DOCUMENT RESUME

ED 336 397 TM 017 040

| | |
|---|---|
| AUTHOR | Chang, Hua-Hua; Stout, William |
| TITLE | The Asymptotic Posterior Normality of the Latent Trait in an IRT Model. |
| INSTITUTION | Illinois Univ., Urbana. Dept. of Statistics. |
| SPONS AGENCY | Office of Naval Research, Arlington, VA. Cognitive and Neural Sciences Div. |
| PUB DATE | 24 May 91 |
| CONTRACT | ONR-N00014-90-J-1940-4421-548 |
| NOTE | 49p. |
| PUB TYPE | Reports - Evaluative/Feasibility (142) |
| EDRS PRICE | MF01/PC02 Plus Postage. |
| DESCRIPTORS | Ability; *Bayesian Statistics; *Equations (Mathematics); *Item Response Theory; *Mathematical Models; Probability; *Psychological Testing; *Sampling |
| IDENTIFIERS | *Asymptotic Distributions; Convergence (Mathematics) |

ABSTRACT

The empirical Bayes modeling approach--latent ability random sampling in the item response theory (IRT) context--to the IRT modeling of psychological tests is described. Under the usual empirical Bayes unidimensional IRT modeling approach, the posterior distribution of examinee ability given test response is approximately normal for a long test. Three theorems are developed to establish the asymptotic posterior normality of latent variable distributions. Implications of the results are discussed. An appendix contains proofs of the theorems, in terms of proof of convergence in probability, proof of strong convergence, and proof of convergence in manifest probability. A 16-item list of references is included. (SLD)

SECURITY CLASSIFICATION OF THIS PAGE

# REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

| | |
|---|---|
| 1a REPORT SECURITY CLASSIFICATION<br>Unclassified | 1b RESTRICTIVE MARKINGS |
| 2a SECURITY CLASSIFICATION AUTHORITY | 3 DISTRIBUTION/AVAILABILITY OF REPORT<br>Approved for public release; distribution unlimited |
| 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE | |
| 4. PERFORMING ORGANIZATION REPORT NUMBER(S)<br>1991 - #4 | 5. MONITORING ORGANIZATION REPORT NUMBER(S) |
| 6a. NAME OF PERFORMING ORGANIZATION<br>University of Illinois<br>Depaitment of Statistics<br>6b OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION<br>Cognitive Science Program<br>Office of Naval Research (Code 1142 CS) |
| 6c. ADDRESS (City, State, and ZIP Code)<br>101 Illini Hall<br>725 S. Wright Street<br>Champaign, IL 61820 | 7b ADDRESS (City, State, and ZIP Code)<br>( ) N. Quincy<br>.. lington, VA 22217-5000 |
| 8a. NAME OF FUNDING/SPONSORING ORGANIZATION<br>8b OFFICE SYMBOL (If applicable) | 9 PROCUREMENT INSTRUMENT IDENTII CATION NUMBER<br>N00014-90-J-1940 |
| 8c. ADDRESS (City, State, and ZIP Coue)<br>101 Illini Hall<br>725 S. Wright Street<br>Champaign, IL 61820 | 10 SOURCE OF FUNDING NUMBERS<br>PROGRAM ELEMENT NO: 61153N<br>PROJECT NO: RR04204<br>TASK NO: RR04204-01<br>WORK UNIT ACCESSION NO: 4421-548 |

11 TITLE (Include Security Classification)
The Asymptotic Posterior Normaliiy of the Latent Trait in an IRT Model

12 PERSONAL AUTHOR(S)
Hua Hua Chang and William Stout

| 13a. TYPE OF REPORT | 13b TIME COVERED | 14 DATE OF REPORT (Year, Month, Day) | 15 PAGE COUNT |
|---|---|---|---|
| technical | FROM 1988 TO 1991 | May 25, 1991 | 43 |

16 SUPPLEMENTARY NOTATION

| 17 COSATI CODES | | | 18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB-GROUP | See reverse |
| | | | |
| | | | |

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

See reverse

| | |
|---|---|
| 20 DISTRIBUTION/AVAILABILITY OF ABSTRACT<br>☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS | 21 ABSTRACT SECURITY CLASSIFICATION |
| 22a NAME OF RESPONSIBLE INDIVIDUAL<br>Dr. Charles E. Davis | 22b TELEPHONE (Include Area Code) (703) 696-4046<br>22c OFFICE SYMBOL ONR-1142-CS |

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

It has long been part of the Item Response Theory (IRT) folklore that under the usual empirical Bayes unidimensional IRT modeling approach, the posterior distribution of examinee ability given test response is approximately normal for a long test. Under very general non-parametric assumptions, we make this claim rigorous for a broad class of latent models.

**Key words**: item response theory, empirical Bayes, posterior distribution, ability estimation, confidence interval, manifest probability.

# 1 Introduction

This article deals with an empirical Bayes modeling approach (by which is meant latent ability random sampling in the IRT context) to the item response theory (IRT) modeling of psychological tests. Suppose we randomly sample N persons from a specified population, and then administer a test consisting of n items. The data structure for a randomly selected examinee can be expressed by a random vector

$$( X_1, \ldots, X_n, \theta),$$

where $X_1, \ldots, X_n$ denote item responses and $\theta$ denotes examinee ability, which is unobservable. Abstractly, in an empirical Bayes problem the data is modeled by independent identically distributed (i.i.d.) random vectors

$$(X_1^{(1)}, ..., X_n^{(1)}, \theta_1),\ (X_1^{(2)}, ..., X_n^{(2)}, \theta_2), .., \ (X_1^{(N)}, ..., X_n^{(N)}, \theta_N).$$

One important measurement goal is the estimation/prediction of each examinee's $\theta$. Clearly one should use the first examinee response $X_1^{(1)}, ..., X_n^{(1)}$ to predict the actual value of $\theta_1$. However, unless the distribution of $\theta$ is completly specified, there is useful information in

$$(X_1^{(2)}, ..., X_n^{(2)}),\ (X_1^{(3)}, ..., X_n^{(3)}), ..., (X_1^{(N)}, ..., X_n^{(N)}),$$

the second through $N$th examinee responses, about the unknown distribution of $\theta$ and thus about the unknown ability $\theta_1$ in particular, which we want to estimate. Thus an alternative approach to using only $(X_1^{(1)}, ..., X_n^{(1)})$ is to use all of the test responses in making inferenses about $\theta_1$.

Let $X_j$ be the score for a randomly selected examinee on the $j$th item; $X_j = 1$ if the answer is correct, $X_j = 0$ if in correct, and let

$$X_j = \begin{cases} 1 & \textit{with probability } P_j(\theta) \\ 0 & \textit{with probability } 1 - P_j(\theta) \end{cases}$$

where $P_j(\theta)$ denotes the probability of correct response for a randomly chosen examinee of ability $\theta$, that is,

$$P_j(\theta) = P\{X_j = 1|\theta\},$$

where $\theta$ is unknown and has the domain $(-\infty, \infty)$ or some subinterval on $(-\infty, \infty)$. We make two assumptions about the IRT models of this paper:

**(a) Local Independence** (also called Conditional Independence)

$$\begin{aligned} P_n(x_1, \ldots, x_n|\theta) &\stackrel{\text{def}}{=} P\{(X_1, \ldots, X_n) = (x_1, \ldots, x_n)|\theta\} \\ &= \prod_{j=1}^{n} P\{X_j = x_j|\theta\} \\ &= \prod_{j=1}^{n} P_j(\theta)^{x_j}[1 - P_j(\theta)]^{1-x_j}. \end{aligned}$$

**(b) Monotonicity:** each $P_j(\theta)$ is strictly increasing in $\theta$.

Lord (1980) makes an interesting remark about the existence of a prior distribution for ability:

> *"In work with published tests, it is usual to test similar groups of examinees year after year with parallel forms of the same test. When this happens, we can form a good picture of the frequency distribution of ability in the next group of examinees to be tested."*

This suggests taking an empirical Bayes approach to IRT modeling, in particular assuming partial knowledge about the distribution of $\theta$ and thereby being able to make efficient use of the response data to make inferences about the distribution of $\theta$ and thus make inferences about the unobservable examinee abilities. The distribution of a test response $X_1, \ldots, X_n$ is indexed by $\theta$, which belongs to the parameter space $\Theta$; that is, each $\theta \in \Theta$ governs a test response distribution. Let $L_n(\theta)$ denote the log-likelihood, that is

$$L_n(\theta) = log\{P_n(X_1, \ldots, X_n|\theta)\}.$$

If we assume that the prior distribution has density $\Pi(\theta)$, according to Bayes' theorem, the posterior density for each given

$$(X_1, \ldots X_n) = (x_1, \ldots x_n)$$

can be written as

$$\begin{aligned} \Pi_n(\theta | x_1, \ldots, x_n) &= \frac{P_n(x_1, \ldots, x_n | \theta)\Pi(\theta)}{P_n(x_1, \ldots, x_n)} \\ &= \frac{\exp\{L_n(\theta)\}\Pi(\theta)}{P_n(x_1, \ldots, x_n)} \end{aligned} \tag{1}$$

where

$$P_n(x_1, \ldots, x_n) = \int_\Theta P_n(x_1, \ldots, x_n | \theta)\Pi(\theta) d\theta.$$

Notice that, the **"prior"** and **"posterior"** refer to the relationship between the distributions and the observation $x_1, \ldots, x_n$. E.g., $\Pi(\theta)$ is prior to $x_1, \ldots, x_n$ and

$$\Pi_n(\theta | x_1, \ldots, x_n)$$

is posterior to $x_1, \ldots, x_n$. These ideas can be easily extended to the study of the asymptotic behaviour of the posterior distribution. In particular, for each $x_1, \ldots, x_n$, what can be said about the posterior probability of $\theta$ as $n$ tends to infinity?

It has long been part of the IRT folklore that under the usual empirical Bayes unidimensional IRT modeling approach, the posterior distribution of $\theta$ given test response is approximately normal for a long test. Holland (1990) indicates:

> *"At present I know of no through discussion of the asymptotic posterior normality of latent variable distributions and this would appear to be an interesting area for further research."*

In classical statistics, when $(X_1, \ldots, X_n)$ are i.i.d., an important result (informally stated) is that, for $n$ large, the posterior density $\Pi_n(\theta | X_1, \ldots, X_n)$ is approximately equal to the normal density $N(\hat{\theta}_n, \hat{\sigma}_n^2)$, where $\hat{\theta}_n$ is the maximum-likelihood estimator (or MLE) of $\theta$ and $\hat{\sigma}_n^2 \stackrel{\text{def}}{=} \{-L_n''(\hat{\theta}_n)\}^{-1}$, where $L_n''(\hat{\theta}_n)$ is the second derivative with respect to $\theta$ of the log-likelihood evaluated at $\hat{\theta}_n$. $\hat{\theta}_n$ and $\hat{\sigma}_n^2$ here are functions of $(X_1, \ldots, X_n)$ only. Intuitively, $\hat{\sigma}_n^2 \to 0$ in applications, usually like $1/n$.

Lindley(1965) proposed a heuristic approach to prove the above result by expanding the log-likelihood in Taylor series in $\theta$ about $\hat{\theta}_n$,

$$L_n(\theta) = L_n(\hat{\theta}_n) + \frac{1}{2}(\theta - \hat{\theta}_n)^2 L_n''(\hat{\theta}_n) + R_n,$$

where $R_n$ is a remainder term. Since the log-likelihood has a maximum at $\hat{\theta}_n$ the first derivative vanishes there. As shown above the posterior density viewed as a function of $\theta$ for fixed $x_1, \ldots, x_n$ is proportional to

$$\Pi(\theta) exp\{L_n(\theta)\}.$$

Therefore,

$$\Pi_n(\theta |\, x_1, \ldots, x_n) \propto \Pi(\theta) exp\{L_n(\hat{\theta}_n) - \frac{(\theta - \hat{\theta}_n)^2}{2\hat{\sigma}_n^2} + R_n\}.$$

Since $L_n(\hat{\theta}_n)$ does not involve $\theta$, it may be absorbed into the omitted constant of proportionality so that

$$\Pi_n(\theta |\, x_1, \ldots, x_n) \propto \Pi(\theta) exp\{-\frac{(\theta - \hat{\theta}_n)^2}{2\hat{\sigma}_n^2} + R_n\}, \quad (2)$$

where the remainder, $R_n$, is claimed to be negligible when compared with the other term in (2). Because $\hat{\sigma}_n^2 \to 0$ like $1/n$, the density in (2) becomes concentrated at $\hat{\theta}_n$ in the limit, thus allowing $\Pi(\theta)$ to also be absorbed into the omitted constant of proportionality. Thus,

$$\Pi_n(\theta |\, x_1, \ldots, x_n) \propto exp\{-\frac{(\theta - \hat{\theta}_n)^2}{2\hat{\sigma}_n^2}\}$$

as desired. However, Lindley (1965) did not give a rigorous proof.

Walker(1969) proved that under certain conditions, the posterior probability of $\hat{\theta}_n + a\hat{\sigma}_n < \theta < \hat{\theta}_n + b\hat{\sigma}_n$, namely

$$\int_{\hat{\theta}_n+a\hat{\sigma}_n}^{\hat{\theta}_n+b\hat{\sigma}_n} \Pi_n(\theta|\ X_1,\dots,X_n)d\theta,$$

converges in probability $P_{\theta_0}$ to

$$(2\pi)^{-1/2}\int_a^b e^{-\frac{1}{2}y^2}dy$$

as $n \to \infty$. Here, as the notation $P_{\theta_0}$ indicates, in the generation of $X_1,\dots,X_n$ we assume $\theta_0$ is the true value of $\theta$. That is $X_1,\dots,X_n$ is generated according to the distribution $P_n(\ x_1,\dots,x_n|\theta_0)$. Then, using the rules of conditional probability computation, it is easy to show that one way to interpret Walker's result is that

$$P[\hat{\theta}_n + a\hat{\sigma}_n < \theta_0 < \hat{\theta}_n + b\hat{\sigma}_n|\ X_1,\dots,X_n,\theta_0]$$

converges in probability to

$$(2\pi)^{-1/2}\int_a^b e^{-\frac{1}{2}y^2}dy$$

as $n \to \infty$. That is, for each fixed (but unknown) $\theta_0$ we have an asymptotic confidence interval for each choice of $a < b$.

