The Use of Probabilistic Models in the Assessment of Mastery

George B. Macready and C. Mitchell Dayton

JOURNAL OF EDUCATIONAL AND BEHAVIORAL STATISTICS 1977 2: 99
DOI: 10.3102/10769986002002099

Journal of Educational Statistics
Summer 1977, Volume 2, Number 2, Pp. 99-120.

THE USE OF PROBABILISTIC MODELS IN THE ASSESSMENT OF MASTERY

George B. Macready
University of Maryland

C. Mitchell Dayton
University of Maryland

Key words: Mastery Testing; Latent Class Models; Domain-Referenced Testing; Criterion-Referenced Testing

ABSTRACT

Descriptions are presented of two related probabilistic models that can be used for making classification decisions with respect to mastery of specific concepts or skills. Included are the development of procedures for: (a) assessing the adequacy of "fit" provided by the models; (b) identifying optimal decision rules for mastery classification; and (c) identifying minimally sufficient numbers of items necessary to obtain acceptable levels of misclassification.

INTRODUCTION

A major use for criterion-referenced tests is to provide information for assessing placement and advancement of students in instructional systems. Such decisions often are made by dichotomously classifying students with respect to mastery (master/non-master) of specific concepts or skills. Typically, the criterion for placement is based on students' responses to a sample of items from an "appropriately" defined domain. The purpose of this paper is to describe and illustrate two related probabilistic models which provide a foundation for: (a) assessing the adequacy of criterion-referenced tests and their constituent items;

(b) deciding upon necessary test length and other characteristics during test construction or revision phases; and (c) identifying optimal decision rules for mastery classification. In addition, statistical procedures are presented for judging the adequacy of the probabilistic models themselves with respect to "fitting" actual test data.

A difficulty in attempting to classify students correctly with respect to mastery stems from item domains which are inappropriately defined for this purpose. If the domain of items which is used in assessing mastery is very broad in scope, then correct responses to different items within the domain may be dependent on the acquisition of different skills. Thus, a dichotomous classification system for mastery is, at best, an oversimplification that obscures what a student has acquired. Macready and Merwin (1973), as well as Harris (1974), have suggested that the construction and revision of item domains should be based on the homogeneity of item content and the internal consistency of examinee responses, so that it is more reasonable to assume that mastery of all items within a domain is an all-or-none process.

Another factor in mastery classification which is not addressed by many classification procedures (e.g., Crehan, 1973; Kriewall, 1969; and Swaminathan, Hambleton and Algina, 1975) is error of classification in assigning item scores. Consideration of this kind of error in the assessment of mastery may be important since a major goal is to identify examinees' true states of skill acquisition. The failure of an examinee to respond correctly to an item selected from an item domain may be conceptualized as being due to one of two factors: (a) the examinee has not acquired the skill defining the domain and (appropriately) has missed the item; or, (b) the examinee has acquired the skill but is unable to produce a correct response due to short-term environmental or psychological factors (e.g., he is a victim of "forgetting"). Conversely, an examinee may respond correctly to the item: (a) because he has acquired the skill defining the domain; or, (b) because he "guesses" the correct response even though he has not acquired the defining skill (this is meant to include the case in which an examinee recalls specific information enabling him to respond correctly to the item even though he has not acquired the general skill underlying all items in the domain). The models developed in this paper attempt to address the above problems by incorporating within their

structure the "desired" nature of item domains and a consideration of classification errors.

THE MODELS

Two related probabilistic models which provide probability estimates of the 2^n possible response patterns on a dichotomously scored, n-item test are developed. The probability of the occurrence of each response pattern is defined in terms of the proportions of individuals who are masters or non-masters, and probabilities of "guessing" or "forgetting" which are associated with the individual test items.

Both models assume that all examinees belong to one of two possible "true score types" for any given domain: masters (M); and non-masters ($\bar{M}$). Masters are those individuals who have acquired the necessary skills to respond correctly to all items within the domain. Thus, for 4 items sampled from a domain, a master's true score response pattern would be 1 1 1 1, where a "one" indicates a correct response to an item. Conversely, non-masters have not acquired the necessary skills to respond correctly to any item within the domain; thus their true score response pattern would be 0 0 0 0, where a "zero" indicates an incorrect response to an item. This dichotomous classification of individuals appears reasonable to the degree that all items within a domain involve the same skills. However, this definition differs from those used in "threshold" models such as that presented by Swaminathan, Hambleton, and Algina (1975).

In general, it is assumed that the only way that any non-true score response pattern can occur is for a non-master to make one or more "guessing-errors" or for a master to make one or more "forgetting-errors." For the first model (Model I), the error probabilities are unrestricted except for the usual 0,1 bounds for probabilities. Thus, we let α_i and β_i represent the probabilities of a "guessing-error" and "forgetting-error," respectively, for item i. Further, Θ and $\bar{\Theta}$ represent the proportions of examinees who are masters and non-masters, respectively, with the usual restrictions of: $0 \leq \Theta \leq 1$ and $\bar{\Theta} + \Theta = 1$. If local independence among responses is assumed (i.e., the occurrence or non-occurrence of "guessing" or "forgetting" is assumed to be independent across items) then the probability of the j^{th} observed response pattern on an n-item test under Model I is:

$$P(j) = P(j \cap \bar{M}) + P(j \cap M) \qquad (1)$$

$$= P(j|\bar{M})\ P(\bar{M}) + P(j|M)\ P(M)$$

$$= \left[\prod_{i=1}^{n} \alpha_i^{a_{ij}} (1 - \alpha_i)^{1-a_{ij}}\right] \bar{\theta} +$$

$$\left[\prod_{i=1}^{n} \beta_i^{1-a_{ij}} (1 - \beta_i)^{a_{ij}}\right] \theta \ ,$$

where: $a_{ij} = \{0,1\}$ is the score on the i^{th} item for the j^{th} response pattern. Thus, for 4 items, the probability of the response pattern 0 1 1 0 occurring is:

$$P(0\ 1\ 1\ 0) = (1 - \alpha_1)\ \alpha_2\ \alpha_3\ (1 - \alpha_4)\ \bar{\theta} +$$

$$\beta_1\ (1 - \beta_2)\ (1 - \beta_3)\ \beta_4\ \theta \ .$$

In general, this model contains 2n+1 independent parameters. Maximum likelihood estimates of these parameters can be obtained by means of the Newton-Raphson iteration procedure, which is utilized because it is not feasible to obtain explicit formulae in terms of sufficient statistics. Furthermore, for cases such as the present one where probabilities can be modeled in terms of multinomial categories, a great deal of computational simplification is introduced by using Fisher's method of scoring. The Fisher method avoids direct computation of the likelihood function for the sample and arrives at the first and second derivatives of the likelihood function by utilizing the partial derivatives of equation (1) with respect to the parameters. Rao (1965, Pp. 302 - 309) provides a detailed presentation and illustration of the method. Assuming that the model is based on q parameters, let $\underline{\phi} = (\phi_{i1} \ \cdots \ \phi_{iq})'$ be a qxl vector of parameter estimates at step i of the iterative procedure. Also, at step i, let L represent the likelihood of the observations (conditional on ϕ), then $\underline{d}_i = (\partial \log_e L/\partial \phi_{i1} \ \cdots \ \partial \log_e L/\partial \phi_{iq})'$, and D_i is a qxq matrix with general element $\{\partial^2 \log_e L/((\partial \phi_{ij})(\partial \phi_{ik}))\}$ in the j,k location. The iterative

procedure involves solving $\underline{\phi}_{i+1} = \underline{\phi}_i - D_i^{-1}\underline{d}_i$ for successive steps until a convergence function such as $\Sigma|\phi_{ij} - \phi_{i-1,j}|$ is sufficiently small (e.g., 10^{-7}). After convergence is reached, the elements $-D^{-1}$ provide estimates of sampling variances and covariances for the maximum likelihood estimators, which are useful for setting up confidence intervals for the parameters in question.

