The Annals of Statistics
1995, Vol. 23, No. 6, 2116–2144

POSTERIOR CONVERGENCE GIVEN THE MEAN

BY B. CLARKE[1] AND J. K. GHOSH[2]

Purdue University

For various applications one wants to know the asymptotic behavior of $w(\theta|\bar{X})$, the posterior density of a parameter θ given the mean $\bar{X}$ of the data rather than the full data set. Here we show that $w(\theta|\bar{X})$ is asymptotically normal in an L^1 sense, and we identify the mean of the limiting normal and its asymptotic variance. The main results are first proved assuming that $X_1, \dots, X_n, \dots$ are independent and identical; suitable modifications to obtain results for the nonidentical case are given separately. Our results may be used to construct approximate HPD (highest posterior density) sets for the parameter which is of use in the statistical theory of standardized educational tests. They may also be used to show the covariance between two test items conditioned on the mean is asymptotically nonpositive. This has implications for constructing tests of item independence.

1. Introduction. Let X_i for $i = 1, 2, \dots$ be a sequence of independently and identically distributed (iid) random variables taking values in a k-dimensional regular minimal lattice of common step length l with probability function $p_\theta(x)$ depending on a d-dimensional Euclidean parameter $\theta = (\theta_1, \dots, \theta_d)$, distributed according to a continuous density w supported on the parameter space Ω. Under strong enough moment assumptions on the X_i's we show that the posterior distribution $w(\theta|\bar{X})$ of θ given the mean $\bar{X}$ is asymptotically normal in an L^1 sense. We identify the location and asymptotic variance of the approximating normal as $\hat{\theta}$ and $J_\mu^t(\hat{\theta})\Sigma^{-1}(\hat{\theta})J_\mu(\hat{\theta})$, where, for $d = k$, $\hat{\theta} = \mu^{-1}(\bar{X})$, at least on a neighborhood of θ, the true value of the parameter J_μ is the $k \times d$ derivative matrix generated by μ as a function of θ, and Σ^{-1} is the covariance matrix of any X_i. A result for $d < k$ is also given.

If $\bar{X}$ is sufficient, then $w(\theta|\bar{X}) = w(\theta|X^n)$, where $X^n = (X_1, \dots, X_n)$, so existing results imply asymptotic normality. When $\bar{X}$ is not sufficient, these results [see Le Cam (1958), Bickel and Yahav (1969) and Walker (1969); there are many others] do not apply. In addition, Le Cam (1953) proves a version of the desired result for the maximum likelihood estimator which is asymptotically sufficient and Doksum and Lo (1990) establish a form of the result for location families and equivariant estimators.

Received September 1992; revised January 1995.
[1]Supported by NSERC Operating Grant 5-81506
[2]Supported by NSF Grant DMS 93-077-27.
AMS 1991 *subject classifications.* Primary 62F15; secondary 62E20.
Key words and phrases. Asymptotic normality, posterior, educational testing, sample mean.

In Section 3 we generalize our results to the case of independent nonidentically distributed (inid) random variables. These comprise a sort of "folk theorem" in the educational testing circle according to Holland (1991), who originally suggested the problem.

In educational testing, the vector θ represents an aptitude and the X_i's are the scores on the ith test item. It is often natural to condition on the total score $n\overline{X}$ [see Yen (1984)] rather than on the full data set to avoid data storage problems. Our results then provide approximate highest posterior density sets for the parameter. In data analysis, practitioners often group data according to the value of a sum. Our theorems allow a form of asymptotic normality to apply within each group. In addition, Ackerman (1991) assumes such a result for the purpose of evaluating the influence of dimensionality of a parameter on test item bias.

An example in which the X_i's are not identical and $\overline{X}$ is not sufficient is a modification of the Rasch model [see Lindsay, Clogg and Grego (1991) and Hambleton (1989)] in which $(\phi_i - \theta_j)$ is replaced by $\alpha_j(\phi_i - \theta_j)$, where the α_j's and θ_j's are known, and the task is to estimate ϕ_i for a fixed value of i. We obtain a general result on the asymptotic normality of $w(\theta|\overline{X})$ applicable in this case.

One of the three main assumptions for many models in educational testing is that the data be conditionally independent, given θ; see Lord (1980) and Bartholomew (1987). In part, Junker (1993) gives a heuristic argument suggesting that a hypothesis test for the conditional independence given θ of test items i and j could be based on the behavior of $\mathrm{Cov}(X_i, X_j|\overline{X})$, provided it is nonpositive. We give conditions under which this expression is asymptotically nonpositive for lattice-valued random variables that are conditionally independent given θ. Note that this expression is a manifest quantity; that is, it can be calculated from the data without reference to the underlying parametric family. This supports Junker's program of characterizing the desired latent properties of standardized tests in terms of manifest quantities.

The structure of the paper is as follows. In Section 2 we state and prove our results for the case of independent and identical lattice-valued random variables. First we consider the case that $d = k$ and the parameter space is compact. Then we give generalizations to $d < k$ and to noncompact parameter spaces. In Section 3 we give analogous results for the case of inid lattice-valued random variables. Section 4 contains the application discussed in the previous paragraph.

2. Identically distributed random variables.

We demonstrate asymptotic normality of $w(\theta|\overline{X})$ when the X_i's are iid in three increasingly general results. The first case is for $d = k$ and a compact parameter space. Our technique is based on a local limit theorem in Bhattacharya and Rao (1986) (hereafter referred to as BR) and a proposition about these quantities which generalizes a result in BR.

We use a three term upper bound on the L^1 distance between the posterior $w(\theta|\bar{X})$ and the target normal, denoted $n(\theta;\theta_0,\hat{\theta})$. The three terms result from using three normal approximations. The first is the target normal itself,

$$n\big(\theta;\theta_0,\hat{\theta}\big) = \sqrt{\left|nJ_\mu^t(\theta_0)\Sigma^{-1}(\theta_0)J_\mu(\theta_0)\right|}/(2\pi)^{d/2} \times \exp\Big(-(n/2)(\theta-\hat{\theta})^t J_\mu^t(\theta_0)\Sigma^{-1}(\theta_0)J_\mu(\theta_0)(\theta-\hat{\theta})\Big) \tag{2.1}$$

where $|\cdot|$ denotes the determinant and $\hat{\theta}=\mu^{-1}(\bar{X})$ for $d=k$ [where $\mu(\theta)=E_\theta X_1$], which need only be defined near θ_0.