As we know, for all realistic applications, the item characteristic curves are not identical. Therefore, the $\{X_j\}$ we have are merely independent, conditional on $\theta$, but not identically distributed. However, the general IRT model enables us to prove, by adapting the approach that Walker (1969) applied to *i.i.d.* random variables,

**(a)** The "weak" convergence, that is, for $-\infty \le a < b \le \infty$,

$$A_n \equiv \int_{\hat{\theta}_n+a\hat{\sigma}_n}^{\hat{\theta}_n+b\hat{\sigma}_n} \Pi_n(\theta|\ X_1,\dots,X_n)d\theta$$

converges in probability $P_{\theta_0}$ to

$$A \equiv (2\pi)^{-1/2} \int_a^b e^{-\frac{1}{2}y^2} dy$$

as $n \to \infty$. That is,

$$P_{\theta_0}\{|A_n - A| < \epsilon\} \to 1, \quad as\ n \to \infty, \quad for\ arbitrary\ \epsilon > 0.$$

**(b)** The strong convergence of $A_n$: that is,

$$P_{\theta_0}\{\lim_{n\to\infty} A_n = A\} = 1;$$

**(c)** Convergence in "manifest" probability, or "$\theta_0$ free" convergence, that is, $A_n$ converges to $A$ in the manifest (or marginal in the sense that $\theta_0$ is integrated out) probability $P$, which is defined, for any fixed $n$

$$\begin{aligned} P\{( X_1, \ldots, X_n) &= ( x_1, \ldots, x_n)\} \\ &= \int_\Theta P_n( x_1, \ldots, x_n|\theta)\pi(\theta)d\theta. \end{aligned}$$

This result is also easily interpretable as an asymptotic confidence inteval for ability. That is, it assures that

$$P\{\hat{\theta}_n + a\hat{\sigma}_n < \theta < \hat{\theta}_n + b\hat{\sigma}_n|\ X_1, \ldots, X_n\}$$

converges in probability to

$$(2\pi)^{-1/2} \int_a^b e^{-\frac{1}{2}y^2} dy$$

as $n \to \infty$. That is, for any randomly sampled examinee, we have an asymptotic confidence inteval for each choice of $a < b$. Here in (c), in contrast to (a), the value of $\theta$ for the randomly sampled examinee is not fixed.

**(d)** The weak and strong consistency of the MLE $\hat{\theta}_n$, which are intermediate results in the proofs of (a) and (b).

Proving (a)-(c) is the main purpose of this paper, thereby meeting the Holland challenge quoted above.

# 2 Further Notation and Assumptions

## 2.1 Basic Notation

$\theta_0$: The true parameter. In saying that $X_j$ is a random variable we infer that $X_j$ has the density

$$P_j(\theta)^{x_j}[1 - P_j(\theta)]^{1-x_j}, \quad x_j = 0, 1,$$

for some fixed value of $\theta$. Denote this value by $\theta_0$, which we refer to as the true parameter.

$\hat{\theta}_n$: The Maximum Likelihood Estimator(MLE) of $\theta$, which is defined as a solution (in general non-unique), of

$$P_n( X_1, \ldots, X_n|\hat{\theta}_n) = \max_{\theta \in \Theta}\{P_n( X_1, \ldots, X_n|\theta)\}, \tag{3}$$

if it exists, or equivalently, of

$$L_n(\hat{\theta}_n) = \max_{\theta \in \Theta}\{L_n(\theta)\}. \tag{4}$$

$I_j(\theta)$: The item information function of item j, which is equal to

$$I_j(\theta) = \frac{\{P_j'(\theta)\}^2}{P_j(\theta)[1 - P_j(\theta)]},$$

where $P_j'(\theta)$ is the first derivative of $P_j(\theta)$ with respect to $\theta$.

$I^{(n)}(\theta)$: The test information function

$$I^{(n)}(\theta) = \sum_{j=1}^{n} I_j(\theta).$$

$\hat{\sigma}_n^2$:

$$\hat{\sigma}_n^2 \stackrel{\text{def}}{=} \{I^{(n)}(\hat{\theta}_n)\}^{-1}, \tag{5}$$

noting that our definition of $\hat{\sigma}_n^2$ used hereafter in the paper differs from the often used $\hat{\sigma}_n^2 \stackrel{\text{def}}{=} \{-L_n''(\hat{\theta}_n)\}^{-1}$ mentioned above.

$\lambda_j(\theta)$: The logit function of item j

$$\lambda_j(\theta) = log\{\frac{P_j(\theta)}{1 - P_j(\theta)}\}. \tag{6}$$

$Z_j(\theta)$:

$$Z_j(\theta) = log\{\frac{P_j(\theta)^{X_j}[1 - P_j(\theta)]^{1-X_j}}{P_j(\theta_0)^{X_j}[1 - Pj(\theta_0)]^{1-X_j}}\}. \tag{7}$$

## 2.2 Regularity Conditions

Some "regularity" conditions and their explanations will be stated before going into details about our theorems. Fix $\theta_0 \in \Theta$: There are five basic assumptions:

**(A1):** Let $\theta \in \Theta$, where $\Theta$ is $(-\infty, \infty)$ or a bounded or unbounded interval in $(-\infty, \infty)$. Let the prior density $\Pi(\theta)$ be continuous and positive at $\theta_0$, where $\theta_0$ is assumed be the true value of $\theta$.

**(A2):** $P_j(\theta)$ is twice continuously differentiable and $P_j'(\theta)$ and $P_j''(\theta)$ are bounded in absolute value uniformly with respect to both $\theta$ and $j$ in some closed interval $N_0$ of $\theta_0 \in \Theta$.

**(A3):** For every fixed $\theta \neq \theta_0$, assume for some given $c(\theta) > 0$

$$\overline{\lim_{n\to\infty}} n^{-1} \sum_{j=1}^{n} E_{\theta_0} Z_j(\theta) \ \leq \ -c(\theta) \tag{8}$$

and

$$\sup_j |\lambda_j(\theta)| < \infty.$$

(See Footnote[1].) Note that

$$L_n(\theta) - L_n(\theta_0) = \sum_{j=1}^{n} Z_j(\theta).$$

[1]For a sequence of real number $\{a_n\}$, if $\lim_{n\to\infty} a_n$ does not exist, then $\{a_n\}$ must have more than one limit point. $\overline{lim}_{n\to\infty} a_n$ denotes the largest limit point (or upper limit).

**(A4):** $\{I_j'(\theta)\}$ and $\{\lambda_j''(\theta)\}$ and $\{\lambda_j'''(\theta)\}$ are bounded in absolute value uniformly in j and in $\theta \in N_0$, $N_0$ specified in (A2) above.

**(A5):**

$$\liminf_{n\to\infty} \frac{I^{(n)}(\theta_0)}{n} > c(\theta_0) > 0.$$

That is, asymptotically, the average information at $\theta_0$ is bounded away from 0.

Although $\Theta$ may be $(-\infty, \infty)$, we always assume without loss of generanality that $\theta_0$ is contained in a finite interval, e.g. $[-a, a]$ for some fixed $a > 0$. This is because from the psychometric viewpoint, taking $var(\theta)$ =1 for convenience, the same educational decision is made about people with $\theta = 4$ and people with $\theta = 24$. Thus, assuming $-5 \leq \theta \leq 5$ does no practical damage.

The condition (8) of assumption (A3), perhaps, looks unfamiliar. But it plays an important role in the proof of Lemma 3.1 below, ensuring the identifiability of $\theta_0$. That is, when $\theta_0$ is the true value of $\theta$, $E\{L_n(\theta) - L_n(\theta_0)\}$ should be sufficiently negative for all values of $\theta \neq \theta_0$ . In other words, this condition allows us to "identify" $\theta_0$ by maximizing the likelihood function. (A3) acts as a remedy in the case that $\{X_j\}$ are merely independent but not identically distributed. In other words, if they are i.i.d., as is the case in Walker's proof, then (A3) is automatically satisfied. To see this, note in the i.i.d. case that

$$n^{-1} \sum_{j=1}^{n} E_{\theta_0}\{Z_j(\theta)\} = E_{\theta_0}\{Z_1(\theta)\}.$$

Note that

$$E_{\theta_0} exp\{Z_1(\theta)\} = P_1(\theta_0)\frac{P_1(\theta)}{P_1(\theta_0)} + (1 - P_1(\theta_0))\frac{1 - P_1(\theta)}{1 - P_1(\theta_0)} \equiv 1.$$

Thus, since $-logx$ is strictly convex, Jensen's inequality (Lehmann, p50) shows that for arbitrary $\theta$

$$E_{\theta_0} Z_1(\theta) \equiv E_{\theta_0}[log\{Y(\theta)\}] < log\{E_{\theta_0}[Y(\theta)]\} \equiv 0, \tag{10}$$

where

$$Y(\theta) = exp\{Z_1(\theta)\}.$$

Thus (8) is satisfied by taking

$$c(\theta) = -E_{\theta_0}\{Z_1(\theta)\}.$$

Unfortunately $\{Z_j(\theta)\}$ in IRT models are not identically distributed, so we have to impose some supplementary condition. According to (10), $n^{-1}\sum_{j=1}^{n} E_{\theta_0} Z_j(\theta)$ will be negative, however, this does not enable ı ȝ to obtain (8). For what classes of IRT models then does (8) hold? Consider the case in which each $E_{\theta_0} Z_j(\theta)$ satisfies, for some $c(\theta)$,

$$E_{\theta_0} Z_j(\theta) \leq -c(\theta) < 0. \tag{11}$$

It is obvious that (8) holds. However, this condition is stronger than needed. It would suffice to merely require that a **"certain proportion" of the $E_{\theta_0} Z_j(\theta)$s satisfy condition (11),** say one in every K, no matter how large the K is. Mathematically speaking, this would imply

$$n^{-1}\sum_{j=1}^{n} E_{\theta_0} Z_j(\theta) \leq n^{-1}\{n\frac{-c(\theta)}{K}\} = \frac{-c(\theta)}{K} \equiv -\tilde{c}(\theta) < 0,$$

and so

$$\overline{\lim_{n\to\infty}} n^{-1}\sum_{j=1}^{n} E_{\theta_0} Z_j(\theta) \leq -\tilde{c}(\theta) < 0.$$

Actually, (8) does not seem very restrictive in IRT models incurred in practice. As evidence, consider a "typical" IRT model of 40 3PL items, in which the item parameters are precalibrated from a real ACT math test. The graphs illustrated in Figure 1 are the $E_{\theta_0} Z_j(\theta)$s computed from this model. Clearly (8) seems to be holding.

(A4) and (A5) are used to make $L_n''(\theta)$ behave sufficiently well for $\theta$ near $\theta_0$. Condition (A5) implies that the test information function evaluated at $\theta_0$ tends to infinity with the same speed as $n$. These five conditions would not be difficult to verify in

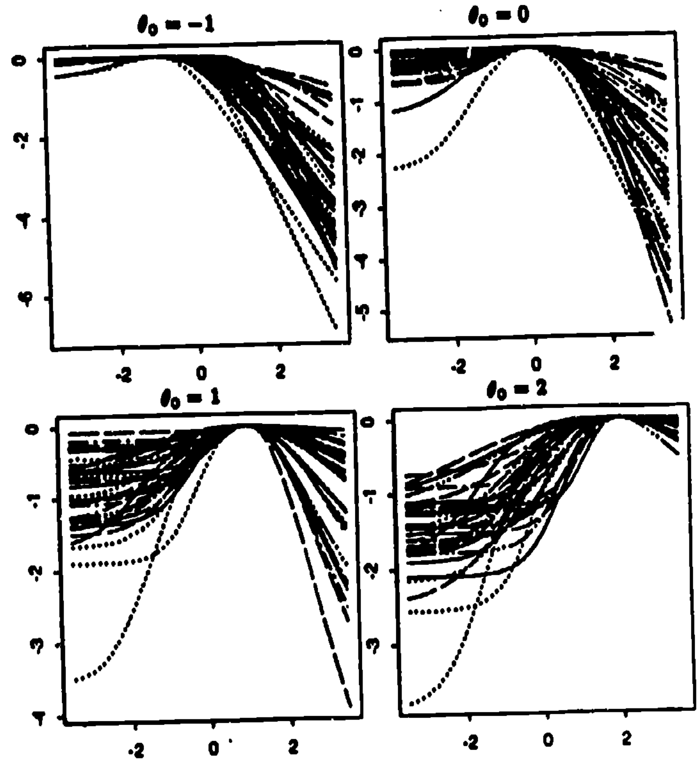


Figure 1: $E_{\theta_0}\{Z_j(\theta)\}$s for 40 items, ACT-MATH Test (Drasgow, 1987).

particular applications and hence are really fairly mild modeling assumptions.

# 3 The Main Theorems

In this section we will introduce three theorems and the major steps of the proof of Theorem 3.1, the basic theorem. The rigorous proofs of these theorems, as well as their related lemmas and corollaries, are contained in an appendix.

