DOCUMENT RESUME

ED 310 155 TM 013 769

AUTHOR	van der Linden, Wim J.
TITLE	On the Estimation of the Proportion of Masters in Criterion-Referenced Testing. Twente Educational Memorandum No. 27.
INSTITUTION	Twente Univ., Enschede (Netherlands). Dept. of Education.
PUB DATE	Apr 81
NOTE	17p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Criterion Referenced Tests; *Estimation (Mathematics); Foreign Countries; *Latent Trait Theory; *Mastery Tests; Mathematical Models; *Maximum Likelihood Statistics; Monte Carlo Methods

ABSTRACT

J. A. Emrick's (1971) model is a latent class model of mastery testing that can be used to estimate the proportion of masters in a given population. A. Hamerle (1980), in a recent paper on this model, has proposed an estimator for the proportion of masters that is claimed to constitute a maximum likelihood approach. It is indicated that Hamerle is not quite correct in his presentation of Emrick's model and that his estimator is not maximum likelihood. An estimator is provided using the method of moments; this estimator appears to have the same shape as Hamerle's estimator, but should be interpreted differently since it is derived under the correct version of Emrick's model. An attractive property of the method of moments is that it also yields simple estimators for the present model if the two success parameters are unknown. It appears that these estimators can be used for tests consisting of three or more items. Results of extensive Monte Carlo studies indicate that the estimators possess excellent statistical properties. (Author/TJH)

TWENTE EDUCATIONAL MEMORANDUM NUMBER 27

ON THE ESTIMATION OF THE PROPORTION OF MASTERS IN CRITERION-REFERENCED TESTING

W.J. van der Linden

TWENTE UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF EDUCATIONAL TECHNOLOGY
P.O. BOX 217
7500 AE ENSCHEDE THE NETHERLANDS

COLOFON:

Publisher	: Department of Educational Technology Twente University of Technology
Typist	: Paula Achterberg
Reproduction	: Central Reproduction Department of the University
Edition	: 60

ON THE ESTIMATION OF THE PROPORTION OF MASTERS IN CRITERION - REFERENCED TESTING

W.J.van der Linden

Twente Educational Memorandum No. 27. Enschede,
The Netherlands: Twente University of Technology,
Department of Educational fechnology, April 1981

Abstract

Emrick's model is a latent class or state model of mastery testing which can be used to estimate the proportion of masters in a given population. Hamerle, in a recent paper on this model, has proposed an estimator for the proportion of masters that is claimed to be maximum likelihood. It is indicated that Hamerle is not quite correct in his presentation of Emrick's model and that his estimator is not maximum likelihood. An estimator is given using the method of moments which appears to have the same shape as Hamerle's estimator but should be interpreted differently since it is derived under the correct version of Emrick's model. An attractive property of the method of moments is that it also yields simple estimators for the present model if the two success parameters are unknown. This paper indicates that these estimators can be used for tests consisting of three or more items and have demonstrated in extensive Monte Carlo studies to possess excellent statistical properties.

ON THE ESTIMATION OF THE PROPORTION OF MASTERS IN CRITERION-REFERENCED TESTING

Recently, Hamerle (1980) has discussed the use of Emrick's model for estimating the proportion of masters in criterion-referenced testing. Emrick's model is a latent class or state model. It assumes that mastery and nonmastery of a well-defined domain of knowledge or skills can be viewed as two different latent states, each characterized by a different set of probabilities of a successful response to the test items. In this respect, the model differs from a continuum model of mastery testing in which a latent continuum is assumed to underly the test items and a cutoff score on the continuum is used to define mastery and nonmastery. When the continuum view of mastery is taken and items are randomly sampled from the domain, the use of the simple binomial model is an appropriate choice (Klauer, 1972). Fricke (1972) advocates the use of the Rasch model for mastery testing with a continuum. For a critical discussion of some differences between state and continuum models for mastery testing, see van der Linden (1978)

Hamerle has also proposed a maximum likelihood estimator for the proportion of masters in Emrick's model. In order to arrive at this estimator, he assumes that the probabilities of a correct item response are known for the class of masters as well as the nonmasters. The same assumption is needed, according to Hamerle, to prevent Emrick's model from being unidentifiable.

