

AD735499

A COMPUTER PROGRAM FOR ESTIMATING
ITEM CHARACTERISTIC CURVE PARAMETERS
USING BIRNBAUM'S THREE-PARAMETER LOGISTIC MODEL

Diana M. Lees
Marilyn S. Wingersky
and Frederic M. Lord

This research was sponsored in part by the
Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research, under
Contract No. N00014-69-C-0017

Contract Authority Identification Number
NR No. 150-303

Frederic M. Lord, Principal Investigator

Educational Testing Service

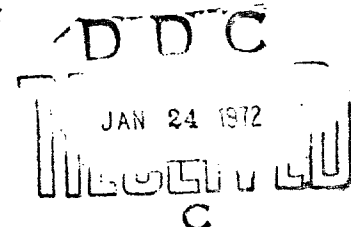
Princeton, New Jersey

January 1972

Reproduction in whole or in part is permitted for
any purpose of the United States Government.

Approved for public release; distribution
unlimited.

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151



R

REPRODUCED FROM
BEST AVAILABLE COPY

A COMPUTER PROGRAM FOR ESTIMATING
ITEM CHARACTERISTIC CURVE PARAMETERS
USING BIRNBAUM'S THREE-PARAMETER LOGISTIC MODEL

Diana M. Lees
Marilyn S. Wingersky
and Frederic M. Lord

This research was sponsored in part by the
Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research, under
Contract No. N00014-69-C-0017

Contract Authority Identification Number
NR No. 150-303

Frederic M. Lord, Principal Investigator

Educational Testing Service
Princeton, New Jersey

January 1972

Reproduction in whole or in part is permitted for
any purpose of the United States Government.

Approved for public release; distribution
unlimited.

A COMPUTER PROGRAM FOR ESTIMATING ITEM CHARACTERISTIC CURVE PARAMETERS
USING BIRNBAUM'S THREE-PARAMETER LOGISTIC MODEL

Diana M. Lees, Marilyn S. Wingersky, and Frederic M. Lord

Abstract

A computer program is described that uses maximum likelihood methods to estimate item parameters under the three-parameter logistic model for item characteristic curves. The procedure also estimates examinee ability levels.

~~Security Classification~~

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Educational Testing Service Princeton, New Jersey 08540		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE A COMPUTER PROGRAM FOR ESTIMATING ITEM CHARACTERISTIC CURVE PARAMETERS USING BIRNBAUM'S THREE-PARAMETER LOGISTIC MODEL			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report			
5. AUTHOR(S) (First name, middle initial, last name) Diana M. Lees, Marilyn S. Wingersky, and Frederic M. Lord			
6. REPORT DATE January 1972		7a. TOTAL NO. OF PAGES 40	7b. NO. OF REFS 3
8a. CONTRACT OR GRANT NO. N00014-69-C-0017		9a. ORIGINATOR'S REPORT NUMBER(S)	
b. PROJECT NO. NR No. 150-303			
c.		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.			
10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Personnel and Training Research Programs Office of Naval Research Arlington, Virginia 22217	
13. ABSTRACT A computer program is described that uses maximum likelihood methods to estimate item parameters under the three-parameter logistic model for item characteristic curves. The procedure also estimates examinee ability levels.			

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6801

Security Classification

3ND PPSO 1315

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Computer Program						
Mental Test Theory						
Computer-based Testing						
Maximum Likelihood						

A COMPUTER PROGRAM FOR ESTIMATING ITEM CHARACTERISTIC CURVE PARAMETERS
USING BIRNBAUM'S THREE-PARAMETER LOGISTIC MODEL*

Given the item responses of a group of examinees to a set of items, the computer program, LOGIST, estimates the parameters of the item characteristic curve for each item, using Birnbaum's three-parameter logistic model. The parameters are estimated by maximum likelihood, using an improved version of the method described by Lord (1968). The program was initially written by Diana Lees and modified later by Marilyn Wingersky.

The parameters of item characteristic curves are needed for effective design of tailored tests, for comparing the relative efficiency of two tests or of two scoring methods at any specified ability level, and also for computing a scoring weight for each item that will maximize the effectiveness of the test at some specified ability level.

The set of items used to estimate the item parameters must meet the assumptions underlying the Birnbaum model. These assumptions are (Lord, 1968)

- "1. The test items have only one psychological dimension in common ...
2. The test items are scored either right or wrong.
3. The probability that an examinee will answer a given item correctly is a three-parameter logistic function of the trait being measured."

(See equation 1 in Appendix A.)

*This research was sponsored in part by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-69-C-0017, Contract Authority Identification Number, NR No. 150-303, and Educational Testing Service. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Method

The parameters estimated by the program are the ability levels, θ_a , for examinee a ($a = 1, 2, \dots, N$); and three parameters for item i ($i = 1, 2, \dots, n$). The item parameters are a_i , the item discriminating power, b_i , the item difficulty, and c_i , the pseudochance-score level. The c_i may be held constant and only the a_i and b_i estimated. Although the origin and unit of measurement for the θ 's are arbitrary, leaving only $N - 2$ unique θ 's, all N θ 's are estimated. The estimated θ 's are linearly transformed to prevent uncontrolled drift.

The likelihood equations for these parameters are solved iteratively in stages using a modified scoring method (Rao, 1965). Each stage consists of two parts; the first part solves for new θ estimates holding the item parameters constant, the second part uses the new θ estimates and, holding them constant, computes new estimates of the item parameters. The initial starting values for the procedure are computed by the program. The likelihood equations and details about the procedure are given in Appendix A.

Convergence

It is important to have a sizable number of items, perhaps $n \geq 40$, and a large number of examinees, perhaps $N \geq 1000$. Unless the sample is large, an estimated item discriminating power may increase without limit, in which case, it may be necessary to remove the unstable item before the process will converge.

To improve convergence, the program automatically excludes examinees with perfect number-right scores on those items included in the estimation procedure. The maximum change of any parameter in an iteration is restricted by the program to dampen wild fluctuations.

The program does not check for oscillating or diverging parameter estimates and it is therefore advisable to periodically check the progress of the iterations. A restart feature is built into the program.