## 3.1 Convergence in Probability

**Theorem 3.1** *Suppose that conditions (A1) through (A5) hold. Let $\hat{\theta}_n$ be an MLE of $\theta_0$, and $\hat{\sigma}_n$ be the square root of $\{I^{(n)}(\hat{\theta}_n)\}^{-1}$. Then, for $-\infty \leq a < b \leq \infty$, the posterior probability of $\hat{\theta}_n + a\hat{\sigma}_n < \theta < \hat{\theta}_n + b\hat{\sigma}_n$, namely*

$$\int_{\hat{\theta}_n + a\hat{\sigma}_n}^{\hat{\theta}_n + b\hat{\sigma}_n} \Pi_n(\theta \mid X_1, \ldots, X_n) d\theta,$$

*tends in $P_{\theta_0}$ to*

$$(2\pi)^{-1/2}\int_a^b e^{-\frac{1}{2}u^2}du,$$

*as $n \to \infty$.*

Theorem 3.1 is the basic result in our asymptotic posterior normality work. Note that $A_n$ is a random variable depending on $X_1, \ldots, X_n$. Thus its distribution is determined by the parameter $\theta_0$ and $A_n \to A$ *in* $P_{\theta_0}$ means

$$\lim_{n\to\infty} P_{\theta_0}\{|A_n - A| < \epsilon\} = 1, \quad \textit{for arbitrary } \epsilon > 0.$$

**Outline of Proof.** To prove the theorem, write

$$\int_{\hat{\theta}_n+a\hat{\sigma}_n}^{\hat{\theta}_n+b\hat{\sigma}_n} \Pi_n(\theta|\, X_1,\ldots,X_n)d\theta = \frac{G}{P_n(\,X_1,\ldots,X_n)}$$
$$= \frac{G}{P_n(\,X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n}\left(\frac{P_n(\,X_1,\ldots,X_n)}{P_n(\,X_1,\ldots,X_n|\hat{\vartheta}_n)\hat{\sigma}_n}\right)^{-1}$$

where

$$G = \int_{\hat{\theta}_n+a\hat{\sigma}_n}^{\hat{\theta}_n+b\hat{\sigma}_n} \Pi(\theta)P_n(\,X_1,\ldots,X_n|\theta)d\theta, \tag{12}$$

and

$$P_n(\,X_1,\ldots,X_n) = \int_\Theta \Pi(\theta)P_n(\,X_1,\ldots,X_n|\theta)d\theta.$$

It suffices to prove

$$\frac{P_n(\,X_1,\ldots,X_n)}{P_n(\,X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} \to (2\pi)^{1/2}\Pi(\theta_0) \tag{13}$$

as $n \to \infty$, in $P_{\theta_0}$, and

$$\frac{G}{P_n(\,X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} \to (2\pi)^{1/2}\Pi(\theta_0)\{\Phi(a) - \Phi(b)\} \tag{14}$$

as $n \to \infty$, in $P_{\theta_0}$, where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-\frac{1}{2}u^2}du$.

In the following we will present the general idea to prove (13). ((14) is proved by the similar method.) First expand $L_n(\theta)$ at $\hat{\theta}_n$ by Taylor expansion: we have

$$\begin{aligned} L_n(\theta) - L_n(\hat{\theta}_n) &= \frac{(\theta-\hat{\theta}_n)^2}{2} L_n''(\theta_n^*) \\ &= -\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1-R_n), \end{aligned} \tag{15}$$

where $\theta_n^*$ is a point between $\theta$ and $\hat{\theta}_n$, and $\hat{\sigma}_n^2$ is defined by (5) and $R_n$ is defined by:

$$\begin{aligned} R_n \stackrel{\text{def}}{=} R_n(\theta,\ X_1,\ldots,X_n) &= 1+\hat{\sigma}_n^2 L_n''(\theta_n^*) \\ &= \{L_n''(\theta_n^*) + I^{(n)}(\hat{\theta}_n)\}/I^{(n)}(\hat{\theta}_n). \end{aligned} \tag{16}$$

Split $P_n(X_1,\ldots,X_n)$ into two parts as follows

$$\begin{aligned} P_n(X_1,\ldots,X_n) &= \int_{|\theta-\theta_0|\geq\delta} \Pi(\theta)P_n(X_1,\ldots,X_n|\theta)d\theta \\ &+ \int_{|\theta-\theta_0|<\delta} \Pi(\theta)P_n(X_1,\ldots,X_n|\theta)d\theta \\ &\stackrel{\text{def}}{=} G_1 + G_2. \end{aligned} \tag{17}$$

Therefore, recalling that $L_n(\theta) = \log P_n(X_1,\ldots,X_n|\theta)$,

$$\begin{aligned} \frac{G_1}{P_n(X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} &= exp\{L_n(\theta_0) - L_n(\hat{\theta}_n)\}\{I^{(n)}(\hat{\theta}_n)\}^{1/2} \\ &\times \int_{|\theta-\theta_0|\geq\delta} \Pi(\theta)exp\{L_n(\theta)-L_n(\theta_0)\}d\theta \end{aligned} \tag{18}$$

and, using (15),

$$\frac{G_2}{P_n(X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} = \frac{\Pi(\theta_0)}{\hat{\sigma}_n}\int_{|\theta-\theta_0|<\delta}\frac{\Pi(\theta)}{\Pi(\theta_0)}\exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1-R_n)\}d\theta. \tag{19}$$

Thus, if

$$\frac{G_1}{P_n(X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} \longrightarrow 0 \ \ in\ P_{\theta_0} \tag{20}$$

and

$$\frac{G_2}{P_n(X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} \rightarrow (2\pi)^{1/2}\Pi(\theta_0) \ \ in\ P_{\theta_0}, \tag{21}$$

then (13) holds. For establishing (20), first consider (18): If $\hat{\theta}_n$ is consistent then $exp\{L_n(\theta_0) - L_n(\hat{\theta}_n)\}$ goes to a constant as $n$ approaches $\infty$. On the other hand, since $\{I^{(n)}(\hat{\theta}_n)\}^{1/2}$ approaches $\infty$ like $n^{1/2}$, we need to make $L_n(\theta) - L_n(\theta_0)$ **"sufficiently negative"** so that the integral of (18) approaches 0 faster than $n^{-1/2}$ and hence the left hand side of (20) can be neglected outside the $\delta$ region of $\theta_0$. As for establishing (21), consider (19): Since $\Pi(\theta)$ is continous, $\Pi(\theta)/\Pi(\theta_0)$ will be close to one for $\delta$ sufficiently small, and we need to make $R_n$ **"sufficiently small"** inside the $\delta$ region so that we can estimate the integral by

$$\int_{|\theta-\theta_0|<\delta} \exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}\} d\theta.$$

Mathematically speaking, we need the following two lemmas.

**Lemma 3.1** *Suppose that conditions (A1) through (A3) hold. For any $\delta > 0$, there exists $k(\delta) > 0$ such that*

$$\lim_{n\to\infty} P_{\theta_0}\{\sup_{|\theta-\theta_0|\geq\delta} n^{-1}[L_n(\theta) - L_n(\theta_0)] < -k(\delta)\} = 1.$$

**Lemma 3.2** *Suppose that conditions (A1) through (A5) hold. Then*

$$L_n(\theta) - L_n(\hat{\theta}_n) = (\theta - \hat{\theta}_n)^2 L_n''(\theta_n^*)/2 = -\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1 - R_n), \qquad (22)$$

*where $\theta_n^*$ is a point between $\theta$ and $\hat{\theta}_n$, and $R_n$ is defined by (16). Also, for any $\varepsilon > 0$, there exists $\delta$ such that*

$$\lim_{n\to\infty} P\{\sup_{|\theta-\theta_0|<\delta} |R_n(\theta, X_1, \ldots, X_n)| < \varepsilon\} = 1. \qquad (23)$$

As a by-product, Lemma 3.1 ensures the consistency of the MLE $\hat{\theta}_n$, which is labeled as Corollary 3.1.

**Corollary 3.1** *Suppose that conditions (A1) through (A3) hold. Than $\hat{\theta}_n$ is weakly consistent, namely*

$$\lim_{n\to\infty} \hat{\theta}_n = \theta_0 \quad in\ P_{\theta_0}. \qquad (24)$$

It can be shown that (22) of Lemma 3.2 makes it possible for us to use the reciprocal of the test information as the variance estimate (see (5)), instead of

$$\hat{\sigma}_n^2 \stackrel{\text{def}}{=} \{-L_n''(\hat{\theta}_n)\}^{-1},$$

as Lindley (1965) and Walker (1969) each suggested. The variance estimate (5) we have chosen has the following advantages:

- The information function $I^{(n)}(\ )$ is always positive. $-L_n''(\ )$, by contrast, could be negative, especially when the sample size is not large enough. So, some times $\{-L_n''(\ )\}^{1/2}$ may not exist.
- The information function is easier to calculate, while the calculation of $L_n''(\ )$ is more complicated.

Future study should be undertaken to compare the speed of the convergence and to explore any further advantages.

## 3.2 Convergence Almost Surely

As discussed in the preceding subsection, the posterior distribution for $\frac{\hat{\theta}_n-\theta}{\hat{\sigma}_n}$, derived from a proper prior density $\Pi(\theta)$, **converges in probability** to the standard normal distribution. In this subsection we will see that a stronger result, **convergence almost surely, (also referred to as strong, almost everywhere, or with probability one convergence)**, can be achieved under the same assumptions.

**Theorem 3.2** *Suppose that conditions (A1) through (A5) hold. Let $\hat{\theta}_n$ be an MLE of $\theta_0$, and $\hat{\sigma}_n$ be the square root of $\{I^{(n)}(\hat{\theta}_n)\}^{-1}$. Then, for $-\infty \le a < b \le \infty$, the posterior probability of $\hat{\theta}_n + a\hat{\sigma}_n < \theta < \hat{\theta}_n + b\hat{\sigma}_n$, namely*

$$A_n \equiv \int_{\hat{\theta}_n+a\hat{\sigma}_n}^{\hat{\theta}_n+b\hat{\sigma}_n} \Pi_n(\theta|\ X_1,\ldots,X_n)d\theta,$$

*tends to*

$$A \equiv (2\pi)^{-1/2} \int_a^b e^{-\frac{1}{2}u^2} du \quad \textit{almost surely},$$

*cs* $n \to \infty$.

What is the difference between the conclusions of Theorem 3.1 and Theorem 3.2? It is instructive to look at the following two statements which are equivalent to these two theorems respectively:

- The sequence $\{A_n\}$ is said to converge in probability $P_{\theta_0}$ to $A$ if and only if for each $\epsilon > 0$,

  $$\lim_{n \to \infty} P_{\theta_0}\{|A_n - A| > \epsilon\} = 0,$$

  or equivalently

  $$\lim_{n \to \infty} P_{\theta_0}\{|A_n - A| \leq \epsilon\} = 1. \tag{25}$$

- The sequence $\{A_n\}$ is said to converge to $A$ almost surely (or in probability one, strongly, almost everywhere, *etc.*) if and only if, for each $\epsilon > 0$,

  $$\lim_{n \to \infty} P_{\theta_0}\{\max_{m \geq n} |A_m - A| \leq \epsilon\} = 1. \tag{26}$$

Since (26) clearly implies (25), we have the immediate conclusion that Theorem 3.2 implies Theorem 3.1.

In order to have a better understanding about **convergence almost surely**, it is interesting to quite the following example by Stout (1974, p9):

> *"In statistics there are certain situations where almost sure convergence seems a more relevant concept than convergence in probability. Consider a physician who treats patients with a drug having the same unknown cure probability of p for each patient. The physician is willing to continue*

*use of the drug as long as no superior drug is found. Along with administering the drug, he estimates the cure probability from time to time by dividing the number of cures up to that point in time by the number of patients treated. If n is the number of patients treated, denote this estimating random variable by $\bar{X}_{(n)}$. Suppose the physician wishes to estimate p within a prescribed tolerance $\epsilon > 0$. He asks whether he will ever reach a point in time such that with high probability, all subsequent estimates will fall within $\epsilon$ of p. That is, he wonders for prescribed $\delta > 0$ whether there exisis an integer N such that*

$$P\{\max_{n \geq N} |\bar{X}_{(n)} - p| \leq \epsilon\} \geq 1 - \delta.$$

*The weak law of 'arge numbers says only that*

$$P\{|\bar{X}_{(n)} - p| \leq \epsilon\} \to 1 \quad as \quad n \to \infty$$

*and hence does not answer his question. It is only by the strong law of large numbers that the existence of such an N is indeed guaranteed."*

## 3.3 Convergence in Manifest Probability

Perhaps it may seem confusing to some readers to simultaneously have $\theta$ fixed at $\theta_0$ and have $\theta$ be a random variable governed by $\Pi(\theta)$, as is the case in Theorems 3.1 and 3.2. Thus some sort of clarification seems needed. The idea that leads to the adoption of the notation $\theta_0$ is the following: For any given response vector

$$( X_1, \ldots, X_n) \ = \ ( x_1, \ldots, x_n),$$

if it comes from a randomly selected examinee we can always assume that he or she has specific ability , say $\theta_0$. However, in most cases $\theta_0$ is unknown but hypothetically specified. Under this assumption, the distribution of $X_1, \ldots, X_n$ is induced by $\theta_0$. On the other hand, the given $x_1, \ldots, x_n$ can also be interpreted just as a pattern.