The purpose of the present note is to indicate that Emrick's model is not adequately represented and interpreted by Hamerle. As a consequence, the maximum likelihood estimator for the proportion of masters given by him is not correct. In the following sections, it will be shown how for the correct version of Emrick's model an estimator can be constructed using the method of moments, which has the same shape as Hamerle's estimator but possesses a different interpretation. A closed-form maximum likelihood estimator does not seem to exist since the likelihood equation is intractable and presumably only to be solved using iterative procedures. An attractive property of moment estimation for Emrick's model is that it yields comparatively simple, explicit estimators when the success probabilities cannot be assumed to be known, whereas maximum likelihood estimation again results in intractable equations entailing the use of iterative procedures. Finally, some light will be shed on the behavior of the moment estimators and some additional comments on the paper by Hamerle will be made.

Emrick's model

Let M denote the state of mastery and $\bar{M}$ the state of nonmastery. Further, let X designate the number-right score on a test of length n sampled from the domain with respect to which mastery has to be assessed for a sample of m examinees with Prob{ M } = μ. Finally, let α denote the probability that an examinee in state $\bar{M}$ will give a correct answer to an item, and β the probability that an examinee in state M will do likewise.

As can be verified from Emrick (1971) and Emrick and Adams (1969), the probability of an observed test score X = x is now equal to

(1) $$p(x) = (1 - \mu)\binom{n}{x}\alpha^x(1 - \alpha)^{n-x} + \mu\binom{n}{x}\beta^x(1 - \beta)^{n-x}.$$

The model given in (1) is generally known as Emrick's model for mastery testing. It is important to note that the model consists of a mixture of two binomials with success parameters α and β and mixing parameter μ.

The model considered by Hamerle is not Emrick's model but differs from (1) in two important aspects. First, the probability function presented by Hamerle as Emrick's model is not a model for the number-right test score but for the two-valued random variable that denotes the mastery, $X \geq c$, and the nonmastery decision, $X < c$ (c being the cutoff score on the test). Second, the success parameters α and β, which represent the probabilities of a nonmaster and a master producing a correct answer to an item, have been replaced in Hamerle's paper by the probabilities that a nonmaster, respectively, a master produces a test score $X \geq c$. Hamerle's model is not a mixture of two binomials, as Emrick's model is, but a mixture of two alternative distributions giving the probability of a mastery decision ($X \geq c$). As a consequence Hamerle's model cannot be viewed as a usual latent class model. It does not define mastery as a latent state underlying the items but instead defines latent states with respect to the mastery decisions to be taken after the test has been administered. It should be noted that these decisions depend on the test cutoff score, c, which, in principle, may be set at any test score value by the teacher or testing agency.

Estimating μ

It will be obvious that Hamerle's estimator is not a maximum likelihood estimator for the proportion of masters in Emrick's model, Further, it appears that writing out the likelihood equation for the model in (1) gives rise to an intractable result which presumably can only be solved iteratively using computer programming. This is due to the fact that the step taken by Hamerle to go from equation (3.2) to the next equation holds for a mixture of two alternative distributions but does not have a generalization that holds for a mixture of two binomials which is the basis of Emrick's model. However, it is possible to construct an estimator for μ in (1) using the method of moments which has the same shape as Hamerle's estimator but should be interpreted differently. This will be demonstrated in the remainder of this section.

Let $X_1, \ldots, X_i, \ldots, X_m$ be a random sample obtained under model (1). The number m is interpreted as the number of randomly selected examinees to which the test will be administered. The method of moments provides estimators by writing the parameters as functions of population moments and replacing the population moments by their corresponding sample functions. It is convenient to take the first r moments (provided that these exist), where r is the number of parameters to be estimated.

For mixtures of binomials it is usual to define

$$(2) \qquad \psi_k = E\left[X(X-1)\ldots(X-k+1)\right]/n(n-1)\ldots(n-k+1),$$

$k = 0, 1, \ldots, n$. The expression defined in (2) is, up to the factor $n(n-1)\ldots(n-k+1)$ in the denominator, equal to the kth factorial moment. Note that for $k = 1$, (2) yields the expected relative test score, $\psi_1 = E\left[X\right]/n$. For $k = 2$, (2) gives a simple linear function of the expected test score and variance. Comparable functions are obtained for larger values of k. The sample functions corresponding to (2) are given by

(3) $$\hat{\psi}_k = \sum_{i=1}^{m}\left[X_i(X_i - 1)\ldots(X_i - k + 1)\right]\Big/mn(n - 1)\ldots(n - k + 1).$$

For k = 1, (3) results in the average relative test score, $\hat{\psi}_1 = \sum_{i=1}^{m} X_i/nm$, whereas for k = 2 a simple linear function of the sample mean and variance is obtained. Larger values of k result in similar functions.