The second normal approximation is obtained from a uniformized local limit theorem. Since the conditional density of $\bar{X}$ given θ, $p_\theta(\bar{x})$, is not known, we require a local limit theorem and its Edgeworth refinements to approximate $p_\theta(\bar{x})$ sufficiently well both for $\|\bar{X}-\theta\|=O(1/\sqrt{n})$ as well as for much larger deviations. The density of $\bar{X}$ can be approximated by a sum whose terms are normal densities multiplied by polynomials. The rate at which the distance between $p_\theta(\bar{X})$ and its closest approximation of this type tends to zero depends on the number of moments assumed to exist. One such result can be found in BR. Let

$$q_{\theta rn}(\bar{X}) = \frac{l}{n^{k/2}}\sum_{i=1}^{r}\frac{f_i\big(\sqrt{n}\,(\bar{X}-\mu(\theta))\big)}{n^{(i-1)/2}}\varphi_{\Sigma(\theta)}\big(\sqrt{n}\,(\bar{X}-\mu(\theta))\big) \tag{2.2}$$

be the r term approximation to $p_\theta(\bar{X})$, where $f_1\equiv 1$, and for $i>1$, f_i is a polynomial of degree at most $3r$ in k variables and $\varphi_{\Sigma(\theta)}$ is the normal density with mean 0 and variance $\Sigma(\theta)$. Here, r will always be a positive integer. The coefficients of f_i depend on θ also; however, we suppress this because it will not affect our arguments.

The third normal approximation is a variant on (2.2), to wit,

$$q_{\theta\theta_0 rn}(\bar{X}) = \frac{l}{n^{k/2}}\sum_{i=1}^{r}\frac{f_i\big(\sqrt{n}\,(\bar{X}-\mu(\theta))\big)}{n^{(i-1)/2}}\varphi_{\Sigma(\theta_0)}\big(\sqrt{n}\,(\bar{X}-\mu(\theta))\big) \tag{2.3}$$

in which the variance matrix is evaluated at θ_0.

We recall that the joint density for $(\Theta,\bar{X})$ is $w(\theta)p_\theta(\bar{X})=w(\theta|\bar{X})m(\bar{X})$, where $m(\bar{X})$ is the mixture of densities $m(\bar{X})=\int_\Omega w(\theta)p_\theta(\bar{X})\,d\theta$. We denote mixtures over approximations (2.2) and (2.3) with respect to w by $m_r(\bar{X})$ and $m_{\theta_0 r}(\bar{X})$, respectively. For brevity we omit subscripts, superscripts and arguments where no confusion will result.

Shrinking neighborhoods in the sample space and in the parameter space are essential to the proof. We denote them $U_{\theta_0 n}=\{X^n\colon \|\bar{X}-\mu(\theta_0)\|\le k_n/\sqrt{n}\}$ and $U'_{n,\theta_0}=\{\theta\colon \|\mu(\theta)-\mu(\theta_0)\|\le k'_n/\sqrt{n}\}$, where $k_n/\sqrt{n}$, $k'_n/\sqrt{n}\to 0$ and $\|\cdot\|$ is a norm on the lattice L, assumed to be embedded in k-dimensional real space. The defining condition on $U=U_{\theta_0 n}$ can be equivalently expressed as $\|\mu(\hat{\theta})-\mu(\theta_0)\|\le k_n/\sqrt{n}$. To permit upper bounds, Taylor expansions of μ can be used to obtain sets containing $U_{\theta_0 n}$ and $U'_{\theta_0 n}$. The defining conditions become $\|\hat{\theta} - \theta_0\| \le k_n/\alpha\sqrt{n}$ and $\|\theta - \theta_0\| \le k'_n/\alpha\sqrt{n}$, where $\alpha = \inf\|\nabla\mu(\theta')\|$ and the infimum is over θ' in a ball of radius ε centered at θ_0. We will use $k_n = c(\ln n)^{1/2}$ and $k'_n = c'(\ln n)^{1/2}$, where $c', c > 0$ and $c' - c > 0$.

First we state and prove a uniform version of Theorem 22.1 in BR.

PROPOSITION 2.1. *For $r \ge 1$ suppose that $E_\theta\|X_1 - \mu(\theta)\|^{r+2}$ is continuous as a function of $\theta \in C$ compact. Assume also that for all x, $p_\theta(x)$ is continuous in θ. Then, provided that $\Sigma(\theta)$ is positive definite on C,*

$$\sup_{\theta\in K}\sup_{\alpha\in L}\left(1 + \left\|\frac{\alpha - n\mu(\theta)}{\sqrt{n}}\right\|^{r+1}\right)\left|p_\theta\left(\frac{\alpha}{n}\right) - q_{\theta r}\left(\frac{\alpha}{n}\right)\right| = O\left(\frac{1}{n^{(k+r)/2}}\right). \tag{2.4}$$

PROOF. For fixed θ we have the desired rate: We use the result of BR and the triangle inequality (add and subtract $q_{\theta r+1}$) to obtain

$$\begin{aligned}
&\left(1 + \left\|\frac{\alpha - n\mu(\theta)}{\sqrt{n}}\right\|^{r+1}\right)\left|p_\theta\left(\frac{\alpha}{n}\right) - q_{\theta r}\left(\frac{\alpha}{n}\right)\right| \\
&\qquad \le O\left(\frac{1}{n^{(k+r)/2}}\right) + \left(1 + \left\|\frac{\alpha - n\mu(\theta)}{\sqrt{n}}\right\|^{r+1}\right)\frac{f_{r+1}\big(\sqrt{n}\,(\alpha/n - \mu(\theta))\big)}{n^{k/2}n^{r/2}} \\
&\qquad\quad \times \varphi_{\Sigma(\theta)}\left(\sqrt{n}\left(\frac{\alpha}{n} - \mu(\theta)\right)\right).
\end{aligned}$$

The last term is seen to be $o(1/n^{(k+r)/2})$.

