# DECISIONS BASED ON TESTS: SOME RESULTS WITH A LINEAR LOSS FUNCTION

Gideon J. Mellenbergh
University of Amsterdam, Amsterdam, The Netherlands
Wim J. van der Linden
Twente University of Technology, Enschede, The Netherlands

Abstract

An important problem in education is determining cutting scores on educational tests consisting of items that can be answered right or wrong. Students with the number of items answered correctly that is equal to or greater than the cutting score pass the test. The others must study the subject again and take a new test later. This problem is comparable to determining the cutting score on a selection test in applied psychology, for instance accepting people for a job, psychotherapy, treatment, and so on. An extra requirement that cutting scores for these procedures should meet, is that they should be fair with respect to the various categories represented among the applicants. A decision theoretic approach with a linear loss function, which results in a simple procedure for determining optimal scores, is discussed.

Paper presented at the European Meeting on Psychometrics and Mathematical Psychology, University of Uppsala, Uppsala, Sweden, June 15-17, 1978.

1.1. Suppose that an educational or psychological test consists of n dichotomously scored (0-1) items and that the random variable X represents the unweighted sum of item scores of a randomly selected person i. Suppose, too that X is used to measure or predict the score of person i on a continuous random variable Z. Finally, suppose that X can be used for this purpose, since $E(Z|X)$, the expectation of Z given X, is monotonically related to X.

1.2. A problem often encountered in educational and psychological testing is the dichotomous decision, which can be represented as follows:

Variable Z

| | Not suitable ($Z < d$) | d | Suitable ($Z \geqslant d$) |
|---|---|---|---|
| Accepted ($X \geqslant c$) | | | |
| Testscore X  c | | | |
| Not-accepted ($X < c$) | | | |

The variable Z is dichotomized into the categories "suitable" ($Z \geqslant d$) and "not-suitable" ($Z < d$), and a cutting score c defines the decisions "accepted"($X \geqslant c$) and "not-accepted" ($X < c$). Examples are pass-fail decisions in educational testing and acceptance-rejectance decisions for job applicants or for psychotherapeutic treatments. The problem is to determine c so that the decisions are optimal in one sense or another.

1.3. Ideally, all suitable persons are accepted and all not-suitable persons not accepted, but because of measurement or prediction errors this state is rare. In order to weight the consequences of these errors a loss function should be specified. Here we consider the following linear loss function that we have introduced elsewhere (van der Linden & Mellenbergh, 1977):

$$L(Z) = \begin{cases} b_0(Z - d) + a_0 & \text{for } X < c \\ b_1(d - Z) + a_1 & \text{for } X \geq c \end{cases} \qquad (1)$$

with $(b_0 + b_1) > 0$. Both parts of Formula 1 contain two different terms and, hence, two different kinds of parameters. The terms $b_0(Z - d)$ and $b_1(d - Z)$ represent the amounts of loss dependent upon the difference between the person's score on Z and the cutting score d; the parameters $b_0$ and $b_1$ are therewith constants of proportionality. The terms $a_0$ and $a_1$ represent the amounts of loss independent of Z and are constant for each decision (see: Figure 1).

1.4. An optimization criterion known from statistical decision theory is the Bayes risk, shortly called the risk. Defining the joint density of X and Z as k(X,Z), the risk is the expected loss with respect to k(X,Z):

$$R = \sum_{x=0}^{c-1} \int_0^d L(Z)k(X,Z)dZ + \sum_{x=0}^{c-1} \int_d^1 L(Z)k(X,Z)dZ +$$

$$\sum_{x=c}^{n} \int_0^d L(Z)k(X,Z)dZ + \sum_{x=c}^{n} \int_d^1 L(Z)k(X,Z)dZ. \qquad (2)$$

Considering R as an function of c, an optimal cutting score c' is that value of c for which R(c) is minimal.

1.5. Substituting Formula 1 into 2 gives

$$R(c) = \sum_{x=0}^{c-1} \int_0^1 \{ b_0(Z - d) + a_0 \} k(X,Z)dZ +$$

$$\sum_{x=c}^{n} \int_0^1 \{ b_1(d - Z) + a_1 \} k(X,Z)dZ. \qquad (3)$$

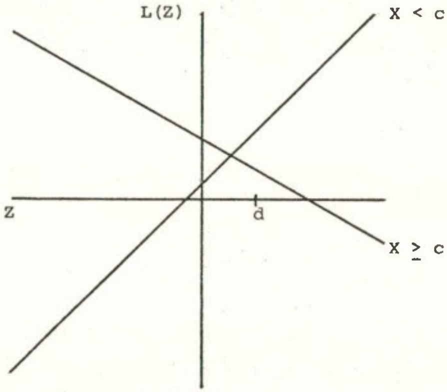
Using

$$k(X,Z) = p(Z|X)h(X) = 1, \qquad (4)$$

$$\int_0^1 p(Z|X)dZ = 1, \qquad (5)$$

Figure 1

An Example of the Linear Loss Function, $a_0 \neq a_1$, $b_0 \neq b_1$.