Our interest is to know the proportion of examinees in the population who would produce response vector $x_1, \ldots, x_n$. Denote this proportion number as

$$P\{(X_1, \ldots, X_n) = (x_1, \ldots, x_n)\} \tag{27}$$

and call it the **manifest probability**. It is clearly that

$$P\{(X_1, \ldots, X_n) = (x_1, \ldots, x_n)\} \geq 0$$

and

$$\sum_{x_1, \ldots, x_n} P\{(X_1, \ldots, X_n) = (x_1, \ldots, x_n)\} = 1.$$

Since we know the prior density $\Pi(\theta)$, (27) can be obtained by integrating the joint probability with respect to $\theta$, that is

$$P\{(X_1, \ldots, X_n) = (x_1, \ldots, x_n)\} = \int_\Theta P_n(x_1, \ldots, x_n|\theta)\Pi(\theta)d\theta.$$

According to Theorem 3.1,

$$\int_{\hat{\theta}_n + a\hat{\sigma}_n}^{\hat{\theta}_n + b\hat{\sigma}_n} \Pi_n(\theta| X_1, \ldots, X_n)d\theta \quad \to \quad \Phi(a) - \Phi(b) \tag{28}$$

in probability $P_{\theta_0}$. It is very interesting to notice that the right hand side of (28) is free of $\theta_0$, which suggests that we can further prove that the convergence is **"free of** $\theta_0$**"**. Since (28) holds for **"every"** $\theta_0$, intuitively speaking, it should be true that (28) holds under the **"average of** $\theta_0$**s"**. Therefore, we ought to be able to substitute the manifest probability $P$ for $P_{\theta_0}$:

**Theorem 3.3** *Suppose that conditions (A1) through (A5) hold. Let $\hat{\theta}_n$ be defined by (3) or (4), and $\hat{\sigma}_n$ be the square root of $\{I^{(n)}(\hat{\theta}_n)\}^{-1}$. Then, for $-\infty \leq a < b \leq \infty$, the posterior probability of $\hat{\theta}_n + a\hat{\sigma}_n < \theta < \hat{\theta}_n + b\hat{\sigma}_n$, namely*

$$\int_{\hat{\theta}_n + a\hat{\sigma}_n}^{\hat{\theta}_n + b\hat{\sigma}_n} \Pi_n(\theta| X_1, \ldots, X_n)d\theta,$$

*tends to*

$$(2\pi)^{-1/2} \int_a^b e^{-\frac{1}{2}u^2} du$$

*in manifest probability P.*

Summarizing the last few paragraphs, Theorem 3.1 implies that the asymptotic posterior normality holds for any randomly chosen examinee with ability $\theta_0$. On the other hand, Theorem 3.3 ensures that this asymptotic property holds for any randomly sampled examinee from the population. In other words, one is sampled from the subpopulation and the other is sampled from the whole population. Therefore, Theorem 3.3 has more general meaning. (*The original idea of Theorem 3.3 was proposed by Brian Junker in personal conversation with one of the authors.)*

# 4 Conclusions

The asymptotic posterior normality of latent variable distributions has been established under very general and appropriate hypotheses. This result has (at least) two important implications. First, it provides a probabilistic basis for assessing ability estimation accuracy in the long test case. Second, it provides an important first step in making rigorous the Dutch Identity conjecture (Holland, 1990), which, roughly speaking, claims that only 2 parameters per item are required in order to obtain good long test model fit for unidimensional test data.

Further, the consistency of MLE of $\theta$ has been discussed. It is very interesting to mention that our proof of the consistency of the $\hat{\theta}_n$ is very similar to the Wald's proof(1949) for the $X_1, \ldots, X_n$ i.i.d. case. It is worth remarking that the general *IRT* model (that is, non identically distributed responses) yields as powerful asymptotic results as the *i.i.d.* model – the favorite model of most statisticians, which has so many good qualities.

Finally we should indicate that for general multidimensional IRT models the asymptotic posterior normality can be proved for the random vector $\underline{\theta}$ given test response $X_1, \ldots, X_n$, under suitable regularity conditions.

# Appendix: Proofs of Main Theorems

In this appendix we will prove the results introduced in Section 3.

## A The Proof of Convergence in Probability

The proof of Theorem 3.1 is based on Lemma 3.1, Lemma 3.2, and Corollary 3.1. Before going to the proofs , two important theorems, from real analysis and probability theory respectively, should be introduced here:

**Theorem A.1 (Heine-Borel covering theorem)** *(Billingsley, p566)*
*If* $[a,b] \subset \cap_{k=1}^{\infty}(a_k, b_k)$, *then* $[a,b] \subset \cap_{k=1}^{n}(a_k, b_k)$ *for some* $n$.

**Remark:** *Equivalent to the above theorem is the assertion that a bounded, closed set is compact*[2].

**Theorem A.2 (Strong law of large number** *(Serfling, p27))*
*Let* $X_1, X_2, \ldots$ *be independent with means* $\mu_1$, $\mu_2$, ...*and variances* ${\sigma_1}^2$, ${\sigma_2}^2$ ,.... *If the series* $\sum_{j=1}^{\infty} \sigma_j^2/j^2$ *converges, then*

$$n^{-1}\sum_{j=1}^{n} X_j - n^{-1}\sum_{j=1}^{n}\mu_j \rightarrow 0 \text{ with probability one.}$$

**Proof of Lemma 3.1:**

<u>Remark</u>: *The proof of Lemma 3.1 is an improvement over Walker's result, which only covers the i.i.d. case. The strategy used in the proof can be described by two steps:*

(a) *to prove, for any* $\theta_i \neq \theta_0$, *there exists* $\delta_i > 0$ *such that*

$$\lim_{n\rightarrow\infty} P_{\theta_0}\{\sup_{|\theta-\theta_i|<\delta_i} n^{-1}[L_n(\theta) - L_n(\theta_0)] < -c_i(\delta_i)\} = 1.$$

*We put the subscript i here because we only need finite number of such* $\theta_i$*s.*

[2] A set $C$ is defined to be **compact** if each cover of it by open sets has a finite subcover – that is, if $[G_\theta : \theta \in \Theta]$ covers $C$ and each $G_\theta$ is open, then some finite subcollection $\{G_{\theta_1}, \ldots, G_{\theta_n}\}$covers $C$.

(b) *to use Theorem A.1 to cover* $\{|\theta - \theta_0| \geq \delta\} \cap C$ *, where* $C$ *is a compact set, by a finite number of open sets* $|\theta - \theta_i| < \delta_i$*, i=1,..,m.*

For any $\theta \neq \theta_0$, recalling from (7), the definition of $Z_j(\theta)$, and (9), it follows that

$$n^{-1}[L_n(\theta) - L_n(\theta_0)] = n^{-1}\sum_{j=1}^{n} Z_j(\theta). \tag{29}$$

Now, from (7),

$$E_{\theta_0} Z_j(\theta) = P_j(\theta_0)\log\{\frac{P_j(\theta)}{P_j(\theta_0)}\} \;+\; [1 - P_j(\theta_0)]\log\{\frac{1 - P_j(\theta)}{1 - P_j(\theta_0)}\}. \tag{30}$$

In order to apply Theorem A.2 to $\{Z_j(\theta)\}$, we need to estimate $var(Z_j(\theta))$. Writing $Z_j(\theta)$ using logit function (see (6)),

$$Z_j(\theta) \;=\; X_j[\lambda_j(\theta) - \lambda_j(\theta_0)] + \log\{\frac{1 - P_j(\theta)}{1 - P_j(\theta_0)}\},$$

it follows that

$$\begin{aligned} var(Z_j(\theta)) &= var(X_j)[\lambda_j(\theta) - \lambda_j(\theta_0)]^2 \\ &= P_j(\theta_0)(1 - P_j(\theta_0))[\lambda_j(\theta) - \lambda_j(\theta_0)]^2. \end{aligned}$$

Since, for any fixed $\theta$, $\lambda_j(\theta)$ is bounded in absolute value uniformly in $j$ (assumption (A3)), this implies that there exists a constant $0 < M(\theta) < \infty$ such that

$$|var(Z_j(\theta))| \;\leq\; M(\theta) \;\; for\ all\ j,$$

and hence

$$\sum_{j=1}^{\infty} \frac{var(Z_j(\theta))}{j^2} \;<\; \infty. \tag{31}$$

Thus we can use the law of large numbers to get

$$n^{-1}\sum_{j=1}^{n} Z_j(\theta) - n^{-1}\sum_{j=1}^{n} E_{\theta_0} Z_j(\theta) \;\rightarrow\; 0 \;\; wp1\ . \tag{32}$$

From (29), (32) and assumption (A3) it follows that

$$P\{\overline{\lim_{n\rightarrow\infty}}\, n^{-1}[L_n(\theta) - L_n(\theta_0)] < -c(\theta) < 0\} = 1 \tag{33}$$

for some $c(\theta) > 0$.

Suppose $N_0$ is the closed interval assumed in condition (A2). For any fixed $\theta' \in N_0 \subset \Theta$ and for any $\theta$ satisfying $|\theta - \theta'| \leq \delta$, define $H_j(\theta', \theta)$ by the following:

$$H_j(\theta',\theta) = |\log \frac{P_j(\theta)}{P_j(\theta')}| + |\log \frac{1 - P_j(\theta)}{1 - P_j(\theta')}|.$$

Since $P_j(\theta)$ is strictly increasing in $\theta$, $P_j(\theta') = 1$ and $P_j(\theta') = 0$ can be ruled out. $H_j(\theta', \theta)$, as a continuous function of $\theta$, will achieve a maximum value over $[\theta' - \delta, \ \theta' + \delta]$. Denote this maximum value as $\hat{H}_j(\delta, \theta')$, that is, there exists $\theta^{(\theta', j, \delta)} \in [\theta' - \delta, \ \theta' + \delta]$ such that

$$\hat{H}_j(\delta,\theta') = H_j(\theta^{(\theta',j,\delta)},\theta') = \max_{|\theta-\theta'|\leq\delta}\{H_j(\theta',\theta)\}. \tag{34}$$

Clearly, for each $j$

$$\lim_{\delta \to 0} \hat{H}_j(\delta,\theta') = 0.$$

Now we have

$$|\log\{P_j(\theta)^{X_j}[1 - P_j(\theta)]^{1-X_j}\} - \log\{P_j(\theta')^{X_j}[1 - P_j(\theta')]^{1-X_j}\}|$$

$$\begin{aligned}
&= | X_j \log\{\frac{P_j(\theta)}{P_j(\theta')}\} + (1 - X_j)\log\{\frac{1 - P_j(\theta)}{1 - P_j(\theta')}\}| \\
&< | \log\{\frac{P_j(\theta)}{P_j(\theta')}\}| + |\log\{\frac{1 - P_j(\theta)}{1 - P_j(\theta')}\}| \qquad (35) \\
&= H_j(\theta',\theta) \leq \hat{H}_j(\delta,\theta') \qquad (36)
\end{aligned}$$

We shall now prove that $\{P_j(\theta)\}$ is equicontinuous[3]. From (A2), $P_j'(\theta)$ is continuous and bounded in absolute value uniformly in $j$ and in $\theta \in N_0$. By the mean value theorem,

$$|P_j(\theta) - P_j(\theta')| = |P_j'(\zeta_j)(\theta - \theta')| \leq \zeta_P|\theta - \theta'| \quad for\ all\ j, \tag{37}$$

[3]A function P defined on $(-\infty, \infty)$ is said to be **equicontinuous** if, given $\epsilon > 0$, there exists a number $\delta > 0$ such that $|x' - x''| < \delta$ implies $|P(x') - P(x'')| < \epsilon$ for all $x'$, $x''$.

where $\zeta_j$ is a point between $\theta$ and $\theta'$ for each $j$, and $\zeta_P = \sup_j\{|P_j'(\zeta_j)|\}$ which is finite. Let $\delta = \epsilon/\zeta_P$ for $\epsilon > 0$, then

$$if\ |\theta - \theta'| \ < \ \delta, \ \ |P_j(\theta) - P_j(\theta')| \ < \ \epsilon \ \ for\ all\ j.$$

Recall that $\theta'$ here is any fixed point in $N_0$. Note that

$$\hat{H}_j(\delta,\theta') \ \leq \ \max_{\theta\in[\theta'-\delta,\theta'+\delta]}\{|\log\frac{P_j(\theta)}{P_j(\theta')}|\} \ + \ \max_{\theta\in[\theta'-\delta,\theta'+\delta]}\{|\log\frac{1-P_j(\theta)}{1-P_j(\theta')}|\}.$$

Since $P_j(\theta)$ is strictly increasing in $\theta$ ,

$$\max_{\theta\in[\theta'-\delta,\theta'+\delta]}\{|\log\frac{P_j(\theta)}{P_j(\theta')}|\} \ \leq \ max\{|\log\frac{P_j(\theta'-\delta)}{P_j(\theta')}|, \ |\log\frac{P_j(\theta'+\delta)}{P_j(\theta')}|\}$$

and

$$\max_{\theta\in[\theta'-\delta,\theta'+\delta]}\{|\log\frac{1-P_j(\theta)}{1-P_j(\theta')}|\} \ \leq \ max\{|\log\frac{1-P_j(\theta'-\delta)}{1-P_j(\theta')}|, \ |\log\frac{1-P_j(\theta'+\delta)}{1-P_j(\theta')}|\}.$$

Therefore,

$$\begin{aligned} n^{-1}\sum_{j=1}^{n}\hat{H}_j(\delta,\theta') \ \leq \ & n^{-1}\sum_{j=1}^{n}|\log\frac{P_j(\theta'-\delta)}{P_j(\theta')}| \ + \ n^{-1}\sum_{j=1}^{n}|\log\frac{P_j(\theta'+\delta)}{P_j(\theta')}| \\ & + \ n^{-1}\sum_{j=1}^{n}|\log\frac{1-P_j(\theta'-\delta)}{1-P_j(\theta')}| \ + \ n^{-1}\sum_{j=1}^{n}|\log\frac{1-P_j(\theta'+\delta)}{1-P_j(\theta')}|. \end{aligned}$$