To finish, we first note that the BR result holds uniformly over compact sets and the characteristic function $\phi(\theta, t) = E_\theta(\exp(i\langle t, \theta\rangle))$ is continuous jointly in t and θ by the continuity of $p_\theta(x)$. Fix $\theta_0 \in K$. For a sufficiently small neighborhood U_{θ_0} of θ_0, the two t-sets in the proof of BR's result can be chosen so as to satisfy (i) the expansion for the characteristic function holds with uniformly small remainder and (ii) on the second t-set, $\phi(\theta, t)$ for $\theta \in U_{\theta_0}$ is uniformly bounded away from unity, which is enough for the BR proof. By the Heine–Borel theorem, the proof is complete. □

THEOREM 2.1. *Let $\Omega \subset \mathbb{R}^d$ be compact. Assume that on Ω, $\mathrm{Var}_\theta\, X_1 = \Sigma(\theta)$ satisfies $\eta_1 I_d \le \Sigma(\theta) \le \eta_2 I_d$ for some $\eta_1, \eta_2 > 0$, where I_d is the $k \times k$ identity matrix, and that the entries of $\Sigma(\theta)$ are continuously differentiable. Assume also that $\mu(\theta) = E_\theta X_1$ has two continuous derivatives, is locally invertible at the interior point θ_0 and its $d \times k$ derivative matrix $J_\mu(\theta)$ has rank d at $\theta = \theta_0$, where $d = k$. Then, if the hypotheses of Proposition 2.1 are satisfied with r replaced by $r + 1$, where $r > \max(0, d/2 - 1, (2/3)d - 4/3)$, we have that*

$$E_{\theta_0}\int\left|w(\theta|\overline{X}) - n\big(\theta; \theta_0, \hat{\theta}\big)\right| d\theta \to 0 \tag{2.5}$$

as $n \to \infty$.

REMARK. Even though $d = k$, we distinguish them here so as to emphasize the role of the sample space and of the parameter space. This permits us to handle the case $d < k$ conveniently later.

REMARK. Replacing θ_0 by $\hat{\theta}$ in the target normal and applying Scheffe's theorem (see BR page 6) we observe that the result continues to hold if we change the variance to $J_\mu^t(\hat{\theta})\Sigma^{-1}(\hat{\theta})J_\mu(\hat{\theta})$. As a consequence, convergence holds with the expectation taken with respect to the mixture density. This applies to Theorems 2.2, 2.3, 3.1 and 3.2 also. Note that for $d = 1$, 2, and 3, four moments are required.

PROOF. We use K to denote a positive constant, not in general the same from occurrence to occurrence. We proceed in four steps. The first step is to obtain lower bounds on $\chi_U m_r(\overline{X})$ and $\chi_U |m(\overline{X}) - m(\overline{X})|$, and note a straightforward upper bound on (2.5) which has three terms. The following three steps will deal with each term in turn.

Step 1, *part* 1: We show that there is a $K > 0$ so that

$$\chi_U m_r(\overline{X}) \geq (K/n^{(k+d)/2})\chi_U. \tag{2.6}$$

First note that products of the form $f_i(\sqrt{n}(\overline{X} - \mu(\theta)))\varphi_{\Sigma(\theta)}(\sqrt{n}(\overline{X} - \mu(\theta)))$ are bounded in absolute value by constants for $i \geq 2$. We can write

$$m_r(\overline{X}) \geq \frac{K}{n^{k/2}} \times \int_{\|\hat{\theta}-\theta\| \leq k'_n/\sqrt{n}} \exp\left(-\left(\frac{n}{2}\right)(\hat{\theta}-\theta)^t J_\mu(\tilde{\theta})^t \Sigma^{-1}(\theta) J_\mu(\tilde{\theta})(\hat{\theta}-\theta)\right) d\theta$$

by a Taylor expansion, where $\tilde{\theta}$ lies on the straight line joining θ and $\hat{\theta}$. Since $J_\mu(\tilde{\theta})^t\Sigma^{-1}(\theta)J_\mu(\tilde{\theta})$ is bounded above and bounded away from singularity, the last expression gives (2.6) by using the transformation $\varphi = \sqrt{n}(\theta - \hat{\theta})$.

Step 1, *part* 2: We show that

$$\begin{aligned} &\chi_U\left|m(\overline{X}) - m_r(\overline{X})\right| \\ &\quad\leq K\chi_U(k'_n)^{d/2}\max\left(\frac{1}{n^{(k+d+r)/2}}, \frac{1}{n^{(k+r+1)/2}}\left(\frac{1}{n^{d(1/2-\delta)}} + \frac{1}{n^{\delta(r+2)}}\right)\right), \end{aligned} \tag{2.7}$$

for any $\delta \in (0, 1/2)$, where χ_A is the indicator function of the set A. Note that the left-hand side is at most

$$\begin{aligned} &\chi_U \int_{U'} \left|p_\theta(\overline{X}) - q_{\theta r}(\overline{X})\right| w(\theta)\, d\theta + \chi_U \int_{U'^c} \left|p_\theta(\overline{X}) - q_{\theta r}(\overline{X})\right| w(\theta)\, d\theta \\ &\quad\leq \chi_U\left[\sup_{\theta\in U'} w(\theta)\right]\mathrm{Vol}(U')\frac{K}{n^{(k+r)/2}} \\ &\qquad+ \chi_U \int_{U'^c} \left|p_\theta(\overline{X}) - q_{\theta, r+1}(\overline{X})\right| w(\theta)\, d\theta \\ &\qquad+ \chi_U \int_{U'^c} \left|q_{\theta, r+1}(\overline{X}) - q_{r\theta}(\overline{X})\right| w(\theta)\, d\theta. \end{aligned} \tag{2.8}$$

The first term in (2.8) is bounded by $\chi_U K(k_n')^{d/2}/n^{(k+d+r)/2}$. To bound the second term, write $V_1 = \{\theta | k_n'/\sqrt{n} \leq \|\mu(\theta) - \mu(\theta_0)\| \leq n^{\delta}/\sqrt{n}\}$ and $V_2 = \{\theta | n^{\delta}/\sqrt{n} \leq \|\mu(\theta) - \mu(\theta_0)\|\}$ for some $\delta \in (0, 1/2)$. By Proposition 2.1, the second term is bounded by

$$\begin{aligned}(2.9)\qquad &\chi_U K \sup_{\theta \in V_1} w(\theta)\left(\frac{n^{\delta}}{\sqrt{n}}\right)^d \frac{1}{n^{(k+r+1)/2}}\\ &\quad + \chi_U K \sup_{\theta \in V_2} w(\theta) \int_{V_2} \left| p_\theta(\bar{X}) - q_{\theta, r+1}(\bar{X}) \right| d\theta.\end{aligned}$$