$$\int_0^1 Zp(Z|X)\,dZ = E(Z|X), \tag{6}$$

where $p(Z|X)$ and $h(X)$ are the probability density of Z given X, and the density of X, it follows that

$$R(c) = \sum_{x=0}^{c-1} \left[b_0 \{ E(Z|X) - d \} + a_0\right]h(x) -$$

$$\sum_{x=c}^{n} \left[b_1 \{ E(Z|X) - d \} + a_1\right]h(x). \tag{7}$$

This is equivalent to

$$R(c) = \sum_{x=0}^{n} \left[b_0 \{ E(Z|X) - d \} + a_0\right]h(x) -$$

$$\sum_{x=c}^{n} \left[(b_0 + b_1) \{ E(Z|X) - d \} + (a_0 - a_1)\right]h(X). \tag{8}$$

Remembering $(b_0 + b_1) > 0$ and eliminating the constant term of Formula 8 , R(c) is minimal for the cutting score c' that maximizes:

$$R'(c) = \sum_{x=c}^{n} \left[(b_0 + b_1) \{ E(Z|X) - d \} + (a_0 - a_1)\right]h(X). \tag{9}$$

Since $h(X) \geqslant 0$ and $E(Z|X)$ is assumed to be a monotonically increasing function, R'(c) is maximal for that value c = c' for which

$$(b_0 + b_1) \{ E(Z|X) - d \} + (a_0 - a_1) \tag{10}$$

is positive for the first time. Using this result, the optimal cutting score can be found if the regression function is specified.

2.1. As a first application of the above result, let us consider the problem of pass-fail decisions in education. Suppose that a test is administered to students in order to examine whether they have mastered a certain subject matter; if not, they should relearn it.

For a specified student the mastery level $T$ with respect to this subject matter can be defined as the expected proportion of items answered correctly

$$T = E(X/n), \tag{11}$$

where the expectation is taken with respect to the propensity distribution of $X/n$ for the specified student.
Suppose that a student is considered to have mastered the subject matter if his mastery level $T$ exceeds a critical value $\tau$; if not, he should relearn the material. Identifying $Z$ with $T$ and $d$ with $\tau$, the above derived results apply and an optimal decision procedure can be designed.

2.2. As Formula 11 is the definition of the true score from classical test theory, a possible regression function for Formula (10) is the linear regression function:

$$E(T|X) = \rho_{XX'} X/n + (1 - \rho_{XX'}) E(X/n), \tag{12}$$

where $\rho_{XX'}$ is the classical reliability coefficient (Lord & Novick, 1968, p. 65). Substituting Formula 12 into 10, setting the result equal to 0, and solving for X yields

$$X' = [E(X) + n \{ d - (a_0 - a_1) / (b_0 + b_1)\} - E(X)] / \rho_{XX'} \tag{13}$$

The cutting score is an integer. For the first integer smaller than $X'$ Formula 10 is negative and for the first integer larger than $X'$ the expression is positive. Therefore, the optimal cutting score is

$$c' = \text{entier } (X') + 1 \tag{14}$$

2.3. The following feature of this optimal procedure for pass-fail decisions is of practical importance. When the restriction $a_0 = a_1$ applies, all parameters of loss function Formula 1 disappear from Formula 10 and the optimal cutting score is that value of c for which $\{E(T|X) - d\}$ is positive for the first time. Or, using the linear regression function Formula 12, the optimal cutting score is the first integer value above

$$X' = E(X) + \{nd - E(X)\} / \rho_{XX'} \qquad (15)$$

In statistical terms this means that loss function Formula 1 is maximally robust under the restriction $a_0 = a_1$. When for each decision the constant components of the loss are equal, the values of the parameters of this loss function need not even to be specified.

3.1. A second application of this approach is in selecting applicants for positions, choosing students for advanced educational programs or for remedial teaching, and so on. An important issue in selection is the fair treatment of applicants from different subpopulations, especially of applicants from disadvantaged subpopulations. Several models for determining different cutting scores in subpopulations have been proposed but most of them are inconsistent (Petersen & Novick, 1976). Novick and Petersen (1976) have argued correctly that acceptance-rejection decisions in selection should be made for each individual separately and should be based on the risk: only individuals with the lowest risk should be accepted. Petersen (1976) has used a threshold loss function for this purpose; in this paper the linear loss function is used (Mellenbergh & van der Linden, 1978).

3.2. In selection two situations are distinguished, quota-free and quota-restricted. In quota-free selection there is no restriction on accepted applicants: all applicants who satisfy the requirements are accepted. In quota-restricted selection, however, only a fixed number of applicants are accepted.