The reason for defining (2) is that it gives rise to an elegant result for the mixture of two binomials given in (1), namely

(4) $$\psi_k = (1 - \mu)\alpha^k + \mu\beta^k$$

(e.g., Johnson & Kotz, 1969, sect. 3.11).

Suppose α and β are known success probabilities. Then, for k = 1, solving (4) for μ and substituting (3) gives the moment estimator

(5) $$\hat{\mu} = \frac{\sum_{i=1}^{m} X_i/mn - \alpha}{\beta - \alpha}.$$

This expression is exactly the same as the one obtained by Hamerle except that the average relative test score, $\sum_{i=1}^{m} X_i/nm$, replaces the proportion of mastery decisions in Hamerle's estimator. This makes sense since the role played by the test score in Emrick's model has been taken over by the 0-1 variable indexing the mastery decisions in Hamerle's model. Note, however, that (5) is a moment estimator obtained under Emrick's model while Hamerle gives a maximum likelihood estimator for a different model.

Estimating α and β

In order to arrive at (5), we have followed Hamerle and supposed that α and β are known. It can be argued, however, that this assumption is not so realistic since no model seems available, which allows us to specify the true values of both α and β. Therefore, estimators for Emrick's model are needed for the case in which not only μ but also β or α and β are unknown. No explicit maximum likelihood estimators for these two cases are known and one has to resort to using one of the available numerical procedures to solve the likelihood equations (Goodman, 1974; Macready & Dayton, 1977). However, the situation is different for moment estimation.

The case of μ and β as unknowns will be considered first. An example is the mastery or random guessing model for multiple-choice items. According to this model, an examinee is either in state M and produces the correct answer with probability β or in state $\bar{M}$ and guesses at random, so that α may be set at q^{-1}(q being the number of options). This case has been considered by van der Linden (1981b).

When two parameters are to be estimated, two moment equations are needed. For k = 1 and 2, (4) yields a system of two equations which can be solved for μ and β. Substituting $\hat{\psi}_1$ and $\hat{\psi}_2$ from (3), then, gives moment estimators for these two parameters. As can be verified from van der Linden (1981b), the result is equal to

$$(6) \qquad \tilde{\beta} = \frac{\hat{\psi}_2 - \alpha^2}{\hat{\psi}_1 - \alpha} - \alpha,$$

while substituting $\tilde{\beta}$ for β in (5) gives the moment estimator for μ. Thus, when β cannot be assumed to be known, β can simply be estimated from the sample mean and variance and the estimator for μ only needs an obvious adaptation.

Suppose that all three parameters are unknown. This will be the case when, for instance, the items are not of the multiple-choice type or the mastery or random guessing model does not hold for other reasons. Now three moment equations are needed to derive the estimators. It can be verified from Blischke (1962) that following the same procedure as above leads to

(7) $$\mathring{\alpha} = \frac{1}{2}a - \frac{1}{2}(a^2 - 4a\hat{\psi}_1 + 4\hat{\psi}_2)$$

and

(8) $$\mathring{\beta} = \frac{1}{2}a + \frac{1}{2}(a^2 - 4a\hat{\psi}_1 + 4\hat{\psi}_2),$$

where a is an auxiliary expression defined as

$$a = \frac{\hat{\psi}_3 - \hat{\psi}_1\hat{\psi}_2}{\hat{\psi}_2 - \hat{\psi}_1^2}.$$

Substituting $\mathring{\alpha}$ and $\mathring{\beta}$ into (5) gives the moment estimator for μ when both α and β cannot be assumed to be known. For further details and existence conditions, see Blischke (1962) and van der Linden (1981a).