Because of the relatively large number of parameters (2n+1) under Model I, there are circumstances in which it is desirable to utilize a more restrictive set of assumptions when modeling criterion-referenced test data. Toward this end, we introduce Model II which includes the additional assumptions that "guessing-errors" for all items are equal (i.e., $\alpha_i = \alpha$) and that "forgetting-errors" for all items are equal (i.e., $\beta_i = \beta$). These assumptions reduce the number of parameters to be estimated to three for tests composed of any number of items and allow for a simplification of the formula defining the probability of the occurrence of the j^{th} response pattern on an n-item test to:

$$P(j) = P(j \cap \bar{M}) + P(j \cap M) \qquad (2)$$

$$= P(j|\bar{M})\ P(\bar{M}) + P(j|M)\ P(M)$$

$$= \left[\alpha^{s_j} (1 - \alpha)^{n-s_j}\right] \bar{\theta} + \left[\beta^{n-s_j} (1 - \beta)^{s_j}\right] \theta\ ,$$

where s_j is the number of correct responses (i.e., 1's) in the j^{th} response pattern. For example, under Model II, the probability of response pattern 0 1 1 0 occurring is:

$$P(0\ 1\ 1\ 0) = \alpha^2 (1 - \alpha)^2 \bar{\theta} + \beta^2 (1 - \beta)^2 \theta.$$

Under Model II, it is apparent that every score pattern, j, with the same number of 1's has the same value for each $P(j \cap \bar{M})$ and for each $P(j \cap M)$. Thus, Model II requires only the frequencies for "number of correct responses" in order to generate maximum likelihood estimators of model parameters. Model I, on the other hand, requires information on the frequency of occurrence of the 2^n distinct response patterns.

Of course, there are other models which are appropriate if specific restrictions are placed on the α_i and/or β_i

parameters. For example, with n = 4 an alternative model might incorporate the following restrictions: $\alpha_1 = \alpha_2 = .25$, $\alpha_3 = \alpha_4 = 0$ and $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta$. While the methodology presented for Models I and II can be generalized to these alternative models, such alternatives are not seen as having as much general applicability to criterion-referenced testing as Models I and II.[1]

Emrick and Adams (1969), as well as Emrick (1971), have discussed a procedure for mastery classification which incorporates assumptions equivalent to Model II. Similarly, Besel (1973, 1975) has presented a model with assumptions equivalent to those of Model I. However, in none of these previous sources have general-purpose statistical procedures been presented for estimating parameters that arise in the models (i.e., previous procedures involve subjectivity in the estimation of parameters). The present approach to these models provides not only a self-contained system for estimating parameters, but, as presented in ensuing sections, procedures for assessing fit of the models to data, for setting up mastery classification rules, and for determining minimally sufficient test lengths.

ASSESSING ADEQUACY OF THE MODELS

The effectiveness of Models I and II in providing an acceptable representation of examinees' true states of mastery may be based on both statistical and judgmental assessments. Assuming that the necessary parameters have been estimated, a chi-square goodness of fit test may be run by utilizing equation (1) or (2) to obtain expected frequencies for the 2^n response patterns. Since it is necessary to estimate a total of 2n + 1 parameters for Model I, and 3 parameters for Model II, the respective degrees of freedom under the models are $2^n - 2n - 2$ and $2^n - 4$. Because of the dependency of the estimated parameters on the sample data it is desirable to assess fit on cross-validation samples; for such tests, the degrees of freedom are $2^n - 1$.

[1]Single copy listings are available of the following FORTRAN IV programs: Program MODEL3G which provides analyses under Model I related to parameter estimation, model fit, and establishing classification and program MODEL3 which provides similar analyses under Model II.

In testing for fit, one potential source of lack of fit is the untenability of the assumption that mastery is "all-or-none." Thus when fit is not obtained, a possible strategy is to subdivide the item domain into two or more new domains which are more restricted in scope and then try again to obtain fit. The process of domain revision may be facilitated by use of task analysis (Gagné, 1962) in which the component skills of a final behavior are identified and a hierarchy describing the conditional relations among the skills is established on the basis of expert judgment. Statistical assessment of such hierarchies can be made by extensions of the current models (Dayton and Macready, 1976). These more general models allow for true score response patterns which are intermediate between mastery and non-mastery (e.g., linear hierarchies).

Comparisons of relative fit provided by Models I and II for a given set of data can be made by means of a chi-square statistic following the procedure presented by Rao (1965). This is done by computing the expected frequencies, E_{Ij} and E_{IIj}, for the j^{th} response pattern under Models I and II respectively. Then a chi-square statistic may be defined as:

$$\sum_{j=1}^{2^n} (E_{Ij} - E_{IIj})^2 / E_{Ij} ,$$

with $2n - 2$ degrees of freedom. This procedure can be used to decide if the additional restrictions of Model II are tenable.

In addition to statistical tests of fit, it is desirable to assess judgmentally the estimated parameters, especially if the model yields a small (non-significant) value for the chi-square goodness-of-fit test. Inspection of the parameter estimates, along with their standard errors, may reveal logical inadequacies in the values obtained. For example, the fit of a model with respect to reproducing observed frequencies may be at the expense of unreasonably large values for the "guessing" and/or "forgetting" parameters. Thus, the investigator may wish to reject the notion of "mastery/non-mastery" for a domain. Under such circumstances, as in the case of statistical lack of fit, it may be desirable to consider subdividing or otherwise restructuring the item domain in order better to meet the assumption that mastery is an "all-or-none" process.

ESTABLISHING DECISION RULES FOR MASTERY/NON-MASTERY CLASSIFICATION

The maximum likelihood estimates of the parameters underlying Models I or II may be used to estimate $P(j \cap \bar{M})$ and $P(j \cap M)$, using equation (1) or (2). These are, respectively, the estimated proportions of examinees who obtain the j^{th} response pattern and are (a) non-masters or (b) masters. A comparison of these values for any given response pattern indicates the theoretical proportion of classification errors resulting from the two possible classification decisions for individuals obtaining that pattern. Thus, when a criterion of minimum expected "cost of misclassification" is incorporated, examinees who obtain response pattern j are classified as non-masters if $\frac{c_1 \hat{P}(j \cap \bar{M})}{c_2 \hat{P}(j \cap M)} > 1.0$ and as masters if $\frac{c_1 \hat{P}(j \cap \bar{M})}{c_2 \hat{P}(j \cap M)} < 1.0$ where c_1 and c_2 are, respectively, the costs of misclassifying a non-master as a master and a master as a non-master. Note that under this strategy, whenever c_1 is set equal to c_2, the expected proportion of "misclassified examinees" is minimized. Recall that under Model II, $P(j \cap \bar{M}) = P(j' \cap \bar{M})$ and $P(j \cap M) = P(j' \cap M)$ for all response patterns j and j' with equal numbers of correct responses. This implies that under Model II mastery classification may be specified in terms of total number of correct responses. It should be noted that the procedure for identifying test cut scores for differentiating between masters and non-masters developed by Emrick and Adams (1969), (also Emrick, 1971) yields the same classification decisions when equivalent parameter estimates are used (given that $\hat{\alpha}, \hat{\beta} \leq .5$).

Once classification decisions have been made for all 2^n response patterns, it is possible to assess classification errors. This may be done by estimating the proportion of examinees who are masters but are classified as non-masters, or who are non-masters but are classified as masters. This proportion of misclassified examinees is estimated as

$$\sum_{j=1}^{2^n} [b_j \hat{P}(j \cap \bar{M}) + (1 - b_j) \hat{P}(j \cap M)] \quad \text{where } b_j = 0$$

if examinees obtaining response pattern j are classified as non-masters and as 1 if they are classified as masters.