The procedure is considered converged when the changes in the parameters become less than the truncation errors, causing the likelihood function to fluctuate. The process ends when the likelihood decreases in at least two out of four successive stages. This may occur in 20 or 30 stages. This convergence criterion is a strict one--fewer stages yield quite satisfactory results.

Program Considerations

The computer program is written in FORTRAN IV with several IBM 360 assembly language subroutines. The assembly language subroutines are for bit manipulation, timer control, and non-FORTRAN input. The program requires 150K for compilation in FORTRAN G and about 126K core for execution on an IBM 360/65.

The dimensions of the program allow for 3000 examinees and 96 items. The item responses are stored in memory in binary, 32 responses per machine word.

The program isn't fast but it is feasible. Some idea of the execution time on a 360/65 can be gotten from the following information. It took approximately 150 seconds per stage to run 64 items with 2861 examinees, requiring about 20 stages to converge. Another set of data with 42 items and 2858 examinees averaged 90 seconds per stage.

A test is made at the end of each stage to see if there is enough time for another stage. If not, the current values of the parameters are punched and the program exits.

Input*

The only input necessary to start the procedure is the item responses for each examinee scored 1 if right, 0 otherwise. Omits are treated as wrong. The input data file may be in binary or BCD. Input options include selecting items from those read and omitting examinees from the input data file.

To restart the procedure where the previous run ended, the punched output from that run is necessary. This output may be modified by removing, adding, or changing items or changing thetas.

Output*

The printout contains sufficient information for following the progress of the procedure. Some of the information printed for each stage is: a) all θ 's whose absolute value is greater than 3; b) the value of the likelihood function after the theta estimation, and again after the item parameter estimation; c) the maximum value of the derivative of the likelihood function for each type of parameter. During each item parameter iteration, the previous value, the current value, the change, and the value of the derivative of the likelihood function are printed for each parameter. At the end of each run the final item parameters and the final thetas are both printed and punched.

Availability

A copy of the program may be obtained upon written request from the authors, Educational Testing Service, Princeton, New Jersey 08540. The user must provide a tape on which the program will be loaded in 80-character card images. The user must specify whether the tape should be blocked, in EBCDIC or BCD, 7-track or 9-track, and the tape density and parity.

*A detailed description is given in Appendix B.

References

- Hastings, C., Jr. Approximations for digital computers. Princeton, N.J.:
Princeton University Press, 1955. P. 192.
- Lord, F. M. An analysis of the Verbal Scholastic Aptitude Test using
Birnbaum's three-parameter logistic model. Educational and
Psychological Measurement, 1968, 28, 989-1020.
- Rao, C. R. Linear statistical inference and its applications. New York:
Wiley, 1965. P. 305.

Appendix A. Formulas and Computational Details

The three-parameter logistic model defines the probability that examinee a will get item i right as

$$(1) \quad P_{ia} = \text{Prob}(x_{ia} = 1 | a_i, b_i, c_i; \theta_a) = c_i + (1 - c_i) \psi(f_{ia})$$

where ψ is the logistic function

$$(2) \quad \psi(f_{ia}) = \frac{1}{1 + \exp(-f_{ia})}$$

$$\text{and } (3) \quad f_{ia} = Da_i(\theta_a - b_i) .$$

D is a constant approximately equal to 1.7, θ_a is the ability level for examinee a , a_i is the item discriminating power, b_i is the item difficulty, and c_i is the pseudo-chance score level for item i . The response of examinee a to item i is x_{ia} coded 1 if the response was correct, 0 otherwise.

Given the $N \times n$ matrix of item responses x_{ia} , the program computes estimates of a_i , b_i , c_i , ($i = 1, 2, \dots, n$) and θ_a ($a = 1, 2, \dots, N$) that maximize the likelihood function

$$(4) \quad L = \prod_{i=1}^n \prod_{a=1}^N P_{ia}^{x_{ia}} Q_{ia}^{1-x_{ia}} ,$$

by maximizing the log of the likelihood function,

$$\log L = \sum_{i=1}^n \sum_{a=1}^N (x_{ia} \log P_{ia} + (1 - x_{ia}) \log Q_{ia}) ,$$

where N is the number of people, n is the number of items, and Q_{ia} is $1 - P_{ia}$. There are N ability parameters and $3n$ item parameters to be estimated.

Method of Estimation

A variation of the scoring method described in Rao (1965) is used to compute the parameter estimates. With trial values for the parameters, the scoring method computes additive corrections to the values by solving a set of simultaneous equations. This process is repeated, correcting the corrected values until the size of the corrections becomes negligible. The estimated values will be denoted by writing the parameters with "hats" over them. The matrix of coefficients for the linear equations is the information matrix. A direct application of the scoring procedure to this problem would require solving an $N + 3n$ system of linear equations.

To simplify the problem, the linear equations are solved in two sections. In the first section the item parameters are held constant and the equations solved for the corrections to the $\hat{\theta}$'s until the corrections become negligible. In the second section the $\hat{\theta}$'s are held constant and the item parameters corrected until the corrections become negligible. The two sections combined are called a stage. The stages are repeated to adjust for dependencies between the abilities and item parameters until the likelihood function no longer increases.

There are three types of convergence, the convergence of the corrections to the abilities within a stage, the convergence of the corrections to the item parameters within a stage, and the convergence over stages of the likelihood function to a maximum.

Explicitly a stage consists of four parts: computing new $\hat{\theta}$'s, scaling the $\hat{\theta}$'s, \hat{a} 's, and \hat{b} 's, computing new item parameter estimates, and testing for convergence of the likelihood function. The following sections explain the four parts in detail. To simplify writing, the estimated parameters will be written without "hats" for the rest of the paper.

Estimating the θ 's

Since the θ 's are independent of each other they can be estimated one at a time. The increment for θ_a , $\Delta\theta_a$, is found by solving the following equation.

$$(5) \quad \Delta\theta_a \cdot I(\theta_a) = S(\theta_a)$$

where

$$(6) \quad S(\theta_a) = \frac{\partial \log L}{\partial \theta_a}$$

$$= D \sum_{i=1}^n a_i \frac{(p_{ia} - c_i)(x_{ia} - p_{ia})}{(1 - c_i)p_{ia}}$$

and

$$(7) \quad I(\theta_a) = \varepsilon(S(\theta_a)^2)$$

$$= D^2 \sum_{i=1}^n a_i^2 \frac{(p_{ia} - c_i)^2 Q_{ia}}{(1 - c_i)^2 p_{ia}}$$

Let θ^r denote the new value of θ_a and θ_a^{r-1} the previous value of θ_a , then $\theta_a^r = \theta_a^{r-1} + \Delta\theta_a$. The iterations for an examinee continue until either the maximum number of iterations is reached or $|\Delta\theta_a|$ becomes less than some tolerance specified by the user.