From the equicontinuity of $\{P_j(\theta)\}$, for arbitrary $\epsilon > 0$, there exist a sufficiently small $\delta > 0$ such that

$$|\log\frac{P_j(\theta'+\delta')}{P_j(\theta')}| < \frac{\epsilon}{4} \quad and \quad |\log\frac{1-P_j(\theta'+\delta')}{1-P_j(\theta')}| \ < \ \frac{\epsilon}{4},$$

where either $\delta' = \delta\ or\ -\delta$. Thus, for all $n$ and for all $\delta$ sufficiently small

$$n^{-1}\sum_{j=1}^{n}\hat{H}_j(\delta,\theta') \ \leq \ \epsilon.$$

Therefore

$$\lim_{n\to\infty} n^{-1}\sum_{j=1}^{n}\hat{H}_j(\delta,\theta') \ = \ 0 \quad as\ \ \delta \ \to \ 0. \tag{38}$$

We shall now prove that for any $\theta_i \neq \theta_0$, there exists a sufficiently small $\delta_i > 0$ and sufficiently small $c_i > 0$ such that

$$\lim_{n\to\infty} P\{\sup_{|\theta-\theta_i|<\delta_i} n^{-1}[L_n(\theta) - L_n(\theta_0)] < -c_i\} = 1. \qquad (39)$$

For $\theta \in \{\theta : |\theta - \theta_i| < \delta\}$, according to (29),(7), and (36),

$$\begin{aligned} n^{-1}[L_n(\theta) - L_n(\theta_0)] &= n^{-1}[L_n(\theta_i) - L_n(\theta_0)] + n^{-1}[L_n(\theta) - L_n(\theta_i)] \\ &\leq n^{-1}[L_n(\theta_i) - L_n(\theta_0)] + n^{-1}\sum_{j=1}^{n} \hat{H}_j(\delta, \theta_i). \end{aligned}$$

So we have

$$\sup_{|\theta-\theta_i|<\delta} n^{-1}[L_n(\theta) - L_n(\theta_0)] \leq n^{-1}[L_n(\theta_i) - L_n(\theta_0)] + n^{-1}\sum_{j=1}^{n} \hat{H}_j(\delta, \theta_i).$$

Substituting $\theta_i$ for $\theta$ in (33), we will have

$$P\{\overline{\lim_{n\to\infty}} n^{-1}[L_n(\theta_i) - L_n(\theta_0)] < -c(\theta_i) \equiv -\tilde{c}_i\} = 1, \qquad (40)$$

where $\tilde{c}_i$ is positive for all $i$, and from (38) we will have for all $i$

$$\lim_{n\to\infty} n^{-1}\sum_{j=1}^{n} \hat{H}_j(\delta, \theta_i) \to 0 \quad as\ \delta \to 0.$$

So there is an open interval $|\theta - \theta_i| < \delta_i$ and a positive number $c_i$, e.g. $c_i = \frac{\tilde{c}_i}{2}$, such that (39) holds.

Recall that in assumption (A1) $\Theta$ can be defined by two different domains. In the following, we will discuss these two cases respectively.

**Case 1:** If $\Theta$ is a bounded closed subset of $(-\infty, \infty)$, then $\Theta - \{\theta : |\theta - \theta_0| < \delta\}$ is compact, according to Theorem A.1 it can be covered by finitely many, say $m$, such open intervals

$$(\theta_1 - \delta_1, \theta_1 + \delta_1),\ (\theta_2 - \delta_2, \theta_2 + \delta_2),\ \ldots, (\theta_m - \delta_m, \theta_m + \delta_m).$$

Define event $A_i^{(n)}$ by

$$A_i^{(n)} = \{\sup_{|\theta-\theta_i|<\delta_i} n^{-1}[L_n(\theta) - L_n(\theta_0)] < -c_i\} \tag{41}$$

From $P\{A_i^{(n)}\} \rightarrow 1$ for each $i$ as $n \rightarrow \infty$, we have

$$P\{\cap_{i=1}^{m} A_i^{(n)}\} \rightarrow 1.$$

Now we replace $c_i$ in (39) with

$$k(\delta) = \min\{c_1, c_2, ......, c_m\}.$$

Therefore, (39) holding for all $i$ implies (24).

**Case 2:** If $\Theta$ is not bounded, such as $\Theta = (-\infty, \infty)$, we will show

$$\lim_{n\rightarrow\infty} P\{\sup_{|\theta|>\Delta} n^{-1}[L_n(\theta) - L_n(\theta_0)] < -c_\Delta < 0\} = 1 \tag{42}$$

for a sufficiently large positive number $\Delta$. Now

$$\Theta - \{\theta : |\theta - \theta_0| < \delta\} \cap \{\theta : |\theta| > \Delta\}$$

is bounded compact set, so finally we can get (24) from (42) by defining

$$k(\delta) = \min\{c_1, c_2, ...., c_m, c_\Delta\}.$$

To complete the proof, we have to prove that (42) is correct. Let $|\theta_\Delta| = \Delta$, rewrite

$$\sup_{|\theta|>\Delta} n^{-1}[L_n(\theta) - L_n(\theta_0)] = n^{-1}[L_n(\theta_\Delta) - L_n(\theta_0)] + \sup_{|\theta|>\Delta} n^{-1}[L_n(\theta) - L_n(\theta_\Delta)], \tag{43}$$

where

$$\frac{1}{n}[L_n(\theta) - L_n(\theta_\Delta)] = \frac{1}{n}X_j \sum_{j=1}^{n} \log \frac{P_j(\theta)}{P_j(\theta_\Delta)} + \frac{1}{n}(1 - X_j) \sum_{j=1}^{n} \log \frac{1 - P_j(\theta)}{1 - P_j(\theta_\Delta)}$$

Since $X_j = 0$ or 1, and $P_j(\theta)$ is strictly increasing in $\theta$, then for $\theta > \Delta$,

$$\sup_{|\theta|>\Delta} n^{-1}[L_n(\theta) - L_n(\theta_\Delta)] \leq \sup_{\theta>\Delta} n^{-1} \sum_{j=1}^{n} \log \frac{P_j(\theta)}{P_j(\Delta)},$$

and for $\theta < -\Delta$,

$$\sup_{|\theta|>\Delta} n^{-1}[L_n(\theta) - L_n(\theta_\Delta)] \leq \sup_{\theta<-\Delta} n^{-1}\sum_{j=1}^{n} \log \frac{1-P_j(\theta)}{1-P_j(-\Delta)}.$$

Since each item response function has horizontal asymptotes as $\theta \to +\infty$ and $\theta \to -\infty$, we can prove that

$$\lim_{n\to\infty} \sup_{\theta>\Delta} n^{-1}\sum_{j=1}^{n} \log \frac{P_j(\theta)}{P_j(\Delta)} \to 0$$

and

$$\lim_{n\to\infty} \sup_{\theta<-\Delta} n^{-1}\sum_{j=1}^{n} \log \frac{1-P_j(\theta)}{1-P_j(-\Delta)} \to 0$$

as $\Delta \to \infty$. Therefore we have

$$\lim_{n\to\infty} \sup_{|\theta|>\Delta} n^{-1}[L_n(\theta) - L_n(\theta_\Delta)] \to 0 \quad as\ \Delta \to \infty. \tag{44}$$

Substituting $\theta_\Delta$ for $\theta$ in (33), we have

$$P\{\overline{\lim_{n\to\infty}} n^{-1}[L_n(\theta_\Delta) - L_n(\theta_0)] < -c_\Delta\} = 1. \tag{45}$$

Formulas (44) and (45) can be used to (43) to get (42). Therefore (42) holds. ■

**Proof of Corollary 3.1:** The $MLE$, if it exists, obviously satisfies

$$L_n(\hat{\theta}_n) - L_n(\theta_0) = \log\{\frac{P_n(X_1,\ldots,X_n|\hat{\theta}_n)}{P_n(X_1,\ldots,X_n|\theta_0)}\} \geq 0 \tag{46}$$

for all n and for all $X_1,\ldots,X_n$. It is sufficient to prove that for any $\epsilon > 0$ and $\delta > 0$, there exists $N(\epsilon,\delta)$ such that

$$Prob\{|\hat{\theta}_n - \theta_0| < \delta\} > 1-\epsilon \quad for\ all\ n > N(\epsilon,\delta).$$

Suppose $\hat{\theta}_n$ is not consistent, then there exist $\epsilon_0$ and $\delta_0$ such that, for any $N$ there exists some $n > N$,

$$Prob\{|\hat{\theta}_n - \theta_0| > \delta_0\} > \epsilon_0.$$

Therefore we can obtain a subsequence $\{\theta_{n_i}\}$ such that

$$Prob\{|\theta_{n_i} - \theta_0| > \delta_0\} > \epsilon_0 \quad for\ all\ n_i. \tag{47}$$

Thus,

$$\epsilon_0 \leq \overline{\lim_{n\to\infty}} Prob\{|\hat{\theta}_n - \theta_0| > \delta_0\} \leq Prob\{\overline{\lim_{n\to\infty}}[|\hat{\theta}_n - \theta_0| > \delta_0]\}.$$

It is obvious that the event

$$\overline{\lim_{n\to\infty}}[|\hat{\theta}_n - \theta_0| > \delta_0]$$

implies that for infinitely many $n$

$$\sup_{|\theta-\theta_0|\geq\delta_0} [L_n(\theta) - L_n(\hat{\theta}_n)] \geq 0 \quad for\ infinitely\ many\ n,$$

because $\theta = \hat{\theta}_n$ is a possible value. But then according to (46) the event

$$\sup_{|\theta-\theta_0|\geq\delta_0} [L_n(\theta) - L_n(\theta_0)] \geq 0 \quad for\ infinitely\ many\ n$$

has a probability greater than or equal to $\epsilon_0$. This contradicts (24), which implies that for any $\epsilon > 0$, there exists $N$ such that

$$Prob\{\sup_{|\theta-\theta_0|\geq\delta_0} [L_n(\theta) - L_n(\theta_0)] \geq 0\} < \epsilon \quad for\ all\ n > N.$$

This completes the proof. ■

**Proof of Lemma 3.2:** Without loss of generality, we first consider that $\hat{\theta}_n \in [|\theta - \theta_0| < \delta] \subset N_0$. Since the $\hat{\theta}_n$ is consistent, the probability of $\hat{\theta}_n$ being contained in the neighborhood of $\theta_0$ will be close to one, when $n$ is sufficiently large.

The second derivative of the log likelihood function can be written as

$$L_n''(\theta) = \sum_{j=1}^{n} \lambda_j''(\theta)[X_j - P_j(\theta)] - \sum_{j=1}^{n} I_j(\theta). \tag{48}$$

To prove (48), first notice that it suffices to prove for $n$=1, that is

$$L_1''(\theta) = \lambda_1''(\theta)[X1 - P_1(\theta)] - I_1(\theta). \tag{49}$$

Note that

$$L_1(\theta) = \lambda'(\theta)X_1 + \log(1 - P_1(\theta)),$$

so that

$$L_1''(\theta) = \lambda_1''(\theta)X_1 + [\log(1 - P_1(\theta))]''.$$

Comparing this with (49) it remains to show that

$$-[\log(1 - P_1(\theta))]'' = \lambda_1''(\theta)P_1(\theta) + I_1(\theta). \tag{50}$$

However by definition,

$$I_1(\theta) = E_{\theta_0}[-L_1''(\theta)] = -\lambda_1''(\theta)P_1(\theta) - [\log(1 - P_1(\theta))]'',$$

which is equivalent to (50).