Similarly the estimated expected cost of misclassification for a single examinee is equal to

$$\sum_{j=1}^{2^n} [c_1 \, b_j \, \hat{P}(j \cap \bar{M}) + c_2 \, (1 - b_j) \, \hat{P}(j \cap M)].$$

SPECIFICATION OF MINIMALLY SUFFICIENT NUMBER OF ITEMS

The strategy presented for identifying the number of items needed for "adequate" mastery classification requires a specification of the maximum acceptable proportion of "misclassified examinees" and the loss ratio, c_1/c_2, to be used in making classification decisions. It is then possible to estimate the minimally sufficient number of items necessary for the proportion of misclassified examinees to be at or below the specified maximum acceptable level. Under Model II this may be done by the following steps:

1. Specify a value for the loss ratio, c_1/c_2.

2. Establish the maximum acceptable expected proportion of "misclassified examinees."

3. Establish estimates of the model parameters, Θ, α, and β.

4. Utilize the values in steps 1 and 3 to calculate

$$\frac{c_1 \, \hat{P}(j \cap \bar{M})}{c_2 \, \hat{P}(j \cap M)} \quad \text{for } j = 1, \ldots, 2^n \text{ , where}$$

$$\hat{P}(j \cap \bar{M}) = \hat{\alpha}^{s_j}(1 - \hat{\alpha})^{n - s_j}\hat{\Theta} \text{ ,}$$

$$\hat{P}(j \cap M) = \hat{\beta}^{n - s_j}(1 - \hat{\beta})^{s_j}\hat{\bar{\Theta}} \text{ ,}$$

and n is the number of items (initially, n = 1).

5. Establish classification decisions for each of the 2^n response patterns (using the rules in the previous section of this paper).

6. Compute the value for the expected proportion of "misclassified examinees" and compare it with the criterion established in step 2.

7. Repeat steps 4 through 6 for a test with one additional item (i.e., replace n by n + 1 in Steps 4 and 5) until a test length is reached which provides an expected "proportion of misclassified examinees" less than the established criterion set in Step 2.

Note that no item is needed for acceptable mastery classification in those cases in which the maximum acceptable proportion of misclassified examinees is greater than or equal to:

(a) $\hat{\theta}$ if $\frac{c_1}{c_2} \frac{\hat{P}(j \cap \overline{M})}{\hat{P}(j \cap M)} \geq 1.0$, or

(b) $\hat{\bar{\theta}}$ if $\frac{c_1}{c_2} \frac{\hat{P}(j \cap \overline{M})}{\hat{P}(j \cap M)} \leq 1.0$

for all response patterns on a one item test. Under these circumstances, acceptable error rates may be obtained by classifying all examinees as non-masters given (a) or as masters given (b).

The above steps may also be used for identifying minimally sufficient numbers of items under Model I. The ordering of the items for inclusion in steps 4 through 7 above should be based on the weighted (by costs) relative "efficiencies" of the items for correctly classifying examinees with respect to state of mastery. Thus, item ordering for test inclusion may be specified by the ordered ranks of $c_1(1 - \hat{\alpha}_i) + c_2(1 - \hat{\beta}_i)$, given that $\hat{\alpha}_i$ and $\hat{\beta}_i$ are less than .5. In practice, estimates for the guessing and forgetting parameters are available often from previous research based on, say, n', items. These estimates may be used in the ordering of items for inclusion. If the above steps result in a minimally sufficient number of items which is greater than n', the arithmetic mean (or median) of the available $\hat{\alpha}_i$ and $\hat{\beta}_i$ values may be used as approximate estimates for the necessary additional items.

Based on the steps described above, Table I provides minimally sufficient numbers of items (the non-parenthesized tabled values) necessary for a variety of specified conditions under Model II. To identify the appropriate tabled values, the user must specify the following: (a) the proportion of non-masters in the population, $\bar{\theta}$ (designated: THETA); (b) the relative costs of classification errors, c_1/c_2 (designated: LOSS RATIO); (c) the maximum acceptable proportion of mis-

TABLE I

Minimally Sufficient Numbers of Items and Cutting Scores for Mastery Classification

THETA = .050

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	7(4)	7(4)	9(6)	14(10)	20(15)	8(4)	9(5)	12(7)	17(11)	26(18)	10(4)	14(6)	21(11)	30(17)	50(31)
.20	.001	6(3)	7(4)	8(5)	12(8)	18(13)	7(3)	8(4)	12(7)	16(10)	25(17)	9(3)	12(5)	20(10)	29(16)	45(27)
1.0	.001	6(3)	6(3)	8(5)	12(8)	17(12)	7(3)	8(4)	11(6)	16(10)	23(15)	9(3)	12(5)	18(8)	27(14)	43(25)
5.0	.001	6(3)	6(3)	8(5)	12(8)	17(12)	7(3)	9(4)	11(6)	17(10)	23(15)	9(3)	13(5)	18(8)	27(14)	45(26)
20.	.001	7(3)	8(4)	10(6)	13(8)	19(13)	8(3)	9(4)	14(7)	19(11)	27(17)	10(3)	14(5)	21(9)	30(15)	48(27)
.05	.01	3(2)	3(2)	6(4)	7(5)	8(6)	4(2)	5(3)	8(5)	9(6)	13(9)	5(2)	7(3)	13(7)	16(9)	26(16)
.20	.01	2(1)	3(2)	5(3)	6(4)	7(5)	4(2)	4(2)	7(4)	8(5)	12(8)	5(2)	6(3)	10(5)	13(7)	22(13)
1.0	.01	2(1)	3(2)	4(3)	5(4)	6(5)	3(1)	4(2)	6(3)	7(4)	10(6)	5(2)	6(2)	10(4)	13(6)	20(11)
5.0	.01	2(1)	4(2)	4(2)	5(3)	8(5)	3(1)	5(2)	6(3)	9(5)	10(6)	6(2)	6(2)	10(4)	13(6)	21(11)
20.	.01	3(1)	4(2)	4(2)	7(4)	8(5)	3(1)	5(2)	7(3)	10(5)	14(8)	7(2)	7(2)	11(4)	16(7)	25(13)
.05	.05	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
.20	.05	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
1.0	.05	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
5.0	.05	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
20.	.05	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
.05	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
.20	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
1.0	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
5.0	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
20.	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)

THETA = .100

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	7(4)	8(5)	9(6)	14(10)	21(16)	8(4)	9(5)	13(8)	21(14)	27(19)	10(4)	15(7)	21(11)	33(19)	53(33)
.20	.001	6(3)	7(4)	9(6)	13(9)	19(14)	7(3)	8(4)	12(7)	19(12)	25(17)	10(4)	12(5)	20(10)	30(18)	47(30)
1.0	.001	6(3)	7(4)	9(6)	13(9)	18(13)	7(3)	8(4)	12(7)	18(11)	25(17)	10(4)	12(5)	19(9)	29(16)	47(29)
5.0	.001	6(3)	7(4)	10(6)	13(9)	19(13)	7(3)	9(4)	13(7)	18(11)	26(17)	11(4)	13(5)	19(9)	30(16)	48(28)
20.	.001	6(3)	8(4)	10(6)	14(9)	20(14)	7(3)	9(4)	13(7)	20(12)	28(18)	12(4)	13(5)	22(10)	33(17)	54(31)
.05	.01	3(2)	3(2)	7(5)	8(6)	13(10)	4(2)	5(3)	8(5)	12(8)	18(13)	5(2)	8(4)	13(7)	20(12)	30(19)
.20	.01	3(2)	3(2)	6(4)	7(5)	11(9)	4(2)	4(2)	7(4)	11(7)	13(9)	5(2)	7(3)	10(5)	16(9)	24(16)
1.0	.01	3(2)	3(2)	5(3)	6(4)	10(7)	4(2)	4(2)	7(4)	10(6)	12(8)	5(2)	7(3)	10(5)	15(8)	24(14)
5.0	.01	4(2)	4(2)	5(3)	8(5)	10(7)	4(2)	4(2)	8(4)	10(6)	14(9)	6(2)	8(3)	11(5)	16(8)	25(14)
20.	.01	4(2)	4(2)	5(3)	8(5)	12(8)	5(2)	5(2)	8(4)	11(6)	15(9)	7(2)	9(3)	12(5)	19(9)	30(16)
.05	.05	1(1)	3(2)	4(3)	4(3)	5(4)	3(2)	3(2)	3(2)	7(5)	8(6)	2(1)	4(2)	7(4)	10(6)	17(11)
.20	.05	1(1)	3(2)	3(2)	3(2)	4(3)	2(1)	2(1)	2(1)	3(2)	6(4)	2(1)	3(2)	4(2)	6(3)	10(6)
1.0	.05	1(1)	2(1)	3(2)	3(2)	3(2)	2(1)	2(1)	2(1)	3(2)	4(3)	2(1)	3(1)	4(2)	4(2)	8(4)
5.0	.05	2(1)	2(1)	3(2)	4(3)	5(3)	2(1)	2(1)	3(2)	4(2)	6(3)	3(1)	3(1)	5(2)	7(3)	10(5)
20.	.05	2(1)	2(1)	4(2)	5(3)	8(5)	3(1)	3(1)	3(2)	5(2)	6(3)	4(1)	4(1)	6(2)	8(3)	13(6)
.05	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
.20	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
1.0	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
5.0	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
20.	.10	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)