Restrictions on θ Estimation

The following restrictions are placed on the magnitude and change in θ_a .

- (a) If θ_a^r has a different sign from θ_a^{r-1} and $|\Delta\theta_a| > 1$, then set $\theta_a^r = \frac{1}{2} \theta_a^{r-1} = 1 * \text{sign}(\theta_a^{r-1})$.
- (b) The amount that θ_a may change is restricted such that $|\Delta\theta| \leq |.1 * \theta_a^{r-1}| + 2$. If $\Delta\theta_a$ is greater than this amount, then set $\theta_a^r = \theta_a^{r-1} + (|.1 * \theta_a^{r-1}| + 2) * \text{sign}(\Delta\theta_a)$.
- (c) The range of θ_a^r is restricted so that $TLO \leq \theta_a^r \leq THI$, where $TLO = -10^{10}$ and $THI = 10^{10}$ unless different extrema have been read. If $\theta_a^r > THI$, θ_a^r is set to THI . If $\theta_a^r < TLO$, θ_a^r is set to TLO .

A message is printed if any adjustments are made in the θ 's.

Scaling

Since the origin and unit of measurement of the θ 's are arbitrary, there are only $N - 2$ unique θ 's. However, all N θ 's are estimated and then linearly transformed to prevent uncontrollable drift. The a 's and b 's are adjusted accordingly.

The standard method of normalizing using the arithmetic mean and standard deviation isn't used since a few of the θ 's may be increasing or decreasing without limit. Instead the θ_a are scaled so that they are centered around the midpoint of those θ_a lying in the interval $[-2/3, 2/3]$; and so that for $2/3$ of the θ_a , $|\theta_a| \leq 1$. This is accomplished approximately using the following algorithm.

1) Let $P23N$ be the proportion of the θ_a which satisfy the condition that $-2/3 \leq \theta_a < 0$ and let $P23P$ be the proportion of the θ_a which satisfy the condition that $0 \leq \theta_a \leq 2/3$. A midpoint, $X23$, is determined (by interpolation) such that $P23P = P23N = 1/2$ and all the θ_a are shifted by the amount $X23$, $\theta_a^* = \theta_a - X23$.

2) let P23 be the proportion of the Θ_n^* which satisfies the condition that $-2/3 \leq \Theta_n^* \leq 2/3$ and let P115 be the proportion which satisfies the condition that $1.15 \leq \Theta_n^* \leq 1.15$. Then a scale factor, X , is determined (by interpolation) such that $2/3$ of the Θ_n^* satisfy (approximately) the condition that $-1 \leq \Theta_n \leq 1$. All the Θ_n are then scaled by X , so that $\Theta_n^{**} = \Theta_n^*/X$.

The item parameters are then scaled appropriately, i.e., $a_i = a_i X$
 $b_i = \frac{b_i - X23}{X}$. As the process reaches convergence the scaling stabilizes so that $P23P \approx 1/2$, $P23N \approx 1/2$, $P23 \approx 1/2$, $P115 \approx 3/4$, $X23 \approx 0$, and $X \approx 1$.

Item Parameter Estimation

Using the new scaled Θ 's and holding them constant, new item parameters are estimated. Since the items are independent, each item can be estimated separately. The new item parameter estimates are found by solving the following 3×3 system of linear equations. This is iterated until the maximum number of iterations has been reached, a_i has become greater than 4 or $|\Delta a_i| < .01$.

The equations to be solved are:

$$\begin{aligned} \Delta a_i I(a_i, a_i) + \Delta b_i I(a_i, b_i) + \Delta c_i I(a_i, c_i) &= S(a_i) \\ (8) \quad \Delta a_i I(a_i, b_i) + \Delta b_i I(b_i, b_i) + \Delta c_i I(b_i, c_i) &= S(b_i) \\ \Delta a_i I(a_i, c_i) + \Delta b_i I(b_i, c_i) + \Delta c_i I(c_i, c_i) &= S(c_i) \end{aligned}$$

where

$$\begin{aligned}
 (9) \quad S(a_i) &= \frac{\partial \log L}{\partial a_i} \\
 &= \frac{D}{1 - c_i} \sum_{a=1}^N (\theta_a - b_i) \frac{(P_{ia} - c_i)}{P_{ia}} (x_{ia} - P_{ia})
 \end{aligned}$$

$$\begin{aligned}
 (10) \quad S(b_i) &= \frac{\partial \log L}{\partial b_i} \\
 &= - \frac{D a_i}{1 - c_i} \sum_{a=1}^N \frac{(P_{ia} - c_i)}{P_{ia}} (x_{ia} - P_{ia})
 \end{aligned}$$

$$\begin{aligned}
 (11) \quad S(c_i) &= \frac{\partial \log L}{\partial c_i} \\
 &= \frac{1}{1 - c_i} \sum_{a=1}^N \frac{x_{ia} - P_{ia}}{P_{ia}}
 \end{aligned}$$

$$\begin{aligned}
 (12) \quad I(a_i, a_i) &= E(S(a_i)^2) \\
 &= \frac{D^2}{(1 - c_i)^2} \sum_{a=1}^N (\theta_a - b_i)^2 (P_{ia} - c_i)^2 \frac{Q_{ia}}{P_{ia}}
 \end{aligned}$$

$$\begin{aligned}
 (13) \quad I(a_i, b_i) &= E(S(a_i)S(b_i)) \\
 &= - \frac{D^2 a_i}{(1 - c_i)^2} \sum_{a=1}^N (\theta_a - b_i) (P_{ia} - c_i)^2 \frac{Q_{ia}}{P_{ia}}
 \end{aligned}$$