Restrict the supremum over the lattice to U and the supremum over θ to V_2 to get $\sqrt{n}\|\bar{X} - \mu(\theta)\| \geq (1-\delta)n^{\delta}$, for n sufficiently large. Proposition 2.1 gives $|p_\theta(\bar{X}) - q_{\theta, r+1}(\bar{X})|$ is less than $(K/n^{(k+r+1)/2})(1/n^{\delta(r+2)})$. Using this in (2.9), the second term in (2.8) is less than

$$\begin{aligned}(2.10)\qquad &K\chi_U \sup_{\theta} w(\theta) \frac{n^{\delta d}}{n^{(k+r+d+1)/2}} + \chi_U K \sup_{\theta} w(\theta) \frac{1}{n^{(k+r+1)/2} n^{\delta(r+2)}}\\ &\quad \leq K\chi_U \frac{1}{n^{(k+r+1)/2}}\left(\frac{1}{n^{d(1/2-\delta)}} + \frac{1}{n^{\delta(r+2)}}\right).\end{aligned}$$

The third term in (2.8) is bounded by

$$(2.11)\qquad \chi_U \int_{V_n} w(\theta) \frac{\left| f_{r+1}\big(\sqrt{n}\,(\bar{X} - \mu(\theta))\big)\right|}{n^{r/2}} \varphi_{\Sigma(\theta)}\big(\sqrt{n}\,(\bar{X} - \mu(\theta))\big)\, d\theta,$$

where $U'^c \subset V_n$ is defined by $V_n = \{\theta | \|\mu(\theta) - \bar{X}\| \geq (c' - c)\sqrt{(\ln n)/n}\}$. This follows by using the triangle inequality since the inequalities in U and U'^c go in opposite directions. To show (2.7) we control the integral term in (2.11). It is bounded by

$$\begin{aligned}(2.12)\qquad &\chi_U \frac{K}{n^{r/2}} \exp\left(-\left(\frac{n}{4}\right)\eta_1 (c' - c)^2((\ln n)/n)\right)\\ &\quad \times \int w(\theta) \left| f_{r+1}\big(\sqrt{n}\,(\bar{X} - \mu(\theta))\big)\right| \varphi_{(1/2)\Sigma(\theta)}\big(\sqrt{n}\,(\bar{X} - \mu(\theta))\big)\, d\theta.\end{aligned}$$

The product $f_{r+1}\varphi_{(1/2)\Sigma(\theta)}$ is uniformly bounded by a constant, so the integral factor can be absorbed into K. The exponential factor is $1/n^{(\eta_1/4)(c'-c)^2}$, so choosing c' large enough gives (2.7).

Step 1, *part* 3: We upper bound the L^1 distance in (2.5) by the sum

$$(2.13)\qquad E_{\theta_0} \int \left| \frac{w(\theta) p_\theta(\bar{X})}{m(\bar{X})} - \frac{w(\theta) q_{\theta r}(\bar{X})}{m_r(\bar{X})} \right| d\theta$$

$$(2.14)\qquad + E_{\theta_0} \int \left| \frac{w(\theta) q_{\theta r}(\bar{X})}{m_r(\bar{X})} - \frac{w(\theta) q_{\theta\theta_0} r(\bar{X})}{m_{\theta_0 r}(\bar{X})} \right| d\theta$$

$$(2.15)\qquad + E_{\theta_0}\int\left|\frac{w(\theta)q_{\theta\theta_0 r}(\overline{X})}{m_{\theta_0 r}(\overline{X})} - n\big(\theta;\theta_0,\hat{\theta}\big)\right| d\theta.$$

Step 2, *part* 1: We use (2.6) and (2.7) with $\delta = 1/4$ to obtain a lower bound for $\chi_U m(\overline{X})$:

$$(2.16)\qquad \chi_U m(\overline{X}) \geq \chi_U\big(m_r(\overline{X}) - |m(\overline{X}) - m_r(\overline{X})|\big) \geq \chi_U K/n^{(k+d)/2},$$

provided $r > \max(0, (1/2)d - 1, (2/3)d - 4/3)$.

Step 2, *part* 2: Expression (2.13) equals

$$(2.17)\qquad E_{\theta_0}\chi_{U^c}\int\left|\frac{w(\theta)p_\theta(\overline{X})}{m(\overline{X})} - \frac{w(\theta)q_{\theta r}(\overline{X})}{m_r(\overline{X})}\right| d\theta$$

$$(2.18)\qquad + E_{\theta_0}\chi_U\int\left|\frac{w(\theta)p_\theta(\overline{X})}{m(\overline{X})} - \frac{w(\theta)q_{\theta r}(\overline{X})}{m_r(\overline{X})}\right| d\theta.$$

For n large enough, the first term in the sum which gives $q_{\theta r}(\overline{X})$ dominates so that $q_{\theta r}(\overline{X})$ is positive everywhere [see the proof of (2.6)]. As a result (2.17) is upper bounded by $E_{\theta_0}\chi_{U^c}(\int w(\theta)p_\theta(\overline{X})/m(\overline{X})\,d\theta + \int w(\theta)q_{\theta r}(\overline{X})/m_r(\overline{X})\,d\theta)$, which is less than $P_{\theta_0}(U^c)$ and so goes to zero. For expression (2.18) we use (2.6) and (2.7) (with $\delta = 1/4$) as well as the fact that $\chi_U\int|p_\theta(\overline{X}) - q_{\theta r}(\overline{X})|w(\theta)\,d\theta$ is bounded above by the right-hand side of (2.7). Now, by adding and subtracting $q_{\theta r}(\overline{X})/m(\overline{X})$ we have that

$$\chi_U\int w(\theta)\left|\frac{p_\theta(\overline{X})}{m(\overline{X})} - \frac{q_{\theta r}(\overline{X})}{m(\overline{X})}\right| d\theta$$

$$(2.19)\qquad \leq \chi_U\int w(\theta)\frac{|p_\theta(\overline{X}) - q_{\theta r}(\overline{X})|}{m(\overline{X})}\,d\theta$$

$$+ \chi_U\int\frac{w(\theta)q_{\theta r}(\overline{X})}{m_r(\overline{X})}\frac{|m(\overline{X}) - m_r(\overline{X})|}{m(\overline{X})}\,d\theta.$$