3.3. In the selection situation the variable Z is interpreted as the score on an external criterion, for example succes on the job, in an educational program, in psychotherapy, and the like. The variable X is the testscore used for predicting the criterion. In this situation different loss functions are necessary, one for each subpopulation. The linear loss function for subpopulation i ($i = 1, 2, \ldots, g$) is :

$$L_i(Z) = \begin{cases} b_{0i}\ (Z - d) + a_{0i} & \text{for } X < c_i \\ b_{1i}\ (d - Z) + a_{1i} & \text{for } X \geq c_i \end{cases} \tag{16}$$

Using the index i at the appropriate positions in Formulas 3 through 7, it follows from Formula 7 that the risk in subpopulation i is:

$$R_i(c) = \sum_{X=0}^{c_i - 1} \left[ b_{0i} \{ E_i(Z|X) - d \} + a_{0i} \right] h_i(X) - \sum_{X=c_i}^{n} \left[ b_{1i} \{ E_i(Z|X) - d \} + a_{1i} \right] h_i(X). \tag{17}$$

The selection process is viewed as a series of separate decisions, each of which involves one random applicant from the total population, and it is assumed that the overall risk of the selection process is the sum of the risks of the applicants. Thus, the overall risk of the selection process is:

$$R_0 = \sum_{i=1}^{g} p_i\ R_i(c) \tag{18}$$

where $p_i$ , $\sum_{i=1}^{g} p_i = 1$, is the proportion of applicants from subpopulation i in the total population of applicants.

3.4. Since, in quota-free selection there is no restriction on applicants who can be accepted, Formula 18 is minimized if the risk of a random applicant is minimized. This is done by minimizing Formula 17 for every subpopulation separately. From Formula 10 follows that the optimal cutting score in subpopulation i is the value for which

$$(b_{0i} + b_{1i})\ \{ E_i(Z|X) - d \} + (a_{0i} - a_{0i}) \tag{19}$$

is positive for the first time. A special case of a monotonically increasing regression function is the regression line:

$$E_i(Z|X) = \alpha_i + \beta_i X, \tag{20}$$

where $\alpha_i$ and $\beta_i$ are the intercept and slope of the regression line in subpopulation i. Substituting Formula 20 into 19, setting the result equal to 0, and solving for X yields:

$$X_i' = (d - \alpha_i) / \beta_i + (a_{0i} - a_{1i}) / \beta_i (b_{0i} + b_{1i}). \tag{21}$$

The cutting score is an integer; for the first integer smaller than $X_i'$ Formula 21 is negative, and for the first integer larger than X' the expression is positive. Therefore, the optimal cutting score in subpopulation i is:

$$c_i' = \text{entier } (X_i') + 1 \tag{22}$$

For the special case $a_{0i} = a_{1i}$ (i = 1, 2, ... , g), it follows from Formula 21 that the optimal cutting score is the first integer larger than

$$X_i' = (d - \alpha_i) / \beta_i \tag{23}$$

In this case the constants of proportionality $b_{0i}$ and $b_{1i}$ are immaterial for determining the optimal cutting scores.
These cutting scores are also derived in the so-called regression model for culture-fair selection; Petersen and Novick (1976) ascertain that this is a consistent model in contrast with other models proposed for culture-fair selection.

3.5. Since, in quota-restricted selection only a fixed proportion p af all applicants is accepted, Formula 18, is minimized under the restriction:

$$\sum_{i=1}^{g} p_i \sum_{X=c_i}^{n} h_i(X) = p \tag{24}$$

Because the testscore is a discrete variable this condition cannot in general be exactly satisfied. Therefore, an upper bound ($p_u$) and a lower bound ($p_\ell$) are fixed: the proportion of accepted applicants from all applicants should be within these bounds. The restriction of Formula 24 becomes:

$$p_\ell < \sum_{i=1}^{g} p_i \sum_{x=c_i}^{n} h_i(X) < p_u. \qquad (25)$$

For determing the optimal cutting scores the following procedure can be used: First, the total number s of sets of cutting scores meeting the restriction of Formula 25 are determined: $A_1 = \{c_{11}, c_{21}, \ldots, c_{g1}\}, \ldots, A_s = \{c_{1s}, c_{2s}, \ldots, c_{gs}\}$. Second, from empirical data the regression functions $E_i(Z|X)$, and the probability densities $h_i(X)$ are estimated. Third, using Formula 18 the overall risk is computed for each set $A_j$ : $R_{0j}$ $(j = 1, 2, \ldots, s)$. Fourth, the minimal value of the risks $R_{0j}$ is determined. If $R_{0r} = \min_j \{R_{0j}\}$ then the set $A_r$ contains the optimal cutting scores for the g sub-populations.

## References


Lord, F.M. & Novick, M.R. Statistical theories of mental test scores. Reading, MA.: Addison-Wesley, 1968.

Mellenbergh, G.J. & van der Linden, W.J. The linear utility model for optimal selection (submitted for publication).

Novick, M.R. & Petersen, N.S. Towards equalizing educational and employment opportunity. Journal of Educational Measurement, 1976, 13, 77-88.

Petersen, N.S. An expected utility model for "optimal" selection. Journal of Educational Statistics, 1976, 1, 333-358.

Petersen, N.S. & Novick, M.R. An evaluation of some models for culture-fair selection. Journal of Educational Measurement, 1977, 1, 593-599.

Van der Linden, W.J. & Mellenbergh, G.J. Optimal cutting scores using a linear loss function. _Applied Psychological Measurement_, 1977, _1_, 593-599.