$$\begin{aligned}
 (14) \quad I(a_i, c_i) &= E(S(a_i)S(c_i)) \\
 &= \frac{D}{(1 - c_i)^2} \sum_{a=1}^N (\theta_a - b_i) (P_{ia} - c_i) \frac{Q_{ia}}{P_{ia}}
 \end{aligned}$$

$$(15) \quad I(b_i, b_i) = E(S(b_i)^2)$$

$$= \frac{D_{a_i}^2}{(1 - c_i)^2} \sum_{a=1}^N (P_{ia} - c_i)^2 \frac{Q_{ia}}{P_{ia}}$$

$$(16) \quad I(b_i, c_i) = E(S(b_i)S(c_i))$$

$$= - \frac{D_{a_i}}{(1 - c_i)^2} \sum_{a=1}^N (P_{ia} - c_i) \frac{Q_{ia}}{P_{ia}}$$

$$(17) \quad I(c_i, c_i) = E(S(c_i)^2)$$

$$= \frac{1}{(1 - c_i)^2} \sum_{a=1}^N \frac{Q_{ia}}{P_{ia}}$$

Restrictions on Item Parameters

The range and maximum change for each item parameter is restricted to prevent large changes. Unless otherwise noted, whenever an item parameter is restricted, the other parameters for that item are restricted by interpolation to corresponding positions. An adjusted item parameter will be allowed to vary in the next iteration unless otherwise specified. A message is printed if adjustments are made. Let r denote the stage, j the iteration within the stage.

Maximum step size:

If the correction to a parameter exceeds the maximum given below, it is set to the limit and the other parameters adjusted.

- 1) $\Delta a_i \leq |.1 \cdot a_i^{r,j-1}| + .2$
- 2) $\Delta b_i \leq |.2 \cdot b_i^{r,j-1}| + .4$
- 3) $\Delta c_i \leq \text{CLAMDA}$. CLAMDA is read in. A good value seems to be .06.

Restrictions on range for item parameters:

1) For a_i , the maximum value $\text{AHI} = 10^{10}$ and the minimum value $\text{ALO} = 0$ are set by the program unless other values are read. If a_i^r exceeds the maximum and the previous value was at the maximum, the simultaneous equations for item i are resolved for b_i and c_i omitting the a_i row and column. Otherwise, if a_i goes out of range, it is set to the bound exceeded.

2) For b_i , the maximum value BHI and the minimum value BLO are set by the program with $\text{BLO} = -10^{10}$ and $\text{BHI} = +10^{10}$, or other values for the extrema may be read. If $b_i^{r,j} > \text{BHI}$, $b_i^{r,j}$ is set to BHI . If $b_i^{r,j} < \text{BLO}$, $b_i^{r,j} = \text{BLO}$.

3) For c_i , the maximum value, CHI , may be read or it is set by the program to 1 divided by the number of choices. Two minimum values for c_i are read. One, CIOM , is used when $b_i \leq 0$, the other CIOP is used when $b_i > 0$. If $c_i^{r,j}$ becomes less than the appropriate minimum, indicating that it is poorly determined, it is set to CHI , which is presumably a better estimate than the appropriate CIO . The other parameters for this item are not adjusted, and c_i is held constant for the rest of this stage and in the following stages.* If $c_i^{r,j} \geq \text{CHI}$, it is set to CHI .

*This procedure is currently being improved.

Item Parameter Options

The following options are available for item parameter iterations (Usually the c's are held constant for the first few stages and then allowed to vary, but this isn't done automatically unless the procedure has converged.):

- a) All item parameters may vary.
- b) All c's are held constant until the procedure converges, then the c's are allowed to vary.
- c) The a's and c's are held constant until the procedure converges, then the a's varied, then the c's varied.
- d) The c's are held constant throughout.
- e) Items will be removed from the iteration procedure at the end of a stage under the following circumstances:
 1. If the matrix of coefficients for this item is singular, the item will be removed. If there is more than one such item, the last one encountered is removed. The others will be picked up in succeeding stages if the singularity still exists.
 2. At the request of the user items may be removed if a_i becomes greater than a specified value.

Initial Estimates

The following initial trial values are computed by the program. The initial estimates for the θ_a are the normalized number-right scores, i.e.,

$$N_a^+ = \sum_{i=1}^n x_{ia}$$

and

$$\theta_a = \frac{N_a^+ - \bar{x}}{s_x}$$

where \bar{x} and s_x are, respectively, the mean and standard deviation of the number right scores. The N_a^+ , \bar{x} , and s_x are printed.

The initial estimates for the item parameters are:

- 1) All of the a 's are set to 1.
- 2) The b_i are computed from the conventional item difficulties π_i (proportion of correct answers) using the following formulas:

$$b_i = (\sqrt{1 + a_i^2}) \gamma_i = \sqrt{2} \gamma_i$$
 , since all $a_i = 1$ initially. The γ_i are computed from the following formula using the approximation given in Hastings (1955), p. 191,

$$\pi_i = \frac{1}{2\pi} \int_{\gamma_i}^{\infty} e^{-t^2/2} dt$$

where

$$\pi_i = \frac{1}{N} \sum_{a=1}^N x_{ia}$$

- 3) If good estimates for the c_i are available, they can be used. Otherwise all c 's are set to $1/K$, where K is the number of choices

-All-

per item. The c's can be set to 0, making equation 1 the two parameter logistic model.

Since the initial estimates of the abilities are better than the initial estimates for the item parameters, the procedure should be started with the second part of a stage (estimating the item parameters).

Termination Procedures

The program will terminate and punch all parameters for the following reasons:

1. The process has converged when the likelihood function has decreased twice out of four successive stages. At this point all of the derivatives of the likelihood function should be approximately 0. Two more stages are computed with detailed printout.
2. The maximum number of stages have been completed.
3. There isn't enough time for another stage.
4. The maximum number of items have been removed, if any were allowed to be removed. Items may be removed if the matrix of coefficients was singular or if the a_i exceeded a maximum of 10 or a maximum read in.

The procedure may be restarted using the punched output.