Using (2.16), the right-hand side of (2.19) can be bounded above by

$$\chi_U K(k_n')^{d/2}\max\left(\frac{1}{n^{r/2}}, \frac{n^{d/2}}{n^{(r+1)/2}}\left(\frac{1}{n^{d(1/2-\delta)}} + \frac{1}{n^{\delta(r+2)}}\right)\right).$$

The first entry of the maximum goes to zero. The second entry is the sum of $1/n^{(r+1)/2-\delta d}$ and $1/n^{(r+1)/2+\delta(r+2)-d/2}$, which goes to zero for $r > \max(0, d/2 - 1, (2d/3) - 4/3)$. Finally, applying E_{θ_0} to (2.19) and its upper bound (which is nonrandom) gives a bound on (2.18) which goes to zero.

Step 3, *part* 1: Next we show (2.14) tends to zero. We upper bound it by

$$(2.20)\qquad E_{\theta_0}\chi_U\int_{U'} w(\theta)\left|\frac{q_{\theta r}(\overline{X})}{m_r(\overline{X})} - \frac{q_{\theta\theta_0 r}(\overline{X})}{m_{\theta_0 r}(\overline{X})}\right| d\theta$$

$$(2.21)\qquad + E_{\theta_0}\chi_U\int_{U'^c}\frac{w(\theta)q_{\theta r}(\overline{X})}{m_r(\overline{X})}$$

$$+ E_{\theta_0} \chi_U \int_{U'^c} \frac{w(\theta) q_{\theta\theta_0 r}(\bar{X})}{m_{\theta_0 r}(\bar{X})}\, d\theta \tag{2.22}$$

$$+ E_{\theta_0} \chi_{U^c} \int \left| \frac{w(\theta) q_{\theta r}(\bar{X})}{m_r(\bar{X})} - \frac{w(\theta) q_{\theta\theta_0 r}(\bar{X})}{m_{\theta_0 r}(\bar{X})} \right| d\theta. \tag{2.23}$$

Step 3, *part* 2: Three of the four terms in the last upper bound are easy to control. Term (2.23) tends to zero by the same reasoning as was used for (2.17): the triangle inequality allows us to use 2 as an upper bound for the integral and gives the convergence to zero.

By reasoning similar to that used to prove (2.6) one can prove

$$\chi_U m_{\theta_0 r}(\bar{X}) \geq \left(K/n^{(k+d)/2}\right)\chi_U. \tag{2.24}$$

By use of (2.24) and (2.6), to prove that (2.21) and (2.22) go to zero it is enough to show

$$E_{\theta_0} \chi_U \int_{U'^c} q_{\theta r}(\bar{X})\, d\theta = o\left(\frac{1}{n^{(k+d)/2}}\right) \tag{2.25}$$

and

$$E_{\theta_0} \chi_U \int_{U'^c} q_{\theta\theta_0 r}(\bar{X})\, d\theta = o\left(\frac{1}{n^{(k+d)/2}}\right). \tag{2.26}$$

We see that the absolute values of the left-hand sides of (2.25) and (2.26) are upper bounded by a sum of r terms that may be controlled alike. So, for $c' - c$ large enough expressions (2.25) and (2.26) can be forced to go to zero at any rate of the form $o(1/n^{\alpha})$ for $\alpha > 0$.

Step 3, *part* 3: For expression (2.20) our technique will be similar to that used for (2.18). By adding and subtracting $q_{\theta\theta_0 r}(\bar{X})/m_r(\bar{X})$ and using (2.6) we see that (2.20) is upper bounded by

$$\begin{aligned}
&E_{\theta_0} \chi_U \int_{U'} w(\theta) \frac{\left| q_{\theta r}(\bar{X}) - q_{\theta\theta_0 r}(\bar{X}) \right|}{m_r(\bar{X})}\, d\theta \\
&\quad + E_{\theta_0} \chi_U \int_{U'} \frac{w(\theta) q_{\theta\theta_0 r}(\bar{X})}{m_{\theta_0 r}(\bar{X})} \frac{\left| m_r(\bar{X}) - m_{\theta_0 r}(\bar{X}) \right|}{m_r(\bar{X})}\, d\theta \\
&\quad \leq K n^{(k+d)/2} \Big[E_{\theta_0} \chi_U \int_{U'} w(\theta) \left| q_{\theta r}(\bar{X}) - q_{\theta\theta_0 r}(\bar{X}) \right| d\theta \\
&\qquad\qquad + E_{\theta_0} \chi_U \left| m_r(\bar{X}) - m_{\theta_0 r}(\bar{X}) \right| \Big] \\
&\quad \leq K n^{(k+d)/2} \Big[2 E_{\theta_0} \chi_U \int_{U'} w(\theta) \left| q_{\theta r}(\bar{X}) - q_{\theta\theta_0 r}(\bar{X}) \right| d\theta \\
&\qquad\qquad + E_{\theta_0} \chi_U \int_{U'^c} w(\theta) \left| q_{\theta r}(\bar{X}) - q_{\theta\theta_0 r}(\bar{X}) \right| d\theta \Big].
\end{aligned}$$

The second term in brackets goes to zero by use of (2.25) and (2.26).