Consider the numerator of $|R_n|$ :

$$\begin{aligned}
|L_n''(\theta_n^*) + I^{(n)}(\hat{\theta}_n)| &= |\sum_{j=1}^{n}[\lambda_j''(\theta_n^*) - \lambda_j''(\theta_0)][X_j - P_j(\theta_n^*)] + \sum_{j=1}^{n}\lambda_j''(\theta_0)[X_j - P_j(\theta_0)] \\
&+ \sum_{j=1}^{n}\lambda_j''(\theta_0)[P_j(\theta_0) - P_j(\theta_n^*)] + \sum_{j=1}^{n}\{I_j(\hat{\theta}_n) - I_j(\theta_n^*)\}| \\
&\leq \sum_{j=1}^{n}|\lambda_j''(\theta_n^*) - \lambda_j''(\theta_0)| \qquad (51) \\
&+ |\sum_{j=1}^{n}\lambda_j''(\theta_0)[X_j - P_j(\theta_0)]| \\
&+ |\sum_{j=1}^{n}\lambda_j''(\theta_0)[P_j(\theta_0) - P_j(\theta_n^*)]| \\
&+ \sum_{j=1}^{n}|I_j(\hat{\theta}_n) - I_j(\theta_n^*)|.
\end{aligned}$$

Note that $\theta_n^*$ depends on $\theta$ and $\hat{\theta}_n$ through the Taylor expansion and that the distribution of $\hat{\theta}_n$ depends on $\theta_0$. From (37)

$$|\sum_{j=1}^{n} \lambda_j''(\theta_0)[P_j(\theta_0) - P_j(\theta_n^*)]| \leq |\theta_n^* - \theta_0| n\zeta_P. \tag{52}$$

From the mean value theorem

$$|\lambda_j''(\theta_n^*) - \lambda_j''(\theta_0)| = |\lambda_j'''(\hat{\theta}_n^{(\lambda,j)})(\theta_n^* - \theta_0)|$$

and

$$|I_j(\hat{\theta}_n) - I_j(\theta_n^*)| = |I_j'(\hat{\theta}_n^{(I,j)})(\hat{\theta}_n - \theta_n^*)|,$$

where $\hat{\theta}_n^{(\lambda,j)}$ is a point between $\theta_n^*$ and $\theta_0$, and $\hat{\theta}_n^{(I,j)}$ is a point between $\hat{\theta}_n$ and $\theta_n^*$. According to assumption (A4), the third derivative of the logit function, $\lambda_j'''(\theta)$, and the first derivative of the information function, $I_j'(\theta)$, are bounded in absolute value uniformly in $j$ and in $\theta$, therefore,

$$\sum_{j=1}^{n} |\lambda_j''(\theta_n^*) - \lambda_j''(\theta_0)| \leq |\theta_n^* - \theta_0| n\zeta_\lambda, \tag{53}$$

and

$$\sum_{j=1}^{n} |I_j(\hat{\theta}_n) - I_j(\theta_n^*)| \leq |\hat{\theta}_n - \theta_n^*| n\zeta_I. \tag{54}$$

Note that $\zeta_P$, $\zeta_\lambda$, and $\zeta_I$ are finite positive numbers and they are independent of $j$. We shall now prove

$$|\sum_{j=1}^{n} \lambda_j''(\theta_0)[X_j - P_j(\theta_0)]| = O_p(n^{1/2}). \tag{55}$$

(See Footnote [4].) Assumption (A4) ensures that $\{\lambda_j''(\theta_0)\}$ is bounded in absolute value uniformly in $j$. By Chebyshev's inequality, for some $M > 0$,

$$P\{|\sum_{j=1}^{n} \lambda_j''(\theta_0)[X_j - P_j(\theta_0)]| > n^{1/2}K\} < \frac{\sum_{j=1}^{n}[\lambda_j''(\theta_0)]^2 P_j(\theta_0)(1 - P_j(\theta_0))}{nK^2} < MK^{-2},$$

[4] The notation of $a_n = O_p(b_n)$ means that $a_n$ is bounded stochasticly by $b_n$ in probability, that is, $a_n = O_p(b_n)$ if and only if for arbitrary $\epsilon > 0$ there exist $M_\epsilon$ and $N_\epsilon$ such that

$$P\{|a_n/b_n| < M_\epsilon\} > 1 - \epsilon \quad \textit{for all } n > N_\epsilon.$$

that is, for arbitrary $\epsilon > 0$, take $K = (M/\epsilon)^{1/2}$, then we have

$$P\{|\sum_{j=1}^{n} \lambda_j''(\theta_0)[X_j - P_j(\theta_0)]/n^{1/2}| < K\} > 1 - \epsilon \quad for\ all\ n$$

that means we have (55).

Formulas (52), (53), (54), and (55) can be applied to (51) to get

$$|L_n''(\theta_n^*) + I^{(n)}(\hat{\theta}_n)| \leq \{|\theta_n^* - \theta_0| + |\hat{\theta}_n - \theta_n^*|\}nC + O_p(n^{1/2}), \tag{56}$$

where

$$C = \zeta_P + \zeta_\lambda + \zeta_I.$$

We shall now prove

$$\lim_{n\to\infty} P\{\ I^{(n)}(\hat{\theta}_n)/n\ \geq c/2 > 0\} = 1. \tag{57}$$

By assumption (A4)

$$\begin{aligned} n^{-1}|I^{(n)}(\hat{\theta}_n) - I^{(n)}(\theta_0)| &\leq n^{-1}\sum_{j=1}^{n}|I_j(\hat{\theta}_n) - I_j(\theta_0)| \\ &\leq |\hat{\theta}_n - \theta_0|\zeta_I. \end{aligned} \tag{58}$$

By using the consistency of $\hat{\theta}_n$ and (58), we get

$$I^{(n)}(\hat{\theta}_n)/n - I^{(n)}(\theta_0)/n \to 0 \quad in\ P_{\theta_0}\ as\ n \to \infty.$$

Thus, by assumption (A5), we have (57).

From (56) and (57) we obtain

$$\begin{aligned} \sup_{|\theta-\theta_0|<\delta} |R_n(\theta,\ X_1,\ldots,X_n)| &\leq \sup_{|\theta-\theta_0|<\delta}\left\{\frac{(|\theta_n^* - \theta_0| + |\hat{\theta}_n - \theta_n^*|)nC}{I^{(n)}(\hat{\theta}_n)}\right\} + O_p\left\{\frac{n^{1/2}}{I^{(n)}(\hat{\theta}_n)}\right\} \\ &= \sup_{|\theta-\theta_0|<\delta}\left\{\frac{(|\theta_n^* - \theta_0| + |\hat{\theta}_n - \theta_n^*|)nC}{I^{(n)}(\hat{\theta}_n)}\right\} + O_p(n^{-1/2}). \end{aligned}$$

Note that

$$|\theta_n^* - \hat{\theta}_n| \leq |\theta_n^* - \theta_0| + |\hat{\theta}_n - \theta_0| \quad and \quad |\theta_n^* - \theta_0| \leq |\theta - \theta_0| + |\hat{\theta}_n - \theta_0|,$$

where the second inequality follows from the fact that $\theta_n^*$ is between $\theta$ and $\hat{\theta}_n$. Therefore

$$\sup_{|\theta-\theta_0|<\delta} |R_n(\theta,\ X_1,\ldots,X_n)| \leq \sup_{|\theta-\theta_0|<\delta} \left\{ \frac{(3|\hat{\theta}_n - \theta_0| + 2|\theta - \theta_0|)C}{\frac{I^{(r)}(\hat{\theta}_n)}{n}} \right\} + O_p(n^{-1/2}).$$

For any $\epsilon > 0$, choose

$$\delta = \frac{\epsilon}{3}\left(\frac{C}{c/2}\right)^{-1},$$

then we have (23), recalling that $\hat{\theta}_n \rightarrow \theta_0$ in $P_{\theta_0}$ and (36).

The above proof is based on the assumption that $\hat{\theta}_n$ is in the neighborhood $(\theta_0 - \delta, \theta_0 + \delta)$, so we just proved that the conditional probability approaches to one:

$$\lim_{n\rightarrow\infty} P[U_n|V_n] = 1, \tag{59}$$

where

$$U_n \equiv \{ \sup_{|\theta-\theta_0|<\delta} |R_n(\theta,\ X_1,\ldots,X_n)| < \epsilon\}$$

and

$$V_n \equiv \{\hat{\theta}_n \in [|\theta - \theta_0| < \delta] \subset N_0\}.$$

Since Corollary 3.1 implies

$$\lim_{n\rightarrow\infty} P[V_n] = 1, \tag{60}$$

it is obvious that (59) and (60) implies $\lim_{n\rightarrow\infty} P[U_n] = 1$. Thus we finish the proof.

■

**Proof of Theorem 3.1:**

**Remark:** *The following proof will use a similar methodology as Walker's(1969). The proof itself will not use any assumption about i.i.d.. Instead, it will just depend on* *the results of Lemma 3.1 and Lemma 3.2.*

As we discussed in section 3.1, it suffices to prove (13) and (14). To prove (13) it suffices to prove (20) and (21). Let us start with (20). Rewrite $G_1$ as

$$G_1 = P_n( X_1,\ldots,X_n|\hat{\theta}_n)\int_{|\theta-\theta_0|\geq\delta}\Pi(\theta)exp\{L_n(\theta)-L_n(\hat{\theta}_n)\}d\theta$$

$$= P_n( X_1,\ldots,X_n|\hat{\theta}_n)exp\{L_n(\theta_0)-L_n(\hat{\theta}_n)\}\int_{|\theta-\theta_0|\geq\delta}\Pi(\theta)exp\{L_n(\theta)-L_n(\theta_0)\}d\theta.$$

Since $\hat{\theta}_n$ is an MLE,

$$L_n(\theta_0)-L_n(\hat{\theta}_n)\leq 0, \tag{61}$$

and therefore $\exp\{L_n(\theta_0)-L_n(\hat{\theta}_n)\} \leq 1$. So we have

$$\frac{G_1}{P_n( X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} = \{I^{(n)}(\hat{\theta}_n)\}^{1/2}\int_{|\theta-\theta_0|\geq\delta}\Pi(\theta)\exp\{L_n(\theta)-L_n(\hat{\theta}_n)\}d\theta$$

$$= \exp\{L_n(\theta_0)-L_n(\hat{\theta}_n)\}\{I^{(n)}(\hat{\theta}_n)\}^{1/2}\int_{|\theta-\theta_0|\geq\delta}\Pi(\theta)\exp\{L_n(\theta)-L_n(\theta_0)\}d\theta$$

$$\leq \{I^{(n)}(\hat{\theta}_n)\}^{1/2}G_0, \tag{62}$$

where

$$G_0 = \int_{|\theta-\theta_0|\geq\delta}\Pi(\theta)\exp\{L_n(\theta)-L_n(\theta_0)\}d\theta.$$

By Lemma 3.1, for any $\delta > 0$, there exists $k(\delta) > 0$ such that

$$\lim_{n\to\infty} P_{\theta_0}\{U_n\} = 1,$$

where

$$U_n = [\sup_{|\theta-\theta_0|\geq\delta} n^{-1}[L_n(\theta)-L_n(\theta_0)] < -k(\delta) < 0]. \tag{63}$$

Define

$$V_n = [G_0 \leq \exp\{-nk(\delta)\}]; \tag{64}$$

notice that

$$\exp\{-nk(\delta)\}\int_{|\theta-\theta_0|\geq\delta}\Pi(\theta)d\theta \leq \exp\{-nk(\delta)\}.$$

Because $U_n \subseteq V_n$, we have

$$\lim_{n\to\infty} P_{\theta_0}\{G_0 \le \exp\{-nK(\delta)\}\} = 1.$$

Since

$$\{I^{(n)}(\hat{\theta}_n)\}^{1/2}\exp\{-nk(\delta)\} \;\to\; 0 \;\; in \; P_{\theta_0}, \;\; as \; n \;\to 0,$$

it follows, (using (62))

$$\lim_{n\to\infty} \frac{G_1}{P_n(\ X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} = 0 \qquad in \; P_{\theta_0}. \tag{65}$$

Thus (20) holds.

**Now we prove (21).** From (15), rewrite $G_2$ as

$$\begin{aligned} G_2 &= P_n(\ X_1,\ldots,X_n|\hat{\theta}_n)\int_{|\theta-\theta_0|<\delta}\Pi(\theta)\exp\{L_n(\theta)-L_n(\hat{\theta}_n)\}d\theta \\ &= P_n(\ X_1,\ldots,X_n|\hat{\theta}_n)\int_{|\theta-\theta_0|<\delta}\Pi(\theta)\exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1-R_n)\}d\theta \\ &= P_n(\ X_1,\ldots,X_n|\hat{\theta}_n)\Pi(\theta_0)\int_{|\theta-\theta_0|<\delta}\frac{\Pi(\theta)}{\Pi(\theta_0)}\exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1-R_n)\}d\theta. \end{aligned}$$

We shall now observe $\frac{G_2}{P_n(\ X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n}$.

$$\frac{G_2}{P_n(\ X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} = \frac{\Pi(\theta_0)}{\hat{\sigma}_n}\int_{|\theta-\theta_0|<\delta}\frac{\Pi(\theta)}{\Pi(\theta_0)}\exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1-R_n)\}d\theta \tag{66}$$

From condition (A1), in particular the continuouity of $\Pi(\theta)$, for any $\epsilon > 0$ we can choose $\delta$ such that $\{\theta : |\theta-\theta_0| < \delta\} \subset N_0$ and

$$1-\epsilon \le \inf_{|\theta-\theta_0|<\delta}\frac{\Pi(\theta)}{\Pi(\theta_0)} \le \sup_{|\theta-\theta_0|<\delta}\frac{\Pi(\theta)}{\Pi(\theta_0)} \le 1+\epsilon. \tag{67}$$

Then, using (66)

$$\frac{(1-\epsilon)\Pi(\theta_0)}{\hat{\sigma}_n}G_3 \le \frac{G_2}{P_n(\ X_1,\ldots,X_n|\hat{\theta}_n)\hat{\sigma}_n} \le \frac{(1+\epsilon)\Pi(\theta_0)}{\hat{\sigma}_n}G_3, \tag{68}$$

where

$$G_3 = \int_{|\theta-\theta_0|<\delta} exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1-R_n)\}d\theta. \tag{69}$$

For any $\epsilon > 0$, define

$$C_n = [\sup_{|\theta-\theta_0|<\delta} |R_n(\theta,\ X_1,\ldots,X_n)| < \epsilon], \tag{70}$$

and

$$D_n = [\int_{|\theta-\theta_0|<\delta} \exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1+\epsilon)\}d\theta \le G_3 \le \int_{|\theta-\theta_0|<\delta} \exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1+\epsilon)\}d\theta] \tag{71}$$

Now we should get rid of $R_n$. Since $C_n \subseteq D_n$, and for any $\epsilon > 0$, from Lemma 3.2,

$$\lim_{n\to\infty} P_{\theta_0}\{C_n\} = 1, \qquad \textit{this implies} \quad \lim_{n\to\infty} P_{\theta_0}\{D_n\} = 1.$$

That is, the probability of the event

$$\int_{|\theta-\theta_0|<\delta} exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1+\epsilon)\}d\theta \le G_3 \le \int_{|\theta-\theta_0|<\delta} exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1-\epsilon)\}d\theta \tag{72}$$

converges to 1 as $n \to \infty$. Therefore, recalling (17),(65),(68), and (69), the only thing left to establish (13) is to observe that

$$\int_{|\theta-\theta_0|<\delta} exp\{-\frac{(\theta-\hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1+\epsilon^*)\}d\theta$$

$$= (2\pi)^{1/2}(1+\epsilon^*)^{-1/2}\hat{\sigma}_n[\Phi\{\hat{\sigma}_n^{-1}(\theta_0+\delta-\hat{\theta}_n)(1+\epsilon^*)^{1/2}\}-\Phi\{\hat{\sigma}_n^{-1}(\theta_0-\delta-\hat{\theta}_n)(1+\epsilon^*)^{1/2}\}], \tag{73}$$

where $\epsilon^* = \epsilon \ or \ -\epsilon$. Since $\hat{\theta}_n$ is consistent and $\hat{\sigma}_n^{-1} \to \infty$ in probability, when $\epsilon < 1$,

$$\theta_0 + \delta - \hat{\theta}_n \to \delta \quad in\ P_{\theta_0},$$

$$\theta_0 - \delta - \hat{\theta}_n \to -\delta \quad in\ P_{\theta_0},$$

$$\hat{\sigma}_n^{-1}(\theta_0+\delta-\hat{\theta}_n)(1+\epsilon^*)^{1/2} \to \infty \quad in\ P_{\theta_0},$$

$$\hat{\sigma}_n^{-1}(\theta_0-\delta-\hat{\theta}_n)(1+\epsilon^*)^{1/2} \to -\infty \quad in\ P_{\theta_0}.$$

So

$$\Phi\{\hat{\sigma}_n^{-1}(\theta_0 + \delta - \hat{\theta}_n)(1+\epsilon^*)^{1/2}\} \rightarrow 1 \quad in\ P_{\theta_0},$$

$$\Phi\{\hat{\sigma}_n^{-1}(\theta_0 - \delta - \hat{\theta}_n)(1+\epsilon^*)^{1/2}\} \rightarrow 0 \quad in\ P_{\theta_0}.$$

Therefore, the difference in the square brackets of (73) converges to unity in probability. Since the $\epsilon$ is arbitrary, this proves (13).