THETA = .250

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	7(4)	8(5)	10(7)	15(11)	22(17)	8(4)	9(5)	16(10)	22(15)	32(25)	12(5)	15(7)	24(13)	37(22)	58(37)
.20	.001	7(4)	8(5)	10(7)	14(10)	20(15)	8(4)	9(5)	15(9)	20(13)	27(19)	10(4)	15(7)	21(11)	33(19)	53(33)
1.0	.001	6(3)	7(4)	10(7)	14(10)	19(14)	8(4)	9(5)	14(8)	19(12)	27(19)	10(4)	14(6)	21(11)	32(18)	51(31)
5.0	.001	6(3)	7(4)	11(7)	14(10)	19(14)	9(4)	10(5)	14(8)	19(12)	28(19)	11(4)	14(6)	23(11)	33(18)	52(31)
20.	.001	6(3)	7(4)	11(7)	15(10)	21(15)	9(4)	10(5)	16(9)	21(13)	30(20)	11(4)	15(6)	23(11)	36(19)	56(33)
.05	.01	3(2)	6(4)	7(5)	9(7)	14(11)	5(3)	5(3)	9(6)	13(9)	19(14)	7(3)	8(4)	14(8)	23(14)	34(22)
.20	.01	3(2)	5(3)	6(4)	8(6)	13(10)	4(2)	5(3)	8(5)	12(8)	17(12)	7(3)	8(4)	13(7)	19(11)	30(19)
1.0	.01	3(2)	5(3)	6(4)	8(6)	11(8)	4(2)	5(3)	8(5)	11(7)	16(11)	6(2)	7(3)	12(6)	18(10)	28(17)
5.0	.01	4(2)	5(3)	7(4)	9(6)	11(8)	5(2)	6(3)	9(5)	11(7)	17(11)	6(2)	10(4)	13(6)	19(10)	30(17)
20.	.01	4(2)	5(3)	7(4)	9(6)	13(9)	5(2)	6(3)	9(5)	14(8)	17(11)	6(2)	10(4)	13(6)	22(11)	35(20)
.05	.05	1(1)	4(3)	4(3)	5(4)	7(6)	3(2)	3(2)	4(3)	8(6)	9(7)	4(2)	5(3)	8(5)	14(9)	21(14)
.20	.05	1(1)	3(2)	4(3)	4(3)	5(4)	2(1)	3(2)	3(2)	7(5)	8(6)	4(2)	4(2)	7(4)	10(6)	14(9)
1.0	.05	1(1)	2(1)	3(2)	4(3)	5(4)	2(1)	3(2)	3(2)	5(3)	8(5)	3(1)	4(2)	6(3)	9(5)	13(8)
5.0	.05	2(1)	2(1)	4(2)	5(3)	6(4)	2(1)	4(2)	4(2)	7(4)	10(6)	3(1)	6(2)	9(3)	13(6)	19(10)
.05	.10	1(1)	1(1)	2(2)	2(2)	7(6)	1(1)	1(1)	4(3)	5(4)	6(5)	3(2)	3(2)	6(4)	9(6)	16(11)
.20	.10	1(1)	1(1)	1(1)	2(2)	5(4)	1(1)	1(1)	3(2)	4(3)	4(3)	2(1)	2(1)	5(3)	6(4)	9(6)
1.0	.10	1(1)	1(1)	1(1)	2(2)	3(3)	1(1)	2(1)	2(1)	3(2)	4(3)	2(1)	2(1)	4(2)	5(3)	7(4)
5.0	.10	1(1)	1(1)	1(1)	2(2)	3(2)	2(1)	2(1)	2(1)	3(2)	5(3)	2(1)	2(1)	4(2)	6(3)	9(5)
20.	.10	2(1)	2(1)	2(1)	3(2)	3(2)	2(1)	2(1)	2(1)	4(2)	7(4)	3(1)	3(1)	5(2)	7(3)	12(6)

TABLE I (continued)

THETA = .400

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	7(4)	9(6)	13(9)	19(12)	23(18)	9(5)	12(7)	17(11)	22(15)	33(24)	12(5)	16(8)	24(13)	37(22)	61(39)
.20	.001	7(4)	8(5)	10(7)	15(11)	21(16)	8(5)	9(5)	15(10)	21(14)	31(22)	10(4)	15(7)	21(11)	34(20)	54(34)
1.0	.001	6(3)	8(5)	10(7)	14(10)	20(15)	8(4)	9(5)	15(9)	20(13)	29(20)	10(4)	14(6)	21(11)	33(19)	52(32)
5.0	.001	6(3)	8(5)	11(7)	14(10)	20(15)	8(4)	9(5)	15(9)	20(13)	29(20)	10(4)	14(6)	22(11)	34(19)	55(33)
20.	.001	6(3)	9(5)	11(7)	16(11)	22(16)	9(4)	10(5)	16(9)	23(14)	31(21)	11(4)	15(6)	25(12)	37(20)	57(34)
.05	.01	3(2)	4(3)	8(6)	9(7)	15(12)	5(3)	8(5)	9(6)	14(10)	20(15)	8(4)	10(5)	14(8)	21(13)	35(23)
.20	.01	3(2)	4(3)	7(5)	8(6)	14(11)	4(2)	5(3)	8(5)	13(9)	18(13)	7(3)	8(4)	13(7)	20(12)	31(20)
1.0	.01	3(2)	4(3)	6(4)	8(6)	12(9)	4(2)	5(3)	8(5)	12(8)	16(11)	7(3)	8(4)	13(7)	19(11)	29(18)
5.0	.01	3(2)	5(3)	6(4)	8(6)	12(9)	4(2)	5(3)	8(5)	12(8)	16(11)	7(3)	9(4)	14(7)	20(11)	32(19)
20.	.01	4(2)	5(3)	6(4)	10(7)	14(10)	4(2)	6(3)	9(5)	13(8)	18(12)	8(3)	10(4)	15(7)	23(12)	36(21)
.05	.05	1(1)	4(3)	5(4)	6(5)	7(6)	3(2)	4(3)	7(5)	8(6)	10(8)	5(3)	5(3)	8(5)	12(8)	22(15)
.20	.05	1(1)	3(2)	4(3)	5(4)	6(5)	3(2)	3(2)	4(3)	7(5)	8(6)	4(2)	5(3)	7(4)	11(7)	18(12)
1.0	.05	1(1)	3(2)	3(2)	4(3)	5(4)	2(1)	3(2)	4(3)	6(4)	8(6)	4(2)	4(2)	6(3)	10(6)	14(9)
5.0	.05	1(1)	3(2)	3(2)	4(3)	5(4)	2(1)	3(2)	4(3)	6(4)	9(5)	5(2)	5(2)	6(3)	11(6)	15(9)
20.	.05	2(1)	3(2)	3(2)	6(4)	7(5)	2(1)	4(2)	5(3)	8(5)	11(7)	5(2)	5(2)	9(4)	12(6)	22(12)
.05	.10	1(1)	2(2)	2(2)	3(3)	7(6)	1(1)	4(3)	5(4)	6(5)	7(6)	3(2)	3(2)	6(4)	10(7)	17(12)
.20	.10	1(1)	1(1)	2(2)	2(2)	6(5)	1(1)	1(1)	4(3)	4(3)	5(4)	2(1)	3(2)	5(3)	6(4)	10(7)
1.0	.10	1(1)	1(1)	2(2)	2(2)	4(3)	1(1)	1(1)	3(2)	4(3)	5(4)	2(1)	3(2)	4(2)	5(3)	9(6)
5.0	.10	1(1)	2(1)	3(2)	3(2)	5(4)	2(1)	2(1)	4(2)	5(3)	8(5)	3(1)	5(2)	7(3)	10(5)	13(7)
20.	.10	2(1)	2(1)	3(2)	3(2)	6(4)	2(1)	2(1)	4(2)	5(3)	8(5)	3(1)	5(2)	7(3)	10(5)	13(7)

