A BAYESIAN PROCEDURE FOR MASTERY DECISIONS BASED ON MULTIVARIATE NORMAL TEST DATA

HUYNH HUYNH
UNIVERSITY OF SOUTH CAROLINA

A Bayesian framework for making mastery/nonmastery decisions based on multivariate test data is described in this study. Overall, mastery is granted (or denied) if the posterior expected loss associated with such action is smaller than the one incurred by the denial (or grant) of mastery. An explicit form for the cutting contour which separates mastery and nonmastery states in the test score space is given for multivariate normal test scores and for a constant loss ratio. For multiple cutting scores in the true ability space, the test score cutting contour will resemble the boundary defined by multiple test cutting scores when the test reliabilities are reasonably close to unity. For tests with low reliabilities, decisions may very well be based simply on a suitably chosen composite score.

Key words: mastery testing, testing for selection, cutting contour, multivariate decisions, Bayesian decisions, error of measurement, test score fallibility.

Introduction

Application of mental measurement to selection or certification problems often involves the use of more than one test score. For example, the selection of students for an advanced program in some subject area may be based on several traits (variables), such as prior achievement, aptitude, interest, etc. Ideally, selection should be based on the subject’s true measures on these traits; in reality, however, decisions are typically based on observed test scores which are contaminated with errors of measurement. Thus, misclassifications are bound to occur, and rules for decisions based on test data are typically formulated in such a way as to minimize the risks incurred by misclassification.

Decision problems based on one variable have been considered at length in the literature. Statistical issues involved in establishing a single cutoff (cutting, passing, or mastery) score are described in detail in a number of sources including Swaminathan, Hambleton, and Algina [1975]; Huynh [1976, 1977, 1979, 1980]; Wilcox [1976]; and van der Linden and Mellenbergh [1977]. Huynh [1979, 1980] also provides an explicit relationship among test cutting score, losses incurred by misclassification, and errors of measurement. In general, within the minimax or empirical Bayes decision framework, it is found that errors in measurement will reduce the test cutting score when a false negative error is more serious than a false positive error. Conversely, the test cutting score will increase when a false negative error is less serious than a false positive error.

The effect of errors of measurement in selection situations involving multiple true cutting scores was considered by Lord [1962]. His procedure involved the use of a regression line (surface) which expresses the amount of “desirability” (Y) assigned to each exam-
inee as a function of his or her true (infallible) scores ($\theta$). When $Y(\theta)$ was continuous, the contour line in the $\theta$-space at which $Y(\theta)$ assumed a constant defined the boundary or cutting contour between subjects deemed acceptable (masters) and those judged as unacceptable (nonmasters). To translate the true score cutting contour, the conditional expected desirability $E_d(Y \mid x)$ was defined at each observed test score vector $x$. The cutting contour in the $x$-space was then taken as the line at which the expected desirability $E_d(Y \mid x)$ assumed a constant. Within the general context as described above, Lord focused on the special case in which $Y(\theta)$ was equal to the smallest component of the true score vector $\theta$. (This implies that the cutting contour in the $\theta$-space consists of several half lines (planes), each being parallel to one of the axes which define the $\theta$-space, and extending in the positive direction of the test scores; in other words, multiple true cutting scores are used to define the true masters.) Using the multivariate normal distribution to describe the true and observed scores, Lord was able to plot the contour line in the observed test score plane which separates the subjects deemed acceptable (masters) from those judged as unacceptable (nonmasters).

Through the conditional expected value $E_d(Y \mid x)$, the Lord procedure is reminiscent of the Bayesian decision procedure based on posterior distributions [Wald, 1950; Ferguson, 1967]; no immediate, natural connection between the two approaches seems apparent to the author of this paper, however. (The conditional nature of both approaches may account for the similarity in the cutting contours subsequently illustrated in this paper and those reported by Lord.) Granting that Bayesian decision theory is a general framework by which decisions based on test scores can be made, this paper will focus on a Bayesian reformulation of the cutting contour problem which was previously explored by Lord [1962].

The purpose of this paper is twofold. First it will describe a Bayesian solution to the “plotting” of a cutting contour in selection situations involving multiple test scores. Second, it will explore the influence of the loss ratio on the cutting contour and will reexamine the distortion caused by errors of measurement [Lord, 1962], using a Bayesian decision-theoretic framework. Examples based on the multivariate normal distribution with constant losses for misdecisions are provided to illuminate various points or procedures put forward in the paper.

**A Bayesian Approach to Cutting Contour**

Now let the column vector $\theta = (\theta_1, \theta_2, \ldots, \theta_k)'$ denote the true scores (measures) of an individual subject on $k$ traits (or selection variables). Let $\Omega$ represent the region in the true score space where a subject must be located in order to qualify for the true state of mastery. Thus a subject is defined as a true master if $\theta \in \Omega$. Let $\Omega'$ be the complement of $\Omega$. Then a subject is declared a true nonmaster when $\theta \in \Omega'$. 

Now let the column vector $x = (x_1, x_2, \ldots, x_k)'$ represent the observed test scores of the subject. On the basis of $x$ and other prior information regarding $\theta$, a decision may be made concerning the subject: either to grant mastery (action $a_1$) or to deny mastery (action $a_2$). When $\theta \in \Omega$, the best course of action is $a_1$, and no loss will be encountered. Similarly, action $a_2$ is best when $\theta \in \Omega'$. For other situations, classification errors occur. To be specific, the choice of action $a_2$ when $\theta \in \Omega$ constitutes a false negative error, whereas the selection of $a_1$ when $\theta \in \Omega'$ produces a false positive error.

Let $C_1(\theta)$ be the loss associated with a false negative error and $C_2(\theta)$ be the loss encountered by a false positive error. Let $p(\theta \mid x)$ be the posterior probability density of $\theta$ given that the test score vector $x$ has been observed. Given $x$, the posterior expected loss encountered in taking action $a_1$ is given by the integral $R(a_1 \mid x) = \int_{\Omega'} C_1(\theta)p(\theta \mid x) \, d\theta$. 