For the first term, it is enough to show

$$\chi_U \chi_{U'} \left| q_{\theta r}(\bar{X}) - q_{\theta\theta_0 r}(\bar{X}) \right| = O\left(\chi_U \chi_{U'} \frac{(k_n + k'_n)^{3r}}{n^{(k+1)/2}} \right), \tag{2.27}$$

for then the integration will give a factor of $K(k'_n/\sqrt{n})^d$ so that term will tend to zero also. Since f_i has degree $3(i-1)$ and on the intersection of U_n and U', $\|\sqrt{n}(\bar{X} - \mu(\theta))\| \le k_n + k'_n$, so we have that the left-hand side of (2.27) is bounded above by

$$\begin{aligned}
&\frac{K\chi_U \chi_{U'}}{n^{k/2}} \sum_{i=1}^{r} \frac{\left| f_i\left(\sqrt{n}\,(\bar{X} - \mu(\theta))\right)\right|}{n^{(i-1)/2}} \\
&\qquad \times \left| \exp\left(-(n/2)\left(\bar{X} - \mu(\theta)\right)^t \Sigma^{-1}(\theta)\left(\bar{X} - \mu(\theta)\right)\right) \right. \\
&\qquad\qquad \left. - \exp\left(-(n/2)\left(\bar{X} - \mu(\theta)\right)\Sigma^{-1}(\theta_0)\left(\bar{X} - \mu(\theta)\right)\right) \right| \\
&\qquad \le \frac{K\chi_U \chi_{U'}}{n^{k/2}} (k_n + k'_n)^{3(r-1)} \left(\frac{n}{2}\right) \left\| \bar{X} - \mu(\theta) \right\|^2 \\
&\qquad\quad \times \left\| \Sigma^{-1}(\theta) - \Sigma^{-1}(\theta_0) \right\|,
\end{aligned}$$

in which we have used $|e^{-x} - e^{-y}| \le |x - y|$ and norm inequalities on the upper bound resulting from that inequality. The matrix norm takes the largest eigenvalue. Using the restriction to U and U' again, we obtain the bound $K\chi_U \chi_{U'}(k_n + k'_n)^{3(r-1)}\|\Sigma^{-1}(\theta) - \Sigma^{-1}(\theta_0)\|/n^{k/2}$. Since all Euclidean norms are equivalent, we can replace the matrix norm with any norm. We choose the norm which sums the absolute values of the entries. Each term in that sum admits a Taylor expansion which can be bounded from above by $(k'_n/\sqrt{n})$ times a positive constant. There are only finitely many constants, so taking the maximum gives an upper bound $K(k'_n/\sqrt{n}) \le K(k_n + k'_n)/\sqrt{n}$ which finishes the proof of (2.27).

Step 4, *part* 1: In this final step we show that (2.15) goes to zero. We start by bounding (2.15) from above by a sum of five terms, two of which are easy. Our bound is

$$E_{\theta_0} \chi_U \int_{U'} \left| \frac{w(\theta) q_{\theta\theta_0 r}(\bar{X})}{m_{\theta_0 r}(\bar{X})} - \frac{w(\theta) q_{\theta\theta_0 1}(\bar{X})}{m_{\theta_0 1}(\bar{X})} \right| d\theta \tag{2.28}$$

$$+ E_{\theta_0} \chi_U \int_{U'} \left| \frac{w(\theta) q_{\theta\theta_0 1}(\bar{X})}{m_{\theta_0 1}(\bar{X})} - n\left(\theta; \theta_0, \hat{\theta}\right) \right| d\theta \tag{2.29}$$

$$+ E_{\theta_0} \chi_U \int_{U'^c} \frac{w(\theta) q_{\theta\theta_0 r}(\bar{X})}{m_{\theta_0 r}(\bar{X})}\, d\theta \tag{2.30}$$

$$+E_{\theta_0}\,\chi_U \int_{U'^c} n\big(\theta;\theta_0,\hat{\theta}\big)\,d\theta \tag{2.31}$$

$$+E_{\theta_0}\,\chi_{U^c}\int\left|\frac{w(\theta)q_{\theta\theta_0}(\overline{X})}{m_{\theta_0 r}(\overline{X})}-n\big(\theta;\theta_0,\hat{\theta}\big)\right|d\theta. \tag{2.32}$$

Step 4, *part* 2: The term (2.30) is handled like (2.22) and (2.32) like (2.23) and (2.17).

Step 4, *part* 3: The next easiest term is (2.31). Since μ is invertible on a neighborhood of θ_0 for any $\eta > 0$ there is an $\varepsilon > 0$ so that $\|\mu(\theta)-\mu(\theta_0)\| < \varepsilon \Rightarrow \|\theta-\theta_0\| < \eta$ and $\|\mu(\hat{\theta})-\mu(\theta_0)\| < \varepsilon \Rightarrow \|\hat{\theta}-\theta_0\| < \eta$. For such a choice of ε we write (2.31) as

$$E_{\theta_0}\,\chi_U \int_{\|\mu(\theta)-\mu(\theta_0)\|>\varepsilon} n\big(\theta;\theta_0,\hat{\theta}\big)\,d\theta \tag{2.33}$$

$$+\,E_{\theta_0}\,\chi_U \int_{\varepsilon \ge \|\mu(\theta)-\mu(\theta_0)\| \ge k'_n\sqrt{n}} n\big(\theta;\theta_0,\hat{\theta}\big)\,d\theta. \tag{2.34}$$

For (2.34), restriction to U and to the domain of integration gives that $\|\theta-\hat{\theta}\| \le 2\eta$, so we can use a Taylor expansion and the triangle inequality to obtain

$$\big\|\nabla\mu(\tilde{\theta})\big\|\,\|\hat{\theta}-\theta| \ge \|\mu(\hat{\theta})-\mu(\theta)\big\| \ge (k'_n-k_n)/\sqrt{n} = (c'-c)\sqrt{\frac{\ln n}{n}}\,,$$

for some $\tilde{\theta}$ lying on the straight line joining θ and $\hat{\theta}$. By the continuity of the derivative we have that $(\hat{\theta}-\theta) \ge K(c'-c)\sqrt{(\ln n)/n}$. Now (2.34) is bounded by

$$\begin{aligned}&K\exp\left(-\left(\frac{n}{4}\right)K(c'-c)^2\left(\frac{\ln n}{n}\right)\right)\\ &\quad\times E_{\theta_0}\,\chi_U\int\exp\left(-\left(\frac{n}{4}\right)(\theta-\hat{\theta})^t\right.\\ &\qquad\left.\times J(\theta_0)^t\Sigma^{-1}(\theta_0)J(\theta)(\theta-\hat{\theta})\right)d\theta,\end{aligned} \tag{2.35}$$

which tends to zero.

