Chapter XI
Identifying Latent Classes and Differential Item Functioning in a Cohort of E-Learning Students

Andrew Sanford
Monash University, Australia

Paul Lajbcygier
Monash University, Australia

Christine Spratt
Royal Australian and New Zealand College of Psychiatrists, Australia

ABSTRACT

A differential item functioning analysis is performed on a cohort of E-Learning students undertaking a unit in computational finance. The motivation for this analysis is to identify differential item functioning based on attributes of the student cohort that are unobserved. The authors find evidence that a model containing two distinct latent classes of students is preferred, and identify those examination items that display the greatest level of differential item functioning. On reviewing the attributes of the students in each of the latent classes, and the items and categories that mostly distinguish those classes, the authors conclude that the bias associated with the differential item functioning is related to the a priori background knowledge that students bring to the unit. Based on this analysis, they recommend changes in unit instruction and examination design so as to remove this bias.
INTRODUCTION

The aim of this chapter is to discuss the identification of latent classes or groups within a cohort of E-Learning students. These latent classes are determined by the existence of differential item functioning, or item bias, experienced by students within these different classes. In our illustrative case study, we are able to identify and interpret a number of these latent classes. Our thesis is that the differential item functioning (DIF), and the latent class structures identified, are a consequence of the students’ diverse educational backgrounds. The case study looks at a unit in computational finance where the students are taking either a single major in commerce or information technology, or double majors in both. We argue that given the multi-disciplinary nature of the unit, those taking a double major are advantaged in terms of background or a priori knowledge over those taking single degrees, resulting in the identified DIF.

DIF analysis seeks to determine the existence of systematic differences in item responses among groups of student, the cause of which is some factor, or factors, other than the innate ability or proficiency of the students. The meaning of ‘innate’ ability is the student trait which the test items have been designed to measure. DIF analysis seeks to identify test items that discriminate amongst students based on factors other than their ability. Item Response Theory (IRT) modeling has a long association with educational and psychometric research, and has proven to be a popular method for detecting DIF. Usually, in investigating DIF with IRT models, students are placed into groups based on the presence or absence of these observable, non-ability factors. Two IRT models are then estimated for each of the groups and their parameters are checked to see whether they are significantly different. If they are, then DIF is considered to exist.

The DIF analysis discussed in this chapter uses examination items and student responses from a unit in computational finance, which has been taught by one of the authors for many years. Materials in this unit have been designed to suit online E-Learning, and all assessment has been prepared in a multiple choice format, appropriate for automated delivery and scoring. The unit attracts students from a diverse range of educational and cultural backgrounds, and thus provides a ready number of observed factors (e.g. gender, major area of study, years of academic attainment, international status, ethnic background, etc.) which can be used to put students into groups.

Although using observed factors is common, it might not be valid in all circumstances. For example, DIF may be due to factors, such as a student’s level of motivation, learning intentions, language deficiencies, anxiety or problem solving strategies, which are not readily observable. An alternative to specifying the student class membership prior to carrying out the DIF analysis is to infer the student membership as an output of the DIF analysis. In our case study, a predefined number of latent classes are specified within the IRT model, and students are allocated to those classes based on their test item responses.

Within the case study, DIF analysis is carried out using a polytomous IRT model previously developed by Bolt, Cohen and Wollack (2001). A valuable output common to most IRT models, and one which provides a useful visual display of DIF, are the Item Response Functions (IRFs) or Item Category Characteristic Curves (ICCCs). These functions display the probabilities associated with the selection of each item and/or category as a function of a student’s ability (or proficiency) level. The correct category response usually records the highest probability. We reproduce a number of ICCC functions to illustrate the differences in response probabilities for the different latent classes of student. (See Figure 2 and Figure 3.)

In the following sections, we review the current IRT and assessment literature; discuss features of the computational finance unit, and the student response data. We then discuss in greater detail