Discussion

In the foregoing, it was indicated that Hamerle's presentation of Emrick's model is not quite correct and that, as a consequence, his estimator for the proportion of masters is not maximum likelihood. An estimator was given using the method of moments which, interestingly, appears to have exactly the same shape as Hamerle's estimator but should be interpreted differently. Also, attention was called to the fact that moment estimators are available in case one or both of the parameters are not known. Two additional comments should be made.

In Hamerle's paper, it is noted that the model presented as Emrick's model is not identifiable, i.e., that distinct sets of parameter values can lead to the same probability distribution for the data. This implies that the parameters cannot be estimated uniquely. Hamerle solves this problem by assuming that α and β are known, thus reducing the number of parameters to be estimated and hence guaranteeing that the condition of identifiability is satisfied. Mixtures of binomials are identifiable if $n \geq r$, where r still denotes the number of parameters to be estimated (Blischke, 1964). For Emrick's model this means that all parameters can be estimated as long as the test has three or more items. This condition will generally be fulfilled in mastery testing and there will thus be no need to assume that some parameters are known to be able to estimate the remaining ones.

Finally, it is observed that moment estimators for Emrick's model have been subjected to extensive Monte Carlo studies of their statistical properties. This has been done for the case of one known success parameter (van der Linden, 1981b) as well as the case of both success parameters being unknown (van der Linden, 1981a). An explicit comparison of the results is given in van der Linden (1980). The general conclusion is that moment estimators have favorable properties and can safely be used in most practical situations. When the mixture of binomials in (1) degenerates into a single model, i.e., when α and β approach each other, the estimators display larger variability. This trend is less observable if one of the success parameters is known, suggesting that it can be expected to be absent for the estimator of μ in (5) obtained under the condition that both success parameters are known.

References

Blischke, W.R. Moment estimators for the parameters of a mixture of two binomial distributions. Annals of Mathematical Statistics, 1962, 33, 444-454.

Blischke, W.R. Estimating the parameters of mixtures of binomial distributions. Journal of the American Statistical Association, 1964, 59, 510-528.

Emrick, J.A. An evaluation model for mastery testing. Journal of Educational Measurement, 1971, 8, 321-326.

Emrick, J.A., & Adams, E.N. An evaluation model for individualized instruction (Report RC 2674). Yorktown Hts., N.Y.: IBM Thomas J. Watson Research Center, October 1969.

Fricke, R. Lehrzielorientierte Messung mit Hilfe stochastischer Messmodelle. In K.J. Klaue., R. Fricke, M. Herbig, H. Rupprecht, & F. Schott, Lehrzielorientierte Tests. Düsseldorf: Schwann, 1972.

Goodman, L.A. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika, 1974, 61, 215-231.

Hamerle, A. Uber die Schätzung des Matery-Anteils bei lehrzielorientierten Tests. Zeitschrift für erziehungswissenschaftliche Forschung, 1980, 14, 117-121.

Johnson, N.L., & Kotz, S. Distributions in statistics: Discrete distributions. Boston: Houghton Miflin Company, 1969.

Klauer, K.J. Zur Theorie und Praxis des binomialen Modells lehrzielorientierter Tests. In K.J. Klauer, R. Fricke, M. Herbig, H. Rupprecht, & F. Schott, Lehrzielorientierte Tests. Düsseldorf: Schwann, 1972.

Macready, G.B., & Dayton, C.M. The use of probabilistic models in the assessment of mastery. Journal of Educational Statistics, 1977, 2, 99-120.

van der Linden, W.J. Forgetting, guessing, and mastery: The Macready and Dayton models revisited and compared with a latent trait approach. Journal of Educational Statistics, 1978, 3, 305-318.

van der Linden, W.J. Simple estimates for the simple latent class mastery testing model. Paper presented at the Fourth International Symposium on Educational Testing, Antwerp, Belgium, June 24-27, 1980.

van der Linden, W.J. Estimating the parameters of Emrick's mastery testing model. Applied Psychological Measurement, 1981, 5, to appear. (a)

van der Linden, W.J The use of moment estimators for mixtures of two binomials with one known success parameter. Submitted for publication, 1981. (b)

Acknowledgement

Thanks are due to Paula Achterberg for typing the manuscript.

Author's Address

Wim J. van der Linden, Onderafdeling Toegepaste Onderwijskunde, Technische Hogeschool Twente, Postbus 217, 7500 AE Enschede, The Netherlands.