THETA = .500

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	8(5)	9(6)	11(8)	16(12)	23(18)	9(5)	10(6)	14(9)	23(16)	33(24)	13(6)	16(8)	25(14)	35(21)	59(38)
.20	.001	7(4)	8(5)	10(7)	15(11)	22(17)	8(4)	9(5)	14(9)	21(14)	28(20)	10(4)	15(7)	22(12)	34(20)	55(35)
1.0	.001	7(4)	8(5)	10(7)	15(11)	20(15)	8(4)	9(5)	14(9)	20(13)	28(20)	10(4)	14(6)	21(11)	33(19)	52(32)
5.0	.001	7(4)	8(5)	10(7)	15(11)	20(15)	8(4)	9(5)	15(9)	20(13)	29(20)	10(4)	14(6)	22(11)	34(19)	54(33)
20.	.001	8(4)	9(5)	13(8)	16(11)	22(16)	9(4)	10(5)	16(9)	22(14)	31(21)	11(4)	14(6)	24(12)	39(21)	60(36)
.05	.01	4(3)	4(3)	8(6)	10(8)	16(13)	5(3)	6(4)	10(7)	14(10)	20(15)	6(3)	11(6)	15(9)	24(15)	36(24)
.20	.01	3(2)	4(3)	7(5)	8(6)	14(11)	5(3)	5(3)	9(6)	13(9)	19(14)	6(3)	8(4)	14(8)	20(12)	31(20)
1.0	.01	3(2)	4(3)	6(4)	8(6)	12(9)	4(2)	5(3)	8(5)	12(8)	17(12)	6(3)	8(4)	13(7)	19(11)	30(19)
5.0	.01	3(2)	5(3)	6(4)	8(6)	12(9)	4(2)	5(3)	8(5)	12(8)	17(12)	7(3)	9(4)	14(7)	20(11)	31(19)
20.	.01	4(2)	5(3)	6(4)	10(7)	12(9)	4(2)	6(3)	11(6)	13(8)	21(14)	8(3)	10(4)	15(7)	24(13)	37(22)
.05	.05	2(2)	4(3)	5(4)	6(5)	8(7)	3(2)	4(3)	5(4)	9(7)	11(9)	5(3)	5(3)	9(6)	15(10)	20(14)
.20	.05	1(1)	3(2)	4(3)	5(4)	6(5)	3(2)	3(2)	4(3)	8(6)	9(7)	4(2)	5(3)	8(5)	11(7)	18(12)
1.0	.05	1(1)	3(2)	4(3)	4(3)	6(5)	3(2)	3(2)	4(3)	6(4)	8(6)	4(2)	4(2)	7(4)	10(6)	15(10)
5.0	.05	1(1)	3(2)	4(3)	4(3)	6(5)	3(2)	3(2)	5(3)	6(4)	10(7)	4(2)	4(2)	8(4)	11(6)	18(11)
20.	.05	2(1)	3(2)	5(3)	6(4)	8(6)	4(2)	4(2)	5(3)	8(5)	12(8)	5(2)	5(2)	9(4)	14(7)	21(12)
.05	.10	2(2)	2(2)	2(2)	3(3)	8(7)	2(2)	2(2)	5(4)	6(5)	7(6)	3(2)	4(3)	7(5)	10(7)	15(11)
.20	.10	1(1)	1(1)	2(2)	2(2)	6(5)	1(1)	1(1)	4(3)	5(4)	6(5)	2(2)	3(2)	6(4)	7(5)	11(8)
1.0	.10	1(1)	1(1)	2(2)	2(2)	6(3)	1(1)	1(1)	3(2)	4(3)	5(4)	2(1)	3(2)	5(3)	6(4)	9(6)
5.0	.10	1(1)	1(1)	3(2)	4(3)	6(4)	2(1)	2(1)	4(2)	5(3)	8(5)	2(1)	5(2)	7(3)	10(5)	16(9)

THETA = .600

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	6(4)	9(6)	11(8)	17(13)	24(19)	9(5)	10(6)	15(10)	23(16)	30(22)	11(5)	16(8)	25(14)	38(23)	60(39)
.20	.001	6(4)	8(5)	10(7)	15(11)	22(17)	8(4)	9(5)	14(9)	22(15)	29(21)	11(5)	15(7)	23(12)	34(20)	55(35)
1.0	.001	6(4)	8(5)	10(7)	15(11)	21(16)	8(4)	9(5)	14(9)	20(13)	28(20)	10(4)	15(7)	21(11)	33(19)	52(33)
5.0	.001	7(4)	8(5)	12(8)	15(11)	21(16)	8(4)	9(5)	15(9)	20(13)	28(20)	10(4)	14(7)	21(11)	35(20)	54(33)
20.	.001	7(4)	9(5)	12(8)	16(11)	23(17)	9(4)	12(6)	16(9)	22(14)	32(22)	13(5)	17(7)	24(12)	36(20)	58(35)
.05	.01	4(3)	4(3)	8(6)	10(8)	12(10)	5(3)	6(4)	10(7)	15(11)	21(16)	6(3)	9(5)	15(9)	22(14)	36(24)
.20	.01	3(2)	4(3)	7(5)	9(7)	11(9)	5(3)	5(3)	9(6)	11(8)	16(12)	6(3)	8(4)	14(8)	21(13)	32(21)
1.0	.01	3(2)	4(3)	7(5)	8(6)	11(9)	4(2)	5(3)	8(5)	11(8)	16(12)	6(3)	8(4)	13(7)	19(11)	30(19)
5.0	.01	4(3)	5(3)	8(6)	8(6)	13(10)	4(2)	5(3)	8(5)	12(8)	17(12)	7(3)	8(4)	13(7)	19(11)	31(19)
20.	.01	3(2)	5(3)	8(5)	10(7)	15(11)	4(2)	8(4)	10(6)	14(9)	19(13)	8(3)	9(4)	14(7)	22(12)	35(21)
.05	.05	2(2)	2(2)	5(4)	7(6)	8(7)	3(2)	4(3)	5(4)	9(7)	11(9)	3(2)	6(4)	9(6)	13(9)	21(15)
.20	.05	1(1)	2(2)	5(4)	6(5)	7(6)	3(2)	3(2)	4(3)	8(6)	9(7)	3(2)	5(3)	6(4)	9(6)	16(11)
1.0	.05	1(1)	2(2)	4(3)	5(4)	6(5)	3(2)	3(2)	4(3)	6(4)	8(6)	3(2)	4(2)	6(4)	9(6)	15(10)
5.0	.05	1(1)	3(2)	4(3)	5(4)	6(5)	3(2)	3(2)	5(3)	6(4)	8(6)	4(2)	4(2)	7(4)	10(6)	16(10)
20.	.05	1(1)	3(2)	5(3)	6(4)	8(6)	4(2)	4(2)	5(3)	8(5)	12(8)	5(2)	7(3)	8(4)	13(7)	22(13)
.05	.10	2(2)	2(2)	3(3)	3(3)	4(4)	2(2)	2(2)	5(4)	6(5)	8(7)	3(2)	4(3)	7(5)	11(8)	16(12)
.20	.10	1(1)	1(1)	2(2)	2(2)	3(3)	1(1)	1(1)	2(2)	5(4)	6(5)	3(2)	3(2)	4(3)	7(5)	11(8)
1.0	.10	1(1)	1(1)	2(2)	2(2)	5(4)	1(1)	1(1)	2(2)	4(3)	5(4)	2(1)	3(2)	4(3)	6(4)	9(6)
5.0	.10	1(1)	1(1)	3(2)	3(2)	5(4)	2(1)	2(1)	4(2)	5(3)	7(4)	2(1)	4(2)	5(3)	7(4)	11(7)
20.	.10	1(1)	1(1)	3(2)	4(3)	7(5)	2(1)	4(2)	5(3)	7(4)	9(6)	2(1)	5(2)	6(3)	10(5)	14(8)