For (2.33) we use a variant of the last argument. We note that local invertibility implies that given $\varepsilon > 0$ there is an $\eta > 0$ so that $\|\mu(\theta)-\mu(\theta_0)\| > \varepsilon \Rightarrow \|\theta-\theta_0\| > \eta$. By restriction to U we have that $\hat{\theta}$ and θ_0 are close, so we Taylor expand to get that there is a $K > 0$ so that $K\|\hat{\theta}-\theta_0\| \le k_n/\sqrt{n}$. Again by the triangle inequality, $\|\hat{\theta}-\theta\| \ge \eta-(Kk_n/\sqrt{n}) \ge \eta/2$. So, in this case we still get a bound much like (2.35). As a result, (2.33) goes to zero.

Step 4, *part* 4: Write expression (2.28) as

$$E_{\theta_0}\chi_U\int_{U'}\frac{w(\theta)q_{\theta\theta_0 1}(\bar{X})}{m_{\theta_0 1}(\bar{X})}$$

$$\times\left|1-[1+F(2,r)]\left(1+\frac{\int F(2,r)w(\theta)\varphi_{\Sigma(\theta_0)}\big(\sqrt{n}\,(\bar{X}-\mu(\theta))\big)\,d\theta}{\int w(\theta)\varphi_{\Sigma(\theta_0)}\sqrt{n}\,\big((\bar{X}-\mu(\theta))\big)\,d\theta}\right)^{-1}\right|d\theta,$$

where $F(2,r)=\sum_{i=2}^{r}f_i(\sqrt{n}\,(\bar{X}-\mu(\theta)))/n^{(i-1)/2}$. We bound each sum by $o(1)$. On U and U' each f_i is bounded by $K(k_n+k_n')^{3(i-1)}$, which is of lower order than $n^{(i-1)/2}$ and so the summation in the numerator goes to zero. Each term in the sum in the denominator is seen to be $o(1)$ by integrating over U' and U'^c, using the lower bound on $m_{\theta_0 r}(\bar{X})$ with the last bound on f_i, and applying the techniques used on the right-hand side of (2.13).

Step 4, *final part*: At last we deal with (2.29). We bound it by adding and subtracting

$$\frac{w(\theta)\exp\Big(-(n/2)\big(\bar{X}-\mu(\theta)\big)^t\Sigma^{-1}(\theta_0)\big(\bar{X}-\mu(\theta)\big)\Big)}{w(\theta_0)(2\pi)^{d/2}\big|nJ(\theta_0)^t\Sigma^{-1}(\theta_0)J(\theta_0)\big|^{-1/2}}$$

and

$$\frac{\exp\Big(-(n/2)\big(\bar{X}-\mu(\theta)\big)^t\Sigma^{-1}(\theta_0)\big(\bar{X}-\mu(\theta)\big)\Big)}{(2\pi)^{d/2}\big|nJ(\theta_0)^t\Sigma^{-1}(\theta_0)J(\theta_0)\big|^{-1/2}}.$$

Our upper bound on (2.29) is now

$$E_{\theta_0}\chi_U\left|1-\frac{\int w(\theta)e^{-(n/2)(\bar{X}-\mu(\theta))^t\Sigma^{-1}(\theta_0)(\bar{X}-\mu(\theta))}\,d\theta}{(2\pi)^{d/2}w(\theta_0)\big|nJ^t(\theta_0)\Sigma^{-1}(\theta_0)J(\theta_0)\big|^{-1/2}}\right| \tag{2.36}$$

$$+E_{\theta_0}\chi_U\int_{U'}\left|\frac{w(\theta)}{w(\theta_0)}-1\right|\frac{e^{-(n/2)(\bar{X}-\mu(\theta))^t\Sigma^{-1}(\theta_0)(\bar{X}-\mu(\theta))}}{(2\pi)^{d/2}\big|nJ^t(\theta_0)\Sigma^{-1}(\theta_0)J(\theta_0)\big|^{-1/2}}\,d\theta \tag{2.37}$$

(2.38)

$$+E_{\theta_0}\chi_U\int_{U'}\frac{\big|e^{-(n/2)(\bar{X}-\mu(\theta))\Sigma^{-1}(\theta_0)(\bar{X}-\mu(\theta))}-e^{-(n/2)(\theta-\hat{\theta})^tJ^t(\theta_0)\Sigma^{-1}(\theta_0)J(\theta_0)(\theta-\hat{\theta})}\big|}{(2\pi)^{d/2}\big|nJ^t(\theta_0)\Sigma^{-1}(\theta_0)J(\theta_0)\big|^{-1/2}}\,d\theta.$$

We note that there is a positive definite matrix M so that $(\hat{\theta}-\theta)^tJ^t(\tilde{\theta})\Sigma^{-1}(\theta_0)J(\tilde{\theta})(\hat{\theta}-\theta)\geq(\hat{\theta}-\theta)^tM(\hat{\theta}-\theta)$. As a result, (2.37) is bounded from above by

$$K\sup_{\theta\in U'}\left|\frac{w(\theta)}{w(\theta_0)}-1\right|E_{\theta_0}\chi_U\int_{U'}n^{d/2}\exp\left(-\left(\frac{n}{2}\right)(\hat{\theta}-\theta)^tM(\theta-\hat{\theta})\right)d\theta,$$

in which the integral is finite, and by the continuity of w the bound goes to zero.

For expression (2.38) we use techniques similar to those used for (2.27). Since $|e^{-x} - e^{-y}| \le |x - y|$, we obtain the upper bound

$$Kn^{d/2} E_{\theta_0} \chi_U \int_{U'} n \left\| (\theta - \hat{\theta}) \right\|^2 \left\| J(\tilde{\theta})^t \Sigma^{-1}(\theta_0) J(\tilde{\theta}) - J(\theta_0) \Sigma^{-1}(\theta_0) J(\theta_0) \right\| d\theta$$

after Taylor expansion of μ, where $\tilde{\theta}$ is on the straight line joining θ and $\hat{\theta}$. By reasoning used in the proof that (2.37) goes to zero, we have that $\|\sqrt{n}(\theta - \hat{\theta})\| \le K(k_n + k_n')^2$. Also since we have restricted to U and U', the norm of the difference of matrices can be controlled by a Taylor expansion.