**Now we prove (14).** First of all we consider (12) and (17) again: $G$ and $G_2$ are the same except for their rigions of integration: one is $(\hat{\theta}_n + a\hat{\sigma}_n,\ \hat{\theta}_n + b\hat{\sigma}_n)$ and the other is$\{\theta : |\theta - \theta_0| < \delta\}$. For the same $\epsilon$ and $\delta$ given by (67), if $(\hat{\theta}_n + a\hat{\sigma}_n,\ \hat{\theta}_n + b\hat{\sigma}_n)$ is a subset of $\{\theta : |\theta - \theta_0| < \delta\}$, we must have

$$1 - \epsilon \leq \inf_{(\hat{\theta}_n + a\hat{\sigma}_n,\ \hat{\theta}_n + b\hat{\sigma}_n)} \frac{\Pi(\theta)}{\Pi(\theta_0)} \leq \sup_{(\hat{\theta}_n + a\hat{\sigma}_n,\ \hat{\theta}_n + b\hat{\sigma}_n)} \frac{\Pi(\theta)}{\Pi(\theta_0)} \leq 1 + \epsilon. \tag{74}$$

Define

$$E_n \equiv [(\hat{\theta}_n + a\hat{\sigma}_n,\ \hat{\theta}_n + b\hat{\sigma}_n) \subseteq \{\theta : |\theta - \theta_0| < \delta\}].$$

Since $\hat{\theta}_n \rightarrow \theta_0$ $in$ $P_{\theta_0}$ and $\hat{\sigma}_n \rightarrow 0$ $in$ $P_{\theta_0}$. Thus,

$$P_{\theta_0}(E_n) \rightarrow 1 \quad as\ n \rightarrow \infty, \tag{75}$$

and hence the probability of (74) converges to 1 as $n \rightarrow \infty$. Consider (68) again. If $(\hat{\theta}_n + a\hat{\sigma}_n,\ \hat{\theta}_n + b\hat{\sigma}_n)$ is a subset of $\{\theta : |\theta - \theta_0| < \delta\}$, and if we substitute the rigions of integration of (68) by $(\hat{\theta}_n + a\hat{\sigma}_n,\ \hat{\theta}_n + b\hat{\sigma}_n)$, then the new inequality (76) below will still hold.

$$\frac{(1-\epsilon)\Pi(\theta_0)}{\hat{\sigma}_n} G_3' \leq \frac{G}{P_n(X_1, \ldots, X_n|\hat{\theta}_n)\hat{\sigma}_n} \leq \frac{(1+\epsilon)\Pi(\theta_0)}{\hat{\sigma}_n} G_3', \tag{76}$$

where

$$G_3' = \int_{\hat{\theta}_n + b\hat{\sigma}_n}^{\hat{\theta}_n + a\hat{\sigma}_n} \exp\{-\frac{(\theta - \hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1 - R_n)\} d\theta. \tag{77}$$

Because of (75), the probability of the event indicated by (76) converges to 1 as $n \to \infty$. For the same $\epsilon$ given by (72) define

$$C_n' = [\sup_{(\hat{\theta}_n + a\hat{\sigma}_n,\ \hat{\theta}_n + b\hat{\sigma}_n)} |R_n(\theta,\ X_1, \ldots, X_n)| < \epsilon], \tag{78}$$

and

$$D_n' = \left[\int_{\hat{\theta}_n + b\hat{\sigma}_n}^{\hat{\theta}_n + a\hat{\sigma}_n} \exp\{-\frac{(\theta - \hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1+\epsilon)\} d\theta \le G_3' \le \int_{\hat{\theta}_n + \hat{\sigma}_n b}^{\hat{\theta}_n + \hat{\sigma}_n a} \exp\{-\frac{(\theta - \hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1-\epsilon)\} d\theta\right]. \tag{79}$$

From (75) and $E_n \subseteq C_n' \subseteq D_n'$,

$$P_{\theta_0}\{D_n'\} \to 1 \;\; as \;\; n \to \infty.$$

Similar to (73), now we shall estimate

$$\int_{\hat{\theta}_n + b\hat{\sigma}_n}^{\hat{\theta}_n + a\hat{\sigma}_n} \exp\{-\frac{(\theta - \hat{\theta}_n)^2}{2\hat{\sigma}_n^2}(1+\epsilon^*)\} d\theta, \tag{80}$$

where $\epsilon^* = \epsilon \; or \; -\epsilon$. It is obvious that the quantity in (80) is equal to

$$(2\pi)^{1/2}\hat{\sigma}_n(1+\epsilon^*)^{-1/2}[\Phi\{a(1+\epsilon^*)^{1/2}\} - \Phi\{b(1+\epsilon^*)^{1/2}\}].$$

Since we can make $\epsilon$ arbitrarily small, therefore, using (76) and (77) we can finally obtain

$$\frac{G}{P_n(\,X_1, \ldots, X_n|\hat{\theta}_n)\hat{\sigma}_n} \to (2\pi)^{1/2}\Pi(\theta_0)\{\Phi(a) - \Phi(b)\}$$

in probability $P_{\theta_0}$. ■

# B The Proof of Strong Convergence

The proof of Theorem 3.2 is analogous to that of Theorem 3.1 and is also based on two lemmas and one corollary. However, these intermediate results are stronger than those used in proving Theorem 3.1.

**Lemma B.1** *Under the assumptions of Lemma 3.1, for any given $\delta > 0$, there exists $k(\delta) > 0$ such that*

$$P_{\theta_0}\{\overline{\lim_{n\to\infty}} \sup_{|\theta-\theta_0|\geq\delta} n^{-1}[L_n(\theta) - L_n(\theta_0)] < -k(\delta)\} = 1. \tag{81}$$

**Proof:** The proof of (81) analogous to that of Lemma 3.1 except the following two changes:

**(1)** replacing (39) by

$$P_{\theta_0}\{\overline{\lim_{n\to\infty}} \sup_{|\theta-\theta_i|<\delta} n^{-1}[L_n(\theta) - L_n(\theta_0)] < -c_i\} = 1; \tag{82}$$

**(2)** replacing (41) by

$$A_i^{(n)} = \{\overline{\lim_{n\to\infty}} \sup_{|\theta-\theta_i|<\delta} n^{-1}[L_n(\theta) - L_n(\theta_0)] < -c_i\}.$$

Now we only need to prove (82). Since

$$\overline{\lim_{n\to\infty}} n^{-1}[L_n(\theta_i) - L_n(\vartheta_0)]$$

is measureable with respect to the tail $\sigma$ field

$$\sigma(Z_n(\theta_i), Z_{n+1}(\theta_i), \ldots.),$$

by the **Kolmogorov's $0-1$ law** (Billingsley, p295) it must be a "nonrandom" constant with probability 1. Denote this constant as $\eta$. According to (40),

$$P_{\theta_0}\{\eta = \overline{\lim_{n\to\infty}} n^{-1}[L_n(\theta_i) - L_n(\theta_0)] \leq -c(\theta_i) < 0\} = 1.$$

Choose

$$\epsilon = \frac{c(\theta_i) - \eta}{2}$$

and choose $\delta$ small enough such that

$$\overline{\lim_{n\to\infty}} n^{-1} \sum_{j=1}^{n} \hat{H}_j(\delta, \theta_i) < \epsilon,$$

(see (34) for the definition of $\hat{H}_j(\delta,\theta_i)$), thus

$$\begin{aligned}\overline{\lim_{n\to\infty}} \sup_{|\theta-\theta_i|<\delta} n^{-1}[L_n(\theta)-L_n(\theta_0)] &\leq \overline{\lim_{n\to\infty}} n^{-1}[L_n(\theta_i)-L_n(\theta_0)] + \overline{\lim_{n\to\infty}} n^{-1}\sum_{j=1}^{n}\hat{H}_j(\delta,\theta_i) \\ &\leq \eta+\epsilon < -c(\theta_i) \quad \textit{almost surely.}\end{aligned}$$

Thus (82) holds. ■

**Corollary B.1** *Lemma B.1 ensures that*

$$P_{\theta_0}\{\lim_{n\to\infty}\hat{\theta}_n = \theta_0\} = 1.$$

**Proof:** Analogous to that of Wald (1949) and omitted. ■

**Lemma B.2** *Under the assumptions of Lemma 3.2, for any $\epsilon > 0$, there exists $\delta$ such that*

$$P_{\theta_0}\{\overline{\lim_{n\to\infty}} \sup_{|\theta-\theta_0|<\delta} |R_n(X_1,\ldots,X_n,\theta)| < \epsilon\} = 1. \tag{83}$$

**Proof:** Analogous to that of Lemma 3.2 and omitted. ■

**Proof of Theorem 3.2:** Based on Lemma B.1, Lemma B.2 and Corrollary B.1. The basic steps are analogous to those of Theorem 3.1 and omitted. ■

# C The Proof of Convergence in Manifest Probability

**Proof of Theorem 3.3:** Theorem 3.1 implies that for arbitrary $\theta$ and arbitrary $\epsilon > 0$,

$$P_\theta\{|A_n(X_1,\ldots,X_n) - A| \geq \epsilon\} \to 0,$$

as $n \to \infty$. Define

$$H_n(\theta, \epsilon) = P_\theta\{| A_n(X_1, \ldots, X_n) - A| \geq \epsilon\}$$

It is clear that for any $\theta$ and $\epsilon > 0$ that

$$0 \leq H_n(\theta, \epsilon) \leq 1 \quad \text{and} \quad \lim_{n \to \infty} H_n(\theta, \epsilon) = 0.$$

By Lebesgue's bounded convergence theorem (Billingsley, p214),

$$\int_\Theta H_n(\theta, \epsilon)\Pi(\theta)d\theta \to 0.$$

That is,

$$\begin{aligned} P\{|A_n(X_1, \ldots, X_n) - A| \geq \epsilon\} &= \int_\Theta P\{|A_n(X_1, \ldots, X_n) - A| \geq \epsilon|\theta\}\Pi(\theta)d\theta \\ &= \int_\Theta H_n(\theta, \epsilon)\Pi(\theta)d\theta \to 0. \end{aligned}$$

This proves Theorem 3.3. ■

# Acknowledgement.

The authors wish to thank Kumar Joag-Dev, Brian Junker, Bert Green, Paul Holland, Robert Mislevy, and especially Zhiliang Ying for their useful comments and discussions.

**References**

**Bishop, A., Fienberg, S., & Holland, P.** (1975). *Discrete multivariate analysis: Theory and practice.* Cambridge, MA: MIT Press.

**Billingsley, P.** (1986). *Probability and measure.* New York: John Wiley & Sons.

**Bock, R.D., & Mislevy R.J.** (1982). Adaptive EAP estimation of ability in a microcomputer environment. *Applied Psychological Measurement, 6, 431-444.*

**Chang, H., & Stout, W.F.** (1990, June). The asymptotic posterior normality on IRT model. Paper presented at 1990 ONR Contractors' Meeting on Model-Based Measurement, Portland, Oregon.

**Drasgow, F.** (1987). A study of measurement bias of two standard psychological tests. *Journal of Applied Psychology, 72*, 19-30.

**Holland, P.W.** (1990). The Dutch identity: a new tool for the study of item response theory models. *Psychometrika, 55*, 5-18.

**Junker, B.W.** (1988). *Statistical aspects of a new latent trait model*, Ph.D. dissertation, Department of Statistics, University of Illinois at Urbana-Champaign.

**Lehmann, E.L.** (1983). *Theory of point estimation.* New York: John Wiley & Sons.

**Lindley, D.V.** (1965). *Introduction to probability and statistics. part 2: Inference.* London: Cambridge University Press.