TABLE I (continued)

THETA = .750

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	6(4)	9(6)	11(8)	17(13)	21(17)	7(4)	10(6)	15(10)	21(15)	31(23)	11(5)	16(8)	23(13)	37(23)	58(38)
.20	.001	6(4)	9(6)	11(8)	16(12)	19(15)	7(4)	10(6)	14(9)	20(14)	30(22)	11(5)	14(7)	22(12)	33(20)	54(35)
1.0	.001	6(4)	8(5)	10(7)	15(11)	19(15)	7(4)	9(5)	14(9)	19(13)	28(20)	10(4)	14(7)	21(11)	32(19)	51(32)
5.0	.001	7(4)	8(5)	10(7)	15(11)	21(16)	8(4)	9(5)	15(9)	21(14)	28(20)	10(4)	15(7)	21(11)	33(19)	53(33)
20.	.001	7(4)	8(5)	12(8)	17(12)	21(16)	8(4)	9(5)	15(9)	22(14)	32(22)	10(4)	16(7)	24(12)	36(20)	59(36)
.05	.01	4(3)	5(4)	6(5)	11(9)	13(11)	5(3)	6(4)	10(7)	13(10)	18(14)	7(4)	9(5)	13(8)	23(15)	35(24)
.20	.01	3(2)	4(3)	6(5)	9(7)	11(9)	5(3)	6(4)	7(5)	11(8)	17(13)	6(3)	9(5)	12(7)	19(12)	30(20)
1.0	.01	3(2)	4(3)	6(5)	8(6)	11(9)	5(3)	5(3)	7(5)	11(8)	16(12)	6(3)	8(4)	12(7)	18(11)	28(18)
5.0	.01	3(2)	4(3)	7(5)	8(6)	11(9)	5(3)	5(3)	8(5)	12(8)	17(12)	6(3)	8(4)	13(7)	19(11)	30(19)
20.	.01	3(2)	5(3)	7(5)	8(6)	13(10)	6(3)	7(4)	8(5)	12(8)	19(13)	7(3)	9(4)	14(7)	23(13)	36(22)
.05	.05	2(2)	2(2)	3(3)	7(6)	9(8)	2(2)	4(3)	6(5)	7(6)	12(10)	4(3)	4(3)	7(5)	14(10)	19(14)
.20	.05	1(1)	2(2)	3(3)	6(5)	8(7)	2(2)	4(3)	5(4)	6(5)	10(8)	3(2)	4(3)	6(4)	10(7)	15(11)
1.0	.05	1(1)	2(2)	3(3)	5(4)	6(5)	2(2)	3(3)	4(3)	5(4)	8(6)	3(2)	3(2)	6(4)	9(6)	13(9)
5.0	.05	1(1)	2(2)	4(3)	5(4)	6(5)	2(2)	3(2)	4(3)	5(4)	8(6)	3(2)	4(2)	6(4)	9(6)	13(9)
20.	.05	1(1)	3(2)	4(3)	7(5)	8(6)	4(2)	4(2)	5(3)	8(5)	10(7)	4(2)	4(2)	8(4)	13(7)	20(12)
.05	.10	2(2)	2(2)	3(3)	4(4)	5(5)	2(2)	2(2)	3(3)	4(4)	9(8)	2(2)	2(2)	5(4)	9(7)	14(11)
.20	.10	1(1)	2(2)	2(2)	3(3)	4(4)	1(1)	2(2)	2(2)	3(3)	7(6)	1(1)	2(2)	4(3)	6(5)	10(8)
1.0	.10	1(1)	1(1)	2(2)	2(2)	3(3)	1(1)	1(1)	2(2)	3(3)	5(4)	1(1)	2(2)	3(3)	5(4)	8(6)
5.0	.10	1(1)	1(1)	2(2)	2(2)	3(3)	1(1)	1(1)	2(2)	4(3)	5(4)	1(1)	3(2)	3(3)	6(4)	9(6)
20.	.10	1(1)	1(1)	3(2)	4(3)	5(4)	1(1)	1(1)	3(2)	6(4)	7(5)	2(1)	4(2)	6(3)	9(5)	15(9)

THETA = .900

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	6(4)	7(5)	12(9)	15(12)	22(18)	8(5)	9(6)	13(9)	19(14)	29(22)	10(5)	13(7)	22(13)	33(21)	55(37)
.20	.001	6(4)	7(5)	11(8)	14(11)	20(16)	7(4)	8(5)	13(9)	18(13)	28(21)	10(5)	13(7)	19(11)	32(20)	47(31)
1.0	.001	6(4)	7(5)	10(7)	13(10)	19(15)	7(4)	8(5)	12(8)	18(13)	26(19)	9(4)	12(6)	19(11)	30(18)	47(31)
5.0	.001	6(4)	7(5)	10(7)	13(10)	19(15)	7(4)	8(5)	12(8)	19(13)	26(19)	9(4)	12(6)	20(11)	30(18)	49(31)
20.	.001	7(4)	8(5)	10(7)	15(11)	21(16)	8(4)	9(5)	15(9)	21(14)	30(21)	9(4)	13(6)	21(11)	33(19)	53(33)
.05	.01	4(3)	5(4)	7(6)	8(7)	14(12)	4(3)	5(4)	9(7)	11(9)	17(14)	5(3)	8(5)	12(8)	19(13)	31(22)
.20	.01	4(3)	4(3)	6(5)	7(6)	13(11)	4(3)	4(3)	8(6)	10(8)	15(12)	5(3)	8(5)	11(7)	16(11)	27(19)
1.0	.01	3(2)	4(3)	5(4)	7(6)	11(9)	3(2)	4(3)	7(5)	9(7)	14(11)	5(3)	7(4)	10(6)	15(10)	25(17)
5.0	.01	3(2)	4(3)	5(4)	7(6)	11(9)	3(2)	5(3)	7(5)	11(8)	15(11)	6(3)	7(4)	10(6)	16(10)	26(17)
20.	.01	3(2)	4(3)	5(4)	9(7)	13(10)	3(2)	5(3)	8(5)	12(8)	15(11)	7(3)	8(4)	13(7)	21(12)	32(20)
.05	.05	2(2)	3(3)	4(4)	5(5)	7(7)	2(2)	3(3)	4(4)	5(5)	10(9)	2(2)	3(3)	6(5)	10(8)	16(13)
.20	.05	2(2)	2(2)	3(3)	4(4)	5(5)	2(2)	2(2)	3(3)	4(4)	8(7)	2(2)	2(2)	5(4)	7(6)	11(9)
1.0	.05	1(1)	2(2)	2(2)	3(3)	4(4)	2(2)	2(2)	3(3)	4(4)	6(5)	2(2)	2(2)	4(3)	6(5)	9(7)
5.0	.05	1(1)	2(2)	3(2)	4(3)	5(4)	2(2)	3(2)	4(3)	5(4)	8(6)	3(2)	3(2)	5(3)	7(5)	11(8)
20.	.05	1(1)	3(2)	4(3)	5(4)	6(5)	3(2)	3(2)	4(3)	7(5)	8(6)	4(2)	4(2)	7(4)	10(6)	15(12)
.05	.10	2(2)	2(2)	3(3)	4(4)	5(5)	7(7)	2(2)	2(2)	3(3)	4(4)	5(5)	7(7)	2(2)	3(3)	4(4)	6(6)	8(8)
.20	.10	2(2)	2(2)	3(3)	4(4)	5(5)	6(6)	2(2)	2(2)	3(3)	4(4)	5(5)	7(6)	2(2)	3(3)	4(4)	6(6)	8(7)
1.0	.10	1(1)	1(1)	2(2)	3(3)	4(4)	5(5)	1(1)	1(1)	2(2)	3(3)	4(4)	5(5)	1(1)	2(2)	3(3)	4(4)	6(5)
5.0	.10	1(1)	1(1)	2(2)	3(3)	4(4)	5(5)	1(1)	1(1)	2(2)	3(3)	4(4)	5(5)	1(1)	2(2)	3(3)	4(4)	6(5)
20.	.10	1(1)	1(1)	2(2)	2(2)	5(4)	1(1)	1(1)	3(2)	4(3)	5(4)	1(1)	3(2)	5(3)	6(4)	9(6)