Appendix B

Input

Options

The user has some control over the execution of the program. To keep track of the progress of the procedure, the user can specify that only a given number of stages be executed and/or that the program execute for only a specified number of seconds. The maximum number of iterations for an item in a stage must be specified. The iterations for an ability parameter are terminated by either exceeding the maximum number of iterations read in or when $\Delta\theta_a$ is less than a user specified tolerance. The first stage in a run may begin with either estimating the abilities or with estimating the item parameters. For any item, c_i , or a_i and c_i , may be held constant while the remaining parameters vary. The bounds on the ranges of the parameters and the number of items that may be removed if a_i exceeds a specified maximum may be supplied by the user. The initial trial values may be computed by the program or supplied by the user.

The item response input must contain one record or card for each examinee with the responses coded 1 if right, 0 otherwise. The item responses may be on tape, disk, or cards in BCD or, if on tape or disk, in binary. Selection of examinees from the data file is possible to the extent of skipping the first m examinees and selecting every i^{th} one thereafter, where m and i are specified by the user. Specific examinees may also be omitted from the input. The items to be selected from the input record are specified.

The procedure can be restarted using the punched output from a previous run. This output may be modified after being read. The modifications

possible are changing θ values, removing examinees and their θ 's, adding examinees and their θ 's, changing item parameters and adding or removing items.

The matrix of item responses may be printed. However it will be in hexadecimal. The matrix of coefficients to the linear equations may be printed for as many stages as desired.

Minimum Setup

The following setup will use the program with a minimum of options. In this setup the program computes the trial values, removes no examinees from the item response data read on unit 10, sets the initial c 's to 1 divided by the number of choices per item and does not use a_1 as a criterion for removing items.

Card

- | | |
|---|--|
| 1 | Punch card 1 as described in setup. |
| 2 | Punch cols. 1-5, 6-10, 16-20, 35, 75, and 80 as described in setup. Leave the rest of the card blank. |
| 3 | Punch card 3 as described in setup. A good value for CLAMDA is .06, for CLOM is .05, and for CLOP is 0. |
| 4 | Punch as described in the setup. |
| 5 | Cols. 1-15 should be blank. Set col. 20 to the maximum number of items that can be removed. The rest of the card should be blank. This is card 6 in setup. |
| 6 | If col. 80 of card 2 is 1, punch card 11 as described in setup. |

Card Input

CARD 11	COLS	FORMAT (10A6, D4, 14, 17)
	1-60	TITLE
	61-62	(not read)
	63-65	n = number of items $n \leq 90$
	66-72	N = number of examinees $N \leq 5000$
CARD 21	COLS	FORMAT (1615)
	1-5	MAXST = Maximum number of stages
	6-10	MAXIT = Maximum number of iterations per item per stage
	15	blank
	19-20	NCH = 0 If c = 0 = 1-9 = Number of choices/item ≥ 10 LIM (col. 80) must be 1. See cards 10 for c input. NCH must be ≥ 10 if the output from a previous run is being read.
	25	NIT = 0 Begin the run with the estimation of the item parameters. > 0 Begin the run with the estimation of abilities.
	26-30	NGT = Number of examinees to be skipped at the beginning of the tape
	35	INPUT = 0 Responses are in BCD on unit 10. = 2 Responses are on cards. = 3 Responses are in binary on unit 10. See description of item response input.

CARD #1

0018

FORMAT (16I5)

36-40

NSKIP =

Number of examinees to be skipped when reading tape, after eliminating the first "NOT" examinees.

If $NSKIP > 0$, then take every "NSKIP"-th examinee, e.g., if $NSKIP = 4$ then the 4th, 8th, 12th, etc. examinees will be used. If $NSKIP = 1$ and $NOT = 0$, the first N examinees will be used.

41-45

NC = 0

Vary all item parameters unless $NCH = 0$ then vary only a_i and b_i .

= 1 Hold c_i constant until the procedure converges, then vary c_i .

= 2 Hold a_i and c_i constant until the procedure converges, then hold c_i constant until the procedure again converges, then vary all parameters, unless $c_i = 0$.

= 3 Same as $NC = 1$, but don't vary c_i .

46-50

IIAST =

Number of last stage for which the matrix of coefficients for the item parameters is printed.

Detailed Input

CARD 2:	COLS	FORMAT (1615)
51-55		IT = 0 Compute initial estimates of θ_a 's. = 1 Read in initial estimates of θ_a 's. See CARD 5. = 2 Read in the output from a previous run.
56-60		IA = 0 Set $a_i = 1.0$ for all items. $\neq 0$ Read in initial estimates of a_i 's. See CARDS 8.
61-65		IB = 0 Compute initial estimates of b_i 's. $\neq 0$ Read in initial estimates of b_i 's. See CARDS 9.
66-70		IY = 0 Do not print matrix of item responses. $\neq 0$ Print matrix of item responses. NOTE: This print is in hexadecimal.
71-75		ITONE = 0 Only 1 iteration per θ_a > 0 Number of iterations per θ_a
76-80		LIM = 0 Use the following as limits on the parameters $10^{-10} \leq a_i \leq 10^{10}$ $10^{-10} \leq b_i \leq 10^{10}$ $10^{-10} \leq \theta_a \leq 10^{10}$ $c_i \leq K$ where $K = NCH$ (col. 20). $\neq 0$ Read the limits. See CARD 11.

Detailed Input

CARD 3	COLS	FORMAT (E15.8, F6.0, 3F6.2)
	1-15	TOLL- iterate on θ_a until $ \Delta\theta_a < \text{TOLL}$
	16-21	TIMAX = Maximum time for run in seconds minus 20 seconds
	22-27	CLAMDA - maximum allowable change in c_i
	28-33	CLOM - minimum value for c_i if $b_i \leq 0$
	34-39	CLOP - minimum value for c_i if $b_i > 0$
CARD(S) 4: (as many as 4 cards)	COLS	FORMAT (26I3, 2X)
	1-3	ITEMNO(1) No. of 1st item
	4-6	ITEMNO(2) No. of 2nd item
	:	
	:	
	:	ITEMNO(LITTLN) No. of n^{th} item
		This card controls the selection of items from the input record, and assumes the items on the input record are numbered consecutively from 1 to the number of items in the record.
CARD(S) 5		Omit if IT (col.55 of card 2) does not equal 1.
	card a	set col. 2 to 0
	card b	col. 64-65 number of items 66-72 number of θ 's being read
	card c	blank
	card d	FORMAT for reading θ 's
	card e	θ 's punched in format given in card 'd.