**Lord, F.M.** (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Lawrence Erlbaum.

**Serfling, R.J.** (1980). *Approximation theorems in mathematical statistics.* New York: John Wiley & Sons.

**Stout, W.F.** (1974). *Almost sure convergence.* New York: Academic Press.

**Stout, W.F.** (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika, 55*, 293-325.

**Wald, A.** (1949). Note on the consistency of the maximum likelihood estimate. *Ann.Math.Statist.*, *20*, 595-601.

**Walker, A.M.** (1969). On the asymptotic behaviour of posterior distributions. *J.R.Statist.Soc.Ser. B,31*, 80-88.

**Wolfowtiz, J.** (1949). On Wald's proof of the consistency of the maximum likelihood estimate. *Ann. Math. Statist., 20*, 602-603.

Distribution List

Dr. Terry Ackerman
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. James Algina
1403 Norman Hall
University of Florida
Gainesville, FL 32605

Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

Dr. Ronald Armstrong
Rutgers University
Graduate School of Management
Newark, NJ 07102

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Laura L. Barnes
College of Education
University of Toledo
2801 W. Bancroft Street
Toledo, OH 43606

Dr. William M. Bart
University of Minnesota
Dept. of Educ. Psychology
330 Burton Hall
178 Pillsbury Dr., S.E.
Minneapolis, MN 55455

Dr. Isaac Bejar
Law School Admissions
Services
P.O. Box 40
Newtown, PA 18940-0040

Dr. Ira Bernstein
Department of Psychology
University of Texas
P.O. Box 19528
Arlington, TX 76019-0528

Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Ramat Aviv 69978
ISRAEL

Dr. Arthur S. Blaiwes
Code N712
Naval Training Systems Center
Orlando, FL 32813-7100

Dr. Bruce Bloxom
Defense Manpower Data Center
99 Pacific St.
Suite 155A
Monterey, CA 93943-3231

Cdt. Arnold Bohrer
Sectie Psychologisch Onderzoek
Rekruterings-En Selectiecentrum
Kwartier Koningen Astrid
Bruijnstraat
1120 Brussels, BELGIUM

Dr. Robert Breaux
Code 281
Naval Training Systems Center
Orlando, FL 32826-3224

Dr. Robert Brennan
American College Testing
Programs
P. O. Box 168
Iowa City, IA 52243

Dr. Gregory Candell
CTB/McGraw-Hill
2500 Garden Road
Monterey, CA 93940

Dr. John B. Carroll
409 Elliott Rd., North
Chapel Hill, NC 27514

Dr. John M. Carroll
IBM Watson Research Center
User Interface Institute
P.O. Box 704
Yorktown Heights, NY 10598

Dr. Robert M. Carroll
Chief of Naval Operations
OP-01B2
Washington, DC 20350

Dr. Raymond E. Christal
UES LAMP Science Advisor
AFHRL/MOEL
Brooks AFB, TX 78235

Mr. Hua Hua Chung
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Norman Cliff
Department of Psychology
Univ. of So. California
Los Angeles, CA 90089-1061

Director, Manpower Program
Center for Naval Analyses
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Hans F. Crombag
Faculty of Law
University of Limburg
P.O. Box 616
Maastricht
The NETHERLANDS 6200 MD

Ms. Carolyn R. Crone
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Timothy Davey
American College Testing Program
P.O. Box 168
Iowa City, IA 52243

Dr. C. M. Dayton
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Ralph J. DeAyala
Measurement, Statistics,
and Evaluation
Benjamin Bldg., Rm. 4112
University of Maryland
College Park, MD 20742

Dr. Lou DiBello
CERL
University of Illinois
103 South Mathews Avenue
Urbana, IL 61801

Dr. Dattprasad Divgi
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Mr. Hei-Ki Dong
Bell Communications Research
Room PYA-1K207
P.O. Box 1320
Piscataway, NJ 08855-1320

Dr. Fritz Drasgow
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
(2 Copies)

Dr. Stephen Dunbar
224B Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. James A. Earles
Air Force Human Resources Lab
Brooks AFB, TX 78235

Dr. Susan Embretson
University of Kansas
Psychology Department
426 Fraser
Lawrence, KS 66045

Dr. George Englehard, Jr.
Division of Educational Studies
Emory University
210 Fishburne Bldg.
Atlanta, GA 30322

ERIC Facility-Acquisitions
2440 Research Blvd, Suite 550
Rockville, MD 20850-3238

Dr. Benjamin A. Fairbank
Operational Technologies Corp.
5825 Callaghan, Suite 225
San Antonio, TX 78228

Dr. Marshall J. Farr, Consultant
Cognitive & Instructional Sciences
2520 North Vernon Street
Arlington, VA 22207

Dr. P-A. Federico
Code 51
NPRDC
San Diego, CA 92152-6800

Dr. Leonard Feldt
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. Richard L. Ferguson
American College Testing
P.O. Box 168
Iowa City, IA 52243

Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

Dr. Myron Fischl
U.S. Army Headquarters
DAPE-MRR
The Pentagon
Washington, DC 20310-0300

Prof. Donald Fitzgerald
University of New England
Department of Psychology
Armidale, New South Wales 2351
AUSTRALIA

Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Alfred R. Fregly
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Robert D. Gibbons
Illinois State Psychiatric Inst.
Rm 529W
1601 W. Taylor Street
Chicago, IL 60612

Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01003

Dr. Drew Gitomer
Educational Testing Service
Princeton, NJ 08541

Dr. Robert Glaser
Learning Research
& Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Sherrie Gott
AFHRL/MOMJ
Brooks AFB, TX 78235-5601

Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Michael Habon
DORNIER GMBH
P.O. Box 1420
D-7990 Friedrichshafen 1
WEST GERMANY

Prof. Edward Haertel
School of Education
Stanford University
Stanford, CA 94305

Dr. Ronald K. Hambleton
University of Massachusetts
Laboratory of Psychometric
and Evaluative Research
Hills South, Room 152
Amherst, MA 01003

Dr. Delwyn Harnisch
University of Illinois
51 Gerty Drive
Champaign, IL 61820

Dr. Grant Henning
Senior Research Scientist
Division of Measurement
Research and Services
Educational Testing Service
Princeton, NJ 08541

Ms. Rebecca Hetter
Navy Personnel R&D Center
Code 63
San Diego, CA 92152-6800

Dr. Thomas M. Hirsch
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Paul W. Holland
Educational Testing Service, 21-T
Rosedale Road
Princeton, NJ 08541

Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 92010

Ms. Julia S. Hough
Cambridge University Press
40 West 20th Street
New York, NY 10011

Dr. William Howell
Chief Scientist
AFHRL/CA
Brooks AFB, TX 78235-5601

Dr. Lloyd Humphreys
University of Illinois
Department of Psychology
603 East Daniel Street
Champaign, IL 61820

Dr. Steven Hunka
3-104 Educ. N.
University of Alberta
Edmonton, Alberta
CANADA T6G 2G5

Dr. Huynh Huynh
College of Education
Univ. of South Carolina
Columbia, SC 29208

Dr. Robert Jannarone
Elec. and Computer Eng. Dept.
University of South Carolina
Columbia, SC 29208

Dr. Kumar Joag-dev
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright Street
Champaign, IL 61820

Dr. Douglas H. Jones
1280 Woodfern Court
Toms River, NJ 08753

Dr. Brian Junker
Carnegie-Mellon University
Department of Statistics
Schenley Park
Pittsburgh, PA 15213

Dr. Michael Kaplan
Office of Basic Research
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Milton S. Katz
European Science Coordination
Office
U.S. Army Research Institute
Box 65
FPO New York 09510-1500

Prof. John A. Keats
Department of Psychology
University of Newcastle
N.S.W. 2308
AUSTRALIA

Dr. Jwa-keun Kim
Department of Psychology
Middle Tennessee State
University
P.O. Box 522
Murfreesboro, TN 37132

Mr. Soon-Hoon Kim
Computer-based Education
Research Laboratory
University of Illinois
Urbana, IL 61801

Dr. G. Gage Kingsbury
Portland Public Schools
Research and Evaluation Department
501 North Dixon Street
P. O. Box 3107
Portland, OR 97209-3107

Dr. William Koch
Box 7246, Meas. and Eval. Ctr.
University of Texas-Austin
Austin, TX 78703

Dr. Richard J. Koubek
Department of Biomedical
& Human Factors
139 Engineering & Math Bldg.
Wright State University
Dayton, OH 45435

Dr. Leonard Kroeker
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Dr. Jerry Lehnus
Defense Manpower Data Center
Suite 400
1600 Wilson Blvd
Rosslyn, VA 22209

Dr. Thomas Leonard
University of Wisconsin
Department of Statistics
1210 West Dayton Street
Madison, WI 53705

Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Charles Lewis
Educational Testing Service
Princeton, NJ 08541-0001

Mr. Rodney Lim
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert L. Linn
Campus Box 249
University of Colorado
Boulder, CO 80309-0249

Dr. Robert Lockman
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

Dr. Richard Luecht
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. George B. Macready
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

Dr. Clessen J. Martin
Office of Chief of Naval
Operations (OP 13 F)
Navy Annex, Room 2832
Washington, DC 20350

Dr. James R. McBride
HumRRO
6430 Elmhurst Drive
San Diego, CA 92120

Dr. Clarence C. McCormick
HQ, USMEPCOM/MEPCT
2500 Green Bay Road
North Chicago, IL 60064

Mr. Christopher McCusker
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert McKinley
Educational Testing Service
Princeton, NJ 08541

Mr. Alan Mead
c/o Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Timothy Miller
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Dr. William Montague
NPRDC Code 13
San Diego, CA 92152-6800

Ms. Kathleen Moreno
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Headquarters Marine Corps
Code MPI-20
Washington, DC 20380

Dr. Ratna Nandakumar
Educational Studies
Willard Hall, Room 213E
University of Delaware
Newark, DE 19716

Library, NPRDC
Code P201L
San Diego, CA 92152-6800

Librarian
Naval Center for Applied Research
in Artificial Intelligence
Naval Research Laboratory
Code 5510
Washington, DC 20375-5000

Dr. Harold F. O'Neil, Jr.
School of Education - WPH 801
Department of Educational
Psychology & Technology
University of Southern California
Los Angeles, CA 90089-0031

Dr. James B. Olsen
WICAT Systems
1875 South State Street
Orem, UT 84058

Office of Naval Research,
Code 1142CS
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Dr. Judith Orasanu
Basic Research Office
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311

Dr. Peter J. Pashley
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Circle, NW
Washington, DC 20036

Dr. James Paulson
Department of Psychology
Portland State University
P.O. Box 751
Portland, OR 97207

Dept. of Administrative Sciences
Code 54
Naval Postgraduate School
Monterey, CA 93943-5026

Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Malcolm Ree
AFHRL/MOA
Brooks AFB, TX 78235

Mr. Steve Reiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
310B Austin Peay Bldg.
Knoxville, TN 37916-0900

Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152-6800

Lowell Schoer
Psychological & Quantitative
Foundations
College of Education
University of Iowa
Iowa City, IA 52242

Dr. Mary Schratz
4100 Parkside
Carlsbad, CA 92008

Dr. Dan Segall
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Robin Shealy
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Kazuo Shigemasu
7-9-24 Kugenuma-Kaigan
Fujisawa 251
JAPAN

Dr. Randall Shumaker
Naval Research Laboratory
Code 5510
4555 Overlook Avenue, S.W.
Washington, DC 20375-5000

Dr. Richard E. Snow
School of Education
Stanford University
Stanford, CA 94305

Dr. Richard C. Sorensen
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Judy Spray
ACT
P.O. Box 168
Iowa City, IA 52243

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Dr. Peter Stoloff
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. William Stout
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

Mr. Brad Sympson
Navy Personnel R&D Center
Code-62
San Diego, CA 92152-6800

Dr. John Tangney
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Kikumi Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Maurice Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

Mr. Thomas J. Thomas
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Mr. Gary Thomasson
University of Illinois
Educational Psychology
Champaign, IL 61820

Dr. Robert Tsutakawa
University of Missouri
Department of Statistics
222 Math. Sciences Bldg.
Columbia, MO 65211

Dr. Ledyard Tucker
University of Illinois
Department of Psychology
603 E. Daniel Street
Champaign, IL 61820

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 440
St. Paul, MN 55114

Dr. Frank L. Vicino
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

Dr. Michael T. Waller
University of Wisconsin-Milwaukee
Educational Psychology Department
Box 413
Milwaukee, WI 53201

Dr. Ming-Mei Wang
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Thomas A. Warm
FAA Academy AAC934D
P.O. Box 25082
Oklahoma City, OK 73125

Dr. Brian Waters
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. Ronald A. Weitzman
Box 146
Carmel, CA 93921

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Douglas Wetzel
Code 51
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Rand R. Wilcox
University of Southern
California
Department of Psychology
Los Angeles, CA 90089-1061

German Military Representative
ATTN: Wolfgang Wildgrube
Streitkraefteamt
D-5300 Bonn 2
4000 Brandywine Street, NW
Washington, DC 20016

Dr. Bruce Williams
Department of Educational
Psychology
University of Illinois
Urbana, IL 61801

Dr. Hilda Wing
Federal Aviation Administration
800 Independence Ave, SW
Washington, DC 20591

Mr. John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering
Cancer Center
1275 York Avenue
New York, NY 10021

Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
Code 51
San Diego, CA 92152-6800

Dr. Kentaro Yamamoto
02-T
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Dr. Wendy Yen
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

Dr. Joseph L. Young
National Science Foundation
Room 320
1800 G Street, N.W.
Washington, DC 20550

Mr. Anthony R. Zara
National Council of State
Boards of Nursing, Inc.
625 North Michigan Avenue
Suite 1544
Chicago, IL 60611