作成される。解答した設問全体の成績(いわゆる総合点)を基準に、成績の低い者から順に並べて受検者数になるべく均等になるように 5 群(場合によっては 3 群)のグループに分け、それぞれのグループごとの正答率を計算する。横軸にこのグループを取り、縦軸にグループごとの正答率を取ってこれら 5 点を 4 本の線分で結ぶと、出題した設問の数だけ折れ線グラフが生成される。

それぞれの設問の正答選択肢に対応する折れ線グラフを観ると、線分の上下方向の位置関係から受検者の勉強度合いに応じた難易度を推測することが出来る。また、この線分の傾きは同一の傾斜角で上昇していくことは少ない。もし、隣り合う群の正答率が極端に異なっていれば、線分の傾斜角が大きくなり、左右の群で分離・識別できていることになるため、設問の特性を視覚的に把握することが出来る。なお、横軸に総合点に基づくグループを取っていることから、この折れ線グラフは基本的には単調増加をするはずであるが、設問によってはそうはなっていない場合もあり、その原因を突き止めることによって作題工程を見直す動機付けにもなる。この図に誤答

選択肢に対応する折れ線グラフも併せて描き加えると、受検者がどの選択肢を魅力的に感じていたかも判る。

加えて、CBT 方式で実施するには出題分野や難易度が偏らないように予め設問を組み合わせた複数のセットを準備しておく必要がある。その際のテクニックとして、出題範囲をいくつかのサブグループに分割してサブグループごとに管理しながら組み合わせる方策が取られている。よって、そのサブグループ内での組み合わせの吟味やサブグループ間での難易度や識別力の差異も評価対象となる。

なお、ここまでは設問を中心に説明を行ってきたが、試験問題は設問が幾つかずつにまとまって大問を構成していることから、設問単位だけでなく、大問単位でも同様に評価する必要がある。また、CBT 方式の場合、受検者群に依って提示される設問や大問が異なるため、受検者群ごとに設問と大問の評価を行う必要もある。このため、大問間の相関や得点分布等に関する指標も算出している。

ここまで説明してきたように、得られた受検者の解答状況を、適宜小セットに分割・結合しながら分析を進める必要があるため、一つの級を取り上げるだけでも非常に緻密に作業を進める必要があり、上記の作業を各級の分科会が定期的に行うには多大な労力が必要となる。そこで、CBT 問題評価分科会では分析結果をシステムチックに提示できるように計画を立てているところである。

3. 今後の課題

前項で述べた各種の分析結果を安定的に提供できるようになると、各級の分科会で行っている分析作業の支援につながるため、まずは評価体制の確立が喫緊の課題と考えている。また、各級の分科会においてどのように慎重・丁寧に作題作業を進めたとしても出題分野のサブグループ内ではばらつきは生じてしまうので、その許容性も検討の必要がある。加えて、CBT 方式を採用していることの宿命とも言えるが、それぞれの設問・大問が複数回の試験で提示されることから、ライフサイクルを考えると、ある程度曝露された設問・大問は廃棄して新しいものと入れ替えるフェーズも必要になる。これらの見極め方の確立も必要であり、順に研究を進めていく予定である。

何れにしても、CBT 方式を導入した統計検定を継続的・安定的に実施していくための評価資料の整備に努め、効率的な検討環境を整えて各級の分科会の支援につなげられればと考えている。

参考文献

- 1) 日本統計学会編(2023), 日本統計学会公式認定 統計検定 2 級 公式問題集 [CBT 対応版], 実務教育出版.
- 2) 統計検定(2018), 統計検定 2 級出題範囲表, <https://www.toukei-kentei.jp/exam/grade2/> (2023 年 2 月 7 日現在)
- 3) 新村 出編(2018), 広辞苑 第七版, 岩波書店.

【注】 予稿原稿提出後の 2023 年 2 月 28 日に参考文献を修正