-B7-

CARD(S)	COLS	FORMAT (16I5)
	1-5	NKILL = number of examinees to be removed from the item response data file. (NKILL \leq 99 .)
	6-10	KILL(1) - sequential position on the input data
	11-15	KILL(2) file of the examinees to be removed. If NKILL is 0, then KILL(1) must be left blank. The total number of KILL(NKILL) examinees to be read is N on card 1 plus NKILL.
		IFILL - blank field following the number of the last examinee to be removed.
		MAXOUT - maximum number of items which may be removed
		ICKA = 0 Do not check a_i as a criterion for removing items > 0 Remove item i if $a_i > ALIM$
		LA = 0 ALIM = 10 \neq 0 ALIM = LA

An example of this card will make it clear. Suppose from a data file of 100 examinees, examinees 47 and 99 are to be removed. Then N on card 1 is 98. NKILL is 2, KILL(1) is 47 and KILL(2) is 99. MAXOUT would be punched in cols. 21-25.

CARD(S) 7 Omit if INPUT (col. 35, card 2) is 0 or 3
Item responses are on cards in BCD.
Card 1: Format for item response cards in A4 format
Card 2: Item responses for each examinee punched in the format specified in card 1. No item selection is done. No examinees can be removed.

Detailed Input

CARD(S) 8: Initial values for a_i (Omit if $IT = 2$ or $IA = 0$
on card 2)

card a cols. 1-80 Format for reading in a's. The a's must be
read as floating-point numbers.

card(s) b $a(i)$, ($i = 1, LITTLN$) - Punched in the format
specified in card a.

CARD(S) 9: Initial values for b_i (Omit if $IT = 2$ or $IB = 0$
on card 2)

card a cols. 1-80 Format for reading in b's. The b's must be
read as floating-point numbers.

card(s) b $b(i)$, ($i = 1, \dots, LITTLN$) - Punched in the format
specified in card a.

CARD(S) 10: Initial values for c_i (Omit if $IT = 2$ or $NCH \leq 10$
on card 2)

$c(i)$, ($i = 1, \dots, LITTLN$) - Punched in 8F10.7 format.

Note: If the c_i are read in, then the limits on the parameters
must be read in, i.e., $LIM \neq 0$ on CARD 2.

CARD(S) 11:	COLS	FORMAT (8F10.0) (Omit if $LIM = 0$ on card 2)
	1-10	$AL\emptyset$ = lower limit on a_i
	11-20	AHI = upper limit on a_i
	21-30	$BL\emptyset$ = lower limit on b_i
	31-40	BHI = upper limit on b_i
	41-50	blank
	51-60	CHI = upper limit on c_i
	61-70	$TL\emptyset$ = lower limit on θ_a
	71-80	THI = upper limit on θ_a

Restart Procedure

The program may be restarted using the punched output from a previous run. It is possible to add or remove examinees, change θ 's, change item parameter estimates, and add or remove items from the output after it has been read in.

If no changes are to be made in the output, the following modifications must be made to the setup to restart after the first run on a set of data.

- a. If any of the examinees were removed from the previous run because they had perfect scores, decrease N on the first card by the number of examinees removed. Modify card 6 to add these examinees to the examinees being removed from the input file.
- b. On card 2 make the following additions. Set col. 17-20 to 2000, col. 25 to 1, col. 55 to 2, and col. 80 to 1.
- c. Substitute the following for card set 5:
 1. Modification card, with col. 2 set to 1.
 2. The punched output from the previous run.
- d. Card 11 must be included. The limits punched are the limits set by the program (see description of card 2, col. 80 in the setup).

Subsequent restarts only require that the punched output from the previous run be replaced with the new punched output.

The following pages describe how to add or remove thetas and the corresponding examinees, add or remove items, change item parameters or change thetas when reading in the output of a previous run. If no changes are necessary, skip to the section titled "Item Response Input" on page B18.

The cards for the changes follow the punched output from a previous run. The cards must be in the order of their description here. As an added reminder the cards will be numbered starting with set c3 in the order that they should be placed. If a specific change isn't made, the cards for that change are omitted.

The modification card mentioned in c1 on page B9 contains an indicator for each change that will be set to 1 if the change is required, blank if no change is required. A brief description of the card follows.

COLS	FORMAT (40I2)
2	INP = 0 0's are punched in E format = 1 0's are punched in hexadecimal in A4 format
4	JADTH = 1 More 0's are to be estimated than those read
6	JSBTH = 1 Remove 0's from the 0's read
8	JCHTH = 1 Change some 0's read
10	JADIT = 1 Add more items than read
12	JSBIT = 1 Remove some of the items from the input
14	JCHIT = 1 Change the a's, b's and c's for some items
16	JCHA = 1 Change the a's for some of the items
18	JCHB = 1 Change the b's for some of the items
20	JCHC = 1 Change the c's for some of the items

The following changes are necessary to add or remove examinees and their corresponding thetas or to change thetas for some examinees. In order to describe these changes a numerical example is essential. Suppose that in the previous run examinees 21, 65, and 250 have been removed from the item response data set. These numbers are the positions of the examinees on the input data set of item responses. Card 5 reads:

cols.

5	3
9-10	21
14-15	65
18-20	250
21-25	blank

Rest of card as described in setup.

The θ 's computed and printed will have the following relationship to the examinees on the data set of item responses.

Examinee	Position in θ 's
1	1
22	21
64	63
66	64
249	247
251	248

A. To change θ 's

1. Set JCHTR (col 8 of the modification card) to 1.
2. Add card set c3 punched as follows.

Card a	COLS	
	1-5	NOUT - Number of θ 's to be changed
Card b		4(15, E15.8)
	1-5	index of first θ to be changed
	6-20	new value for that θ
	21-25	index of 2nd θ to be changed
	26-40	new value for that θ

The index of the θ 's corresponds to the indices listed under the position in θ column above and corresponds also to the index in the listing of final θ 's printed in the corresponding output.

B. Add thetas for examinees that have previously been removed.

1. N on card 1 is the previous N plus the number of thetas being added.
2. Set JADTH (col. 4 of modification card) to 1.
3. Add card set c4 punched as follows.

