

PISA-Vertiefungsprojekt: Genauigkeitsanalysen zu den Österreich-Ergebnissen der PISA-Studie 2009

Andreas Quatember

Johannes Kepler Universität Linz, IFAS - Institut für Angewandte Statistik

Abstract:

Die im Dreijahresabstand durchgeführte PISA-Studie (*PISA*: Programme for International Student Assessment) hat das Ziel, die Fähigkeit der Schülerinnen und Schüler einer bestimmten Altersklasse (konkret: der 15- bis 16-jährigen) eines Landes in verschiedenen Kernkompetenzen (Lesen, Mathematik, Naturwissenschaft) zu bestimmen. Die in der Studie erzielten Ergebnisse basieren dabei natürlich nicht auf einer *Vollerhebung* der betreffenden Zielpopulation, sondern auf einer diesbezüglichen *Stichprobenerhebung*. Die interessierenden Parameter sind die Mittelwerte der Lese-, Mathematik- und naturwissenschaftlichen Fähigkeiten in der Zielpopulation. Bei diesem Rückschluss von aus einer Stichprobe gewonnenen Ergebnissen auf Parameter stellt sich automatisch die Frage nach der Genauigkeit der Schätzung – gemessen beispielsweise durch die Breite eines diesbezüglichen Konfidenzintervalls. Dessen Berechnung ist in diesem Fall nicht trivial, weil zum Einen das Stichprobendesign hochkomplex ist und zum Anderen die Messung der Fähigkeiten der einzelnen Schülerinnen und Schüler keinen einzelnen *Fähigkeitsmesswert*, sondern eine *Fähigkeitsverteilung* liefert, welche die diesbezügliche individuelle Fähigkeit repräsentiert.

Eine Möglichkeit der Schätzung der Genauigkeit bieten in solchen Fällen sogenannte Resamplingverfahren wie das Balanced Repeated Half Sampling, das Jackknifen oder das Bootstrappen. Die offizielle Vorgehensweise zur Genauigkeitsschätzung der PISA-Ergebnisse ist eine Variante der erstgenannten Methode. Da jedoch die exakte Vorgehensweise von der OECD nicht öffentlich zugänglich gemacht wird und für die gestellten Aufgabe zu unflexibel erscheint, wurde für dieses Vertiefungsprojekt, das sich mit der Auswirkung der Verwendung unterschiedlicher Stichprobendesigns auf die Genauigkeit der PISA-Resultate beschäftigen sollte, ein intuitiv zugänglicheres, adaptiertes Bootstrapverfahren verwendet.

Im Vortrag werden das komplexe Stichprobendesign der PISA-Studie 2009 in Österreich und die für die Anwendung des Bootstrapverfahrens für endliche Grundgesamtheiten im vorliegenden Fall nötigen Adaptierungen vorgestellt. Ferner werden natürlich die Ergebnisse der Simulationsstudie zum Vergleich verschiedener Stichprobendesigns auf Simulationsbasis diskutiert.