THETA = .950

LOSS RATIO	MAX ERROR RATE	AF05 AG05	AG10	AG20	AG30	AG40	AF10 AG05	AG10	AG20	AG30	AG40	AF20 AG05	AG10	AG20	AG30	AG40
.05	.001	7(5)	8(6)	10(8)	16(13)	19(16)	6(4)	9(6)	14(10)	20(15)	27(21)	10(5)	14(8)	21(13)	32(21)	50(34)
.20	.001	6(4)	7(5)	9(7)	14(11)	17(14)	6(4)	9(6)	13(9)	19(14)	25(19)	8(4)	11(6)	18(11)	28(18)	46(31)
1.0	.001	6(4)	7(5)	9(7)	13(10)	17(14)	6(4)	8(5)	12(8)	17(12)	24(18)	8(4)	11(6)	18(11)	27(17)	44(29)
5.0	.001	6(4)	7(5)	9(7)	13(10)	17(14)	7(4)	8(5)	12(8)	17(12)	24(18)	8(4)	12(6)	20(11)	28(17)	45(29)
20.	.001	7(4)	8(5)	10(7)	13(10)	19(15)	7(4)	9(5)	14(9)	19(13)	28(20)	9(4)	12(6)	21(11)	33(19)	49(31)
.05	.01	3(3)	3(3)	7(6)	9(8)	11(10)	4(3)	5(4)	7(6)	11(9)	14(12)	4(3)	7(5)	11(8)	18(13)	27(20)
.20	.01	2(2)	3(3)	5(5)	8(7)	10(9)	4(3)	5(4)	6(5)	10(8)	13(11)	4(3)	6(4)	10(7)	14(10)	23(17)
1.0	.01	2(2)	3(3)	5(4)	7(6)	9(8)	3(2)	4(3)	6(5)	9(7)	12(10)	4(3)	6(4)	9(6)	13(9)	21(15)
5.0	.01	2(2)	4(3)	5(4)	7(6)	9(8)	3(2)	4(3)	7(5)	9(7)	13(10)	5(3)	7(4)	10(6)	15(10)	22(15)
20.	.01	3(2)	4(3)	5(4)	9(7)	11(9)	3(2)	5(3)	7(5)	11(8)	15(11)	6(3)	8(4)	13(7)	18(11)	28(18)
.05	.05	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
.20	.05	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
1.0	.05	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
5.0	.05	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
20.	.05	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
.05	.10	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
.20	.10	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
1.0	.10	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
5.0	.10	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)
20.	.10	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)	0(1)

classified examinees (designated: MAX ERROR RATE); (d) the probability of a master making a "forgetting-error," α, (designated: AF); and (e) the probability of a non-master making a "guessing-error," β, (designated: AG).[2]

Note that for some of the sets of specified conditions, the minimally sufficient number of items is 0. This indicates that an acceptable level of misclassification can be obtained without testing by either classifying all examinees as non-masters (when the minimally acceptable test score for mastery classification is 1), or classifying all examinees as masters (when the minimally acceptable test score for mastery classification is 0).

The values found in Table I for minimally sufficient number of items may be used to obtain estimates (under Model II) for the values obtained using the iterative procedure described above. These estimates may be obtained by using for table entry, the closest tabled values of AF, AG, MAX ERROR RATE, THETA and LOSS RATIO in place of the actual values of interest.

EXAMPLES

In order to exemplify the use of the procedures described in earlier sections of this paper, results obtained on portions of a domain-referenced test dealing with multiplication of whole numbers are used. The analyses were based on item scores obtained on 4 randomly selected items from each of 2 domains for 284 fourth-grade students. The first domain (D_1) contains items involving integer multiplication in which: (a) the multiplier has 2 digits; (b) the multiplicand has either 3 or 4 digits; and (c) there are no "carry" operations involved. The second domain (D_2) contains items involving integer multiplication in which: both (a) and (b) above occur and there is at least one "carry" operation for each digit in the multiplier.

[2]Single copies of more complete tabled values may be obtained by writing the authors; this table includes entries for the following specified values: THETA = .05 through .95 by increments of .05; LOSS RATIO = .05, .20, 1.0, 5.0 and 20.0; MAX ERROR RATE = .001, .01, .05, .10, .15 and .20; AF = .01 and .05 through .35 by increments of .05; and AG = .01 and .05 through .40 by increments of .05.

The item scores from D_1 and D_2 for 142 randomly selected students (50% of the original sample) were used to generate maximum likelihood estimates of the parameters and their standard errors under Models I and II (presented in Table II), while the data for the remaining 142 students were used as a cross-validation sample.

Note that the estimated α's for D_1 and D_2 under both models are relatively small in magnitude (except $\hat{\alpha}_1$ for D_2) when compared to their standard errors and the corresponding $\hat{\beta}$-values. This outcome was expected since the items were presented in free-response format.

On the basis of the parameter estimates presented in Table II, expected frequencies corresponding to each of the 16 possible response patterns were generated. These expected frequencies along with the observed frequencies for both the validation and cross-validation samples are presented in Table III. These frequencies were used in the statistical assessment of fit provided by the models for both domains.

Table IV presents results of chi-square tests used in assessing both absolute and relative fit provided by Models I and II. Chi-square results related to model validation and cross-validation suggest reasonable absolute fit under both models for domain D_1; while for domain D_2, reasonable fit was obtained only under Model I.

Chi-square results related to relative fit provided by the two models suggest that for domain D_1 the simpler Model II provides a fit which is comparable to that obtained under Model I. Thus, Model II is to be preferred for use with this domain. Assessment of relative fit for domain D_2 resulted in significantly better fit under Model I. Thus, for this domain, the use of Model I is to be preferred. However, the large estimated value for β_3 under D_2 appears to be a logically unreasonable estimate. This suggests the possible need for subdividing or otherwise restructuring this domain.

In order to identify the desired mastery classification for students obtaining each response pattern, the values of $\hat{P}(j \cap \bar{M})$ and $\hat{P}(j \cap M)$ were computed and are presented in Table V.