COLS

Card a	1-5	N O UT - Number of θ 's to be added
Card b		4(I5, E15.8)
	1-5	Position that the first added θ will have in the new sequence
	6-20	Initial value for the first added θ
	etc.	etc. (θ 's must be punched in ascending order by position)

Using our example given above, suppose we want to add examinees 21 and 250. The positions that these examinees will have in the new list of θ 's is 21 and 249 respectively. Thus N~~O~~UT equals 2, card c4 contains 21 in cols. 4-5, cols. 6-20 contain the initial value of theta for examinee 21, and 23-25 contains 249 and cols. 26-40 contain the initial value of theta for examinee 250.

4. Card 6, set NKILL equal to NKILL minus the number of added θ 's, and remove the examinee numbers from the KILL vector. In the example, NKILL will now be 1, and KILL(1) = 65 .

C. Remove θ 's and corresponding examinees.

1. N on card 1 becomes the previous value of N minus the number of examinees to be removed.
2. Set JSBTH (col. 6 of modification card) to 1.
3. Add card set c5 punched as follows.

COLS	(16I5)
1-5	NOUT - Number of θ 's to be removed
6-10	Sequence number of first θ to be removed (numbers should be in ascending numerical order)
11-15	Sequence number of second θ to be removed
etc.	etc.

The sequence number is the position of the theta in the input vector of thetas. Referring back to the example given in the beginning of this section suppose that we want to remove the 15th and the 80th theta. Card set c5 would then contain

col.	(16I5)
5	2
9-10	15
14-15	80

4. Card 6. Since we are removing two more thetas we must also remove their associated examinees. The examinee number corresponding to the 15th θ is 15, that corresponding to the 80th θ is 82. Card 5 should read, referring back to the initial example.

cols.	
5	5
9-10	15
14-15	21
19-20	65
24-25	82
28-30	250
31-35	blank
Rest of card as described	

The following cards describe the additions to the setup to add or remove items or change the parameters for an item. The number of the item is its item number read in input card 4.

A. Add items.

1. Set JADIT (col. 10 of modification card) to 1.
2. Add card set c6 punched as follows.

COLS

card a	1-5	NOIT - number of items to be added
card b		(15, 3015.8, 15)
	1-5	Number of test item to be added
	6-20	Initial a value for that item
	21-35	Initial b value for that item
	36-50	Initial c value for that item
	51-55	ITEST; number of parameters to vary*

3. The next run that uses the output of this run will require that n on card 1 be the n of this run plus the number of items added and the item numbers of the added items must be added to card(s) 4.

B. Subtract items.

1. Set JSBIT (col. 12 of modification card) to 1.
2. Add card set c7 punched as follows.

COLS

(1615)

1-5	NOIT - Number of items to be removed
6-10	Number of 1st item to be removed
11-15	Number of 2nd item to be removed
etc.	etc.

*The number of parameters allowed to vary is 1 if only b_i is allowed to vary, 2 if a_i and b_i are allowed to vary, 3 if a_i , b_i , and c_i are allowed to vary.

4. The next run that uses the output of this run will require the following changes: (a) n on card 1 will be the n of this run minus the number of items removed and (b) the item numbers for the items removed must be removed from the item numbers on card(a) 4, (usually requires repunching).

C. Change all parameters for some items.

1. Set JCHIT (col. 14 of modification card) to 1.
2. Add card set c8, punch as described below

COLS

card a	1-5	NGIT - Number of items changed
card(s) b		(15, 3E15.8, 15)
	1-5	Number of item for which the item parameters are to be changed
	6-20	New a
	21-35	New b
	36-50	New c
	51-55	ITEST(1) ; number of parameters allowed to vary*

D. Change only some a's.

1. Set JCHA (col. 16 of modification card) to 1.
2. Add card set c9, punched as described below.

COLS

card a	1-5	NGUT - Number of items for which the initial values are to be changed
card(s) b		(4(15, E15.8))
	1-5	Number of item for which a is to be changed
	6-20	New value for a

*The number of parameters allowed to vary is 1 if only b_1 is allowed to vary, 2 if a_1 and b_1 is allowed to vary, 3 if a_1 , b_1 , and c_1 are allowed to vary.

COLS

E1-25	Number of next item for which a is to be changed
26-40	New value for that a
etc.	etc.

E. Change only b's.

1. Set JCIB (col. 18 of modification card) to 1.
2. Add card set c10, punched as described below.

COLS

card a	1-5	NOT - Number of b's to be changed
card(s) b		(4(I5, E15.8))
	1-5	Number of 1st item for which b is to be changed
	6-20	New value of b for that item
	21-25	Number of 2nd item for which b is to be changed
	26-40	New value of b for that item
	etc.	etc.

F. Change only the c's.

1. Set JCHC (col. 20 of modification card) to 1.
2. Add card set c11, punched as described below.

COLS

card a	1-5	NOT - Number of c's to be changed
card(s) b		(4(I5, E15.8))
	1-5	Number of 1st item for which c is to be changed
	6-20	New value of c for that item

COLS

21-25 Number of 2nd item for which initial c
 is to be changed

26-40 New value of c for that item

- G. To change the number of parameters that are allowed to vary for an item, either change the punched output (the number of parameters that are allowed to vary are punched 16 per card in 16I5 format) or follow the procedure for changing all of the parameters for an item, see C on page B15, punching the current values for a , b , and c but changing the number of parameters allowed to vary.

Item Response Input

The item responses for a test can be on disk or tape in BCD or binary with one record per examinee. The examinees are numbered in the order that they are read. Specific examinees can be omitted. Examinees skipped at the beginning of the data set or when selecting every k^{th} examinee are not included in the numbering. The program reads the responses from unit 10. Under OS 360, the responses can be on cards and read on unit 10. However for debugging purposes the option to read card input with the setup is included.

Each record contains the item responses coded 1 if right, 0 otherwise, with item 1 the first digit in the record and the last item is the m^{th} digit in the record, for m test items. The items to be used are specified on card 4 of the setup. For example if there are 10 responses per examinee and only the 2nd, 5th, 7th, and 8th are to be used, then card 4 would contain 2 in col. 3, 5 in col. 6, 7 in col. 9, and 8 in col. 12.

