研究紀要論文抄録

試験問題統計情報のデータベース化と利用環境の整備

大津　起夫[i]　石岡　恒憲[i]　橋本　貴充[i]

大学入試センターによる全国統一の大学入学試験は，昭和54（1979）年に共通第１次学力試験が実施されて以来，平成21（2009）年で31回目の実施を迎えた。平成２（1990）年度以降は制度変更されて大学入試センター試験となり，科目選択の自由度が増えた。制度変更以降は，アラカルト方式の科目選択が可能であるため，受験者数は科目により異なる。少なくとも１科目以上を受験した者の数は，平成21年では約50万８千人であった。

大学入試センター研究開発部では，これまでに毎年度，受験者の個人情報を秘匿した成績データを作成し研究に利用してきた。しかしながら，データの量が膨大であり，また過去の業務処理が古典的な大型計算機システムによって行われてきたこともあり，これらのデータは特別な電子的アノーテーションのないテキストファイルとして保存されている。このためファイル中に含まれる情報を取り出すためには，紙媒体のファイルフォーマット資料を参照する必要があり，設問正答率等の統計情報を導出するのはかなり煩雑な作業となる。

本研究では，昭和54年から平成20年に至るすべての本試験科目について，成績データから平均点，項目別得点率等の統計情報を抽出し，さらにこれらの統計情報を電子的に組織化することにより容易な参照を可能とした。過去の研究開発部の共同研究においては，一部の年度について統計情報を整備し，PC上のデスクトップデータベースを用いて閲覧可能とした。これらの成果を受け継ぎつつも，本研究においては以下の新たな方針をとった。

１）試験問題作成における利用を想定すると，掲載情報を少数の重要な指標に絞ることが実用的であると判断

i　大学入試センター 研究開発部試験評価解析研究部門
本論文は大学入試センター研究開発部におけるプロジェクト研究Ⅳ「センター試験に関わる統計情報の蓄積と利用に関する研究」（平成18年度～平成20年度）の成果に基づく。

された。このため整備する統計情報を精選し，情報の単純な呈示方法を考案した。

2）最近の大学入試センター試験問題の採点においては，部分点の付加はほとんど行われていないが，過去の資料を検討すると部分点の付与が行われている場合がしばしば生じている。これらの多様な採点方法に対応し得るよう採点集計システムを設計した。統計情報の導出という目的に限定するならば，ごく少数の部分点の付与を正確に再現することにはさして重要性はないが，配点規則の正確な復元のためには，すべてのケースにおいて正しく採点が行われることが望ましい。このため，部分点の付与を忠実に再現する労力を払った。

3）統計情報の利用と更新を考慮すると，個別のPCにシステムを配置することは管理上煩雑であり，ローカルネットワークを介し，サーバコンピュータ上の資料をウェブブラウザを用いて閲覧することが望ましい。利用者PC上にそれぞれデータをインストールして利用を行うことは不可能ではないが，データの更新を行うことが煩雑になり，運用上の困難をもたらす。このため統計情報はサーバで一括管理し，リモートPCから閲覧を可能にした。

4）市販システムによって実現された試験問題全文検索機能との統合を行い，試験問題文の検索結果から関係する統計情報の参照を可能にした。

論文の前半では，センター試験成績データからの統計量の導出方法について示した。特に採点規則の復元方法と，プログラムにおける採点処理方法の技術的な詳細を説明した。後半では，導出された統計情報を用いて閲覧用のXHTMLファイルを作成するための方法，およびウェブサービスを利用してこれらを閲覧可能とし，さらに一部の統計情報を図示する方法を示した。また，試験問題の全文検索方法について説明し，さらに過去資料において見られる項目別得点率の統計的特徴について検討を加えた。