If c_1 and c_2 are assigned equal values (i.e., misclassification errors of both types are equally costly) then, for Model I on domain D_1, a non-mastery designation would be assigned to those students attaining the response patterns

TABLE II

Maximum Likelihood Parameter Estimates and Their Standard Errors

Model I					Model II				
		D_1		D_2			D_1		D_2
Parameter	Est. value	Std. error	Est. value	Std. error	Parameter	Est. value	Std. error	Est. value	Std. error
$\bar{\theta}$	.23	.038	.41	.063	$\bar{\theta}$	.23	.037	.40	.068
α_1	.00	.015	.21	.067	α	.02	.015	.08	.036
α_2	.00	.016	.07	.062					
α_3	.05	.043	.02	.029					
α_4	.02	.032	.05	.053					
β_1	.13	.035	.25	.059	β	.13	.017	.34	.041
β_2	.12	.033	.22	.062					
β_3	.13	.034	.57	.063					
β_4	.13	.035	.29	.065					

TABLE III

Observed and Expected Response Frequencies

Response Pattern	D_1 Observed Freq. Val. sample	D_1 Observed Freq. Cross-val. sample	D_1 Expected Freq. Model I	D_1 Expected Freq. Model II	D_2 Observed Freq. Val. sample	D_2 Observed Freq. Cross-val. sample	D_2 Expected Freq. Model I	D_2 Expected Freq. Model II
0 0 0 0	31	18	30.67	30.97	41	41	41.04	41.07
1 0 0 0	0	0	.18	.77	13	12	12.91	5.95
0 1 0 0	0	1	.22	.77	6	10	5.62	5.95
0 0 1 0	2	2	1.93	.77	1	3	1.30	5.95
0 0 0 1	1	1	.94	.77	4	3	4.04	5.95
1 1 0 0	1	4	1.40	1.35	7	8	8.92	4.68
1 0 1 0	3	1	1.28	1.35	3	1	1.93	4.68
1 0 0 1	2	1	1.18	1.35	6	2	6.13	4.68
0 1 1 0	1	2	1.52	1.35	2	2	2.08	4.68
0 1 0 1	5	1	1.40	1.35	5	5	6.61	4.68
0 0 1 1	0	1	1.32	1.35	4	1	1.42	4.68
1 1 1 0	9	7	9.79	9.18	7	8	6.19	8.32
1 1 0 1	5	10	9.01	9.18	23	16	19.74	8.32
1 0 1 1	7	11	8.26	9.18	1	4	4.22	8.32
0 1 1 1	8	7	9.81	9.18	4	6	4.90	8.32
1 1 1 1	67	75	63.07	63.15	15	20	14.95	15.82

TABLE IV

Statistical Tests of Model Fit

	D_1[a]		D_2	
Assessment	Model I	Model II	Model I	Model II
Model Validation				
Chi-square	16.757	20.182	9.459	51.758
P-Value	.010	.064	.149	.000
Model Cross-Validation[b]				
Chi-Square	18.284	17.903	12.997	34.173
P-Value	.248	.268	.603	.003
Comparison of Models				
Chi-Square	4.276		52.643	
P-Value	.639		.000	

[a]The p-values related to domain D_1 are probably too small because of the large number of expected values less than 2, (i.e., 10 out of the 16 expected values).

[b]Model cross-validation was based on fit provided by the original expected frequencies to the observed frequencies obtained form the 142 students not used in parameter estimation.

TABLE V

Estimated Joint Proportions of Response Patterns and Mastery States

Response Pattern	D_1				D_2			
	Model I		Model II		Model I		Model II	
	$\hat{P}(j \cap \bar{M})$	$\hat{P}(j \cap M)$	$\hat{P}(j \cap \bar{M})$	$\hat{P}(j \cap M)$	$\hat{P}(j \cap \bar{M})$	$\hat{P}(j \cap M)$	$\hat{P}(j \cap \bar{M})$	$\hat{P}(j \cap M)$
0 0 0 0	.2158	.0002	.2179	.0002	.2837	.0053	.2808	.0084
1 0 0 0	.0000	.0013	.0041	.0014	.0748	.0161	.0259	.0160
0 1 0 0	.0000	.0015	.0041	.0014	.0208	.0188	.0259	.0160
0 0 1 0	.0122	.0014	.0041	.0014	.0052	.0040	.0259	.0160
0 0 0 1	.0053	.0013	.0041	.0014	.0156	.0128	.0259	.0160
1 1 0 0	.0000	.0099	.0001	.0094	.0055	.0573	.0024	.0306
1 0 1 0	.0000	.0090	.0001	.0094	.0014	.0123	.0024	.0306
1 0 0 1	.0000	.0083	.0001	.0094	.0041	.0391	.0024	.0306
0 1 1 0	.0000	.0107	.0001	.0094	.0004	.0142	.0024	.0306
0 1 0 1	.0000	.0099	.0001	.0094	.0011	.0454	.0024	.0306
0 0 1 1	.0003	.0091	.0001	.0094	.0003	.0097	.0024	.0306
1 1 1 0	.0000	.0690	.0000	.0647	.0001	.0435	.0002	.0583
1 1 0 1	.0000	.0635	.0000	.0647	.0003	.1387	.0002	.0583
1 0 1 1	.0000	.0582	.0000	.0647	.0001	.0297	.0002	.0583
0 1 1 1	.0000	.0691	.0000	.0647	.0000	.0345	.0002	.0583
1 1 1 1	.0000	.4441	.0000	.4447	.0000	.1053	.0000	.1114

0 0 0 0, 0 0 1 0, and 0 0 0 1. This results in an expected misclassification error of .0032. The response patterns designating non-mastery status for domain D_2 are 0 0 0 0, 1 0 0 0, 0 1 0 0, 0 0 1 0, and 0 0 0 1, which result in an expected misclassification error of .0703.

For Model II, the "number correct" scores designating non-mastery status for both domains are: 0 and 1, which result in expected misclassification errors of .0064 and .0876 respectively for D_1 and D_2. Note that for both of the domains, the resulting misclassification errors obtained under both models are similar in magnitude, with the error rate occurring under Model I being slightly smaller. This is to be expected since only Model I allows for differential mastery designation of response patterns with the same number of correct responses.

If the parameter estimates obtained under Model II for D_1 are used to estimate (via Table I) the minimally sufficient number of items necessary for a maximum acceptable proportion of misclassified examinees of .05, when the loss ratio used in classification is 1.0, then the estimated number of items is two. This value may be compared with the actual number of items required which is also two. Under the same specified conditions for D_2, the estimated number of items is six (if reversals are made for table entry between THETA and THETA as well as AF and AG), while the actual number required is also six.

REFERENCES

Besel, R. R. Using group performance to interpret individual responses to criterion-referenced tests. Professional Paper 25, SWRL Educational Research and Development, Los Alamitos, California, June 1973.

Besel, R. R. Mixed group validation and the problem of mastery-learning decisions. Paper presented at the 1975 annual meeting of the American Educational Research Association, Washington, D.C., April 1975.

Crehan, K. D. Item analysis for teacher-made mastery tests. Unpublished doctoral dissertation, State University of New York at Buffalo, 1973.

Dayton, C. M., & Macready, G. B. A probabilistic model for validation of behavioral hierarchies. Psychometrika, 1976, 41, 189-204.

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

Emerick, J. A., & Adams, E. N. An evaluation model for individualized instruction. Research Report RC 2674, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, 1969.

Gagné, R. M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Harris, C. W. Some technical characteristics of mastery tests. In C. W. Harris, M. C. Alkin & W. J. Popham (Eds.), Problems in Criterion-Referenced Measurement. Los Angeles: Center for the Study of Evaluation, U.C.L.A., 1974, 98-115.

Kriewall, T. E. Applications of information theory and acceptance sampling principles to the management of mathematics instruction. Technical Report No. 103, Wisconsin Research and Development Center for Cognitive Learning, Madison, Wisconsin, 1969.

Macready, G. B., & Merwin, J. C. Homogeneity within item forms in domain referenced testing. Educational and Psychological Measurement, 1973, 33. 351-360.

Rao, C. R. Linear statistical inference and its applications. New York: Wiley, 1965.

Swaminathan, H., Hambleton, K., & Algina, J. A Bayesian decision-theoretic procedure for use with criterion-referenced tests. Journal of Educational Measurement, 1975, 12, 87-98.

AUTHORS

MACREADY, GEORGE B. Address: Department of Measurement and Statistics, College of Education, University of Maryland, College Park, MD 20742. Title: Associate Professor. Degrees: B.A. Willimette University, M.A. University or Oregon, Ph.D. University of Minnesota. Specialization: Research Methodology, Measurement Theory.

DAYTON, C. MITCHELL. Address: Department of Measurement and Statistics, College of Education, University of Maryland, College Park, MD 20742. Title: Professor. Degrees: B.A. University of Chicago, M.A., Ph.D. University of Maryland. Specialization: Applied Statistics, Computer Applications.

[*Manuscript received September 1976; revised March 1977.*]