The input options in card 2 are

1. INPUT = 0. The item responses are on unit 10 in BCD.
2. INPUT = 3. The item responses are on unit 10 in binary.
3. INPUT = 2. The item responses are punched packed ready for use one card per examinee. An A4 format for reading the responses is read in at execution time. No examinees may be skipped or items selected.

Printed Output

1st Page

1. Number right score - mean and variance followed by the number right scores for each examinee printed 20 per line, after the first number in the line which is the sequence number for the first examinee in that row.
2. If input is output from a previous run, the title, no. of examinees, no. of items, time and date of the previous run is printed.
3. If any examinees are skipped, their sequence numbers are printed.
- 4.. If the input is binary, the total input record is printed in hexadecimal for the first 10 examinees, followed by the packed responses for the items selected printed in hexadecimal. The words in which the selected responses are stored are not zeroed beforehand. Consequently if the number of item responses does not fill a whole word, the remainder of the word will be garbage.

2nd Page

- 1st line: title, number of items, number of examinees, time and date.
- 2nd line: maximum number of stages, maximum number of iterations for items, and NCH, the number of choices. If the c's are read in, the number of choices will print out asterisks.
- 3rd line: the number of items that may be removed if their a's become greater than the limit for a that has been read.
- 4th line: number of people removed because they had perfect scores, followed by their sequence numbers.
- 5th line: if the initial item parameters are computed, a list of the percentage of people who got the item right is printed.
- 6th line: initial estimates for a , b , and c , and the number of item

parameters that are allowed to vary. 2 means a and b are allowed to vary, 3 that a, b, and c vary, 1 that only b varies.

7th line: initial estimates for θ 's, printed 12 per line, the first number in a line is the sequence number of the first theta in the line.
Note: The sequence numbers are from 1 to the number of examinees minus the number of the examinees removed.
8th line: initial likelihood function is printed.

Page 2

1. "Beginning of stage" followed by stage number.

2. Thetas with absolute value greater than 3. If there are none, only the message will be printed.

3. Normalized a_i and b_i , $i = 1, 2, \dots, n$.

4. Information about the normalization of θ 's.

$P(2/3)$ is the proportion of θ 's minus $X(2/3)$ with absolute values less than $2/3$.

$P(115)$ is the proportion of θ 's minus $X(2/3)$ with absolute value less than 1.15.

X is the scaling factor - refer to the description of the program for more information.

$P(-2/3)$ is the proportion of θ 's such that $-2/3 \leq \theta < 0$.

$P(2/3)$ is the proportion of θ 's such that $0 \leq \theta \leq 2/3$.

$X(2/3)$ is the mean of θ 's for θ_i , such that $-2/3 \leq \theta_i \leq 2/3$.

5. Time for computation of θ 's is printed.

Page(s) 3

1. If detailed printout is requested, the following is printed:

- a. Item number and iteration number
- b. Information matrix and score vector

If c is held constant, only a 2×2 information matrix and the score vector will be printed as follows

$$\begin{array}{ccc} I_{aa} & I_{ab} & S_a \\ I_{ba} & I_{bb} & S_b \end{array}$$

If c isn't being held constant a 3×3 information matrix and the score vector is printed

$$\begin{array}{cccc} I_{aa} & I_{ab} & I_{ac} & S_a \\ I_{ab} & I_{bb} & I_{bc} & S_b \\ I_{ac} & I_{bc} & I_{cc} & S_c \end{array}$$

c. After the information matrix the following is printed

DEL A = change in a_i

A = new value for a_i

DEL B = change in b_i

B = new value for b_i

DEL C = change in c_i

C = new value for c_i

d. If any of the parameters change by more than the maximum change or any go out of range, a message that the parameters have been adjusted is printed followed by a repeat of line (c) for the adjusted values.

2. At the end of the detailed printout for each item or if the detailed printout isn't requested, a line described by the title at the top of page 3 is printed. This contains

- a. Item number
 - b. Derivatives of the log of the likelihood function for a , b , and c
 - c. Deltas for a , b , and c
 - d. Iteration number
 - e. An "*" if the maximum number of iterations has been reached
 - f. If a is fixed, the value that a is fixed at is printed
 - g. If c is fixed, the value that c is fixed at is printed and whether it was fixed by exceeding the upper or lower bound.
3. Following all the item iterations, the new values of a , b , and c are printed.

Page 4

1. Maximum derivatives for a , b , and c and the maximum number of iterations is printed.
2. Maximum theta derivative and maximum number of iterations for the thetas.
3. The likelihood function after the last θ estimations is printed.
4. Time for the total item iterations.
5. "End of stage" followed by the stage number.
6. Total time for the stage.

Page 5

1. If any $\psi_{ia} > 19$, and the corresponding x_{ia} is 1, a message is printed giving the index of the ability. Ignore this message.
2. Likelihood function after change in item parameters.
3. Message that the likelihood function has increased, if it has since the initial computation of the likelihood function or since the previous computation of the likelihood function after a change in the item

parameters. If the likelihood function hasn't increased, no message is printed.

Repeat from page 2 through page 5 until either the procedure converges, the maximum number of stages has been reached, or the time limit is exceeded.

Convergence is determined when the likelihood function computed after a change in item parameters, has decreased in two out of four stages. After it has converged, two more stages are done with detailed printout. If the maximum number of stages has been reached or there isn't enough time for another stage, the program ends after the item parameter estimation when it prints "Time for Stage."

Last Page

At the end of the run the final θ 's are printed, 12 θ 's per line, with the first number in the line the sequence number of the first θ in the line.

Punched Output

At the end of a run the following cards are punched so that the program can be restarted.

Card 1

1-60	Title
64-65	Number of items
66-72	Number of examinees

Card 2

Date and time of run (5A4)

Card Set 3 (I4, 19A4)

Θ's punched 19 per card. The first 4 columns contain the sequence number for the card, the Θ's are punched in 19A4 format.

Card Set 4

Format for the a's which is (5E15.8).

The a's, punched in 5E15.8 format.

Card Set 5

Format for the b's (5E15.8).

The b's punched in 5E15.8 format.

Card Set 6

The c's punched in 8F10.7 format.

Card Set 7

The number of parameters that are allowed to vary for each item punched in 1615 format.