Finally, for (2.36), consider the integral

$$\begin{aligned} &\chi_U \int_{U'} w(\theta) \exp\left(-\left(\frac{n}{2}\right) \left(\overline{X} - \mu(\theta)\right)^t \Sigma^{-1}(\theta_0) \left(\overline{X} - \mu(\theta)\right) \right) d\theta \\ &\quad + \chi_U \int_{U'^c} w(\theta) \exp\left(-\left(\frac{n}{2}\right) \left(\overline{X} - \mu(\theta)\right)^t \Sigma^{-1}(\theta_0) \left(\overline{X} - \mu(\theta)\right) \right) d\theta. \end{aligned} \tag{2.39}$$

For a lower bound we drop the second term and Taylor expand μ in the first. For an upper bound, observe that the second term in (2.39) can be bounded by K/n^{α}, where $\alpha > 0$ is an increasing function of $c' - c$. The first term in (2.39) can be bounded above by Taylor expanding μ. Thus, there are functions g_1, g_2 with $g_1(\varepsilon), g_2(\varepsilon) \to 1$ as $\varepsilon \to 0$, and on U,

$$g_1(\varepsilon) \le \frac{\int w(\theta) \exp\left(-(n/2) \left(\overline{X} - \mu(\theta)\right) \Sigma^{-1}(\theta_0) \left(\overline{X} - \mu(\theta)\right) \right) d\theta}{\left(w(\theta_0)(2\pi)^{d/2} \right) / \left| nJ^t(\theta_0) \Sigma^{-1}(\theta_0) J(\theta_0) \right|^{1/2}} \le g_2(\varepsilon).$$

Using the last pair of inequalities it is seen that (2.36) tends to zero also. □

REMARK. A version of Theorem 2.1 holds for continuous random variables. Indeed, a version of Proposition 2.1 can be obtained from Theorem 19.2 in BR. The only extra assumption is that for some $p \ge 1$ the pth power of the characteristic function of X be integrable. Doing this, the proof of Theorem 2.1 here applies to the continuous case also.

Next, we extend Theorem 2.1 to noncompact parameter spaces Ω. For C compact we define mixtures

$$m_C(\overline{X}) = \int_C \frac{w(\theta)}{W(C)} p_\theta(\overline{X}) \, d\theta \quad \text{and} \quad m_{C^c}(\overline{X}) = \int_{C^c} \frac{w(\theta)}{W(C^c)} p_\theta(\overline{X}) \, d\theta,$$

where W is the prior probability with density w. Again we use local invertibility of μ at θ_0. This means there is an open set O containing θ_0 so that the restriction of μ to O, $\mu|_O$: $O \to \mu(O)$ is invertible and that for $\theta \in O^c$, $\mu(\theta) \in \mu(O)^c$. Our result is the following.

THEOREM 2.2. *Assume the hypotheses of Theorem* 2.1, *including the lower bound on* r *and* $k = d$. *Also, assume that for all* k *components* $X_{(i)}$ *of* X_1 *the* $k + d + 1$ *central moment is uniformly bounded in* θ, *that is,* $\sup_{\theta \in \Omega} E_\theta |X_{(i)} - \mu_{(i)}(\theta)|^{k+d+1}$ *is finite for* $i = 1, \ldots, k$. *Then,*

$$E_{\theta_0} \int \left| w(\theta|\overline{X}) - n\left(\theta; \theta_0, \hat{\theta}\right) \right| d\theta \to 0. \tag{2.40}$$

PROOF. Let C be a compact set, to be specified shortly. Write the integral in (2.40) as a sum of an integral over C and an integral over C^c and let $w_C(\theta) = w(\theta)|_C/W(C)$. In the integral over C, add and subtract $w_C(\theta|\overline{X})$, apply the triangle inequality and then pull out $w_C(\theta|\overline{X})$ as a factor in the term which is a difference of posteriors to see that (2.40) is bounded from above by

$$E_{\theta_0}\int_C\left|\frac{w_C(\theta)p_\theta(\overline{X})}{m_C(\overline{X})} - n\big(\theta;\theta_0,\hat{\theta}\big)\right|d\theta \tag{2.41}$$

$$+ E_{\theta_0}\left|1 - \frac{1}{1+\int_{C^c}w(\theta)p_\theta(\overline{X})\,d\theta\Big/\int_C w(\theta)p_\theta(\overline{X})\,d\theta}\right| \tag{2.42}$$

$$+ E_{\theta_0}W(C^c|\overline{X}) + E_{\theta_0}N\big(C^c;\theta_0,\hat{\theta}\big). \tag{2.43}$$

By Theorem 2.1, expression (2.41) tends to zero. Since the quantity in absolute value bars in (2.42) is between zero and one, expression (2.42) will tend to zero if we show that

$$\frac{\int_{C^c}w(\theta)p_\theta(\overline{X})\,d\theta}{\int_C w(\theta)p(\overline{X}|\theta)\,d\theta} \to_{P_{\theta_0}} 0. \tag{2.44}$$

To prove (2.44) we first show that

$$P_{\theta_0}\Big(m_{C^c}(\overline{X})n^{(k+d+1/2)/2} > p_{\theta_0}(\overline{X})\Big) = o(1). \tag{2.45}$$

Set $C = \{\theta|\,\|(\mu(\theta) - \mu(\theta_0))\| \le \delta\}$. Now, the left-hand side of (2.45) is less than

$$\begin{aligned}
&P_{\theta_0}\left(\left|\overline{X} - \mu(\theta_0)\right| > \frac{\delta}{2}\right)\\
&\quad + P_{\theta_0}\left(\left|\overline{X} - \mu(\theta_0)\right| \le \frac{\delta}{2},\, m_{C^c}(\overline{X})n^{(k+d+1/2)/2} > p_{\theta_0}(\overline{X})\right)\\
&\quad \le \frac{K}{n} + n^{(k+d+1/2)/2}\int_{C^c} w_{C^c}(\theta)\sum_{|\bar{x}-\mu(\theta)|\ge\delta/2} P_\theta(\bar{x})\,d\theta\\
&\quad \le \frac{K}{n} + n^{(k+d+1/2)/2}\int_{K^c} w_{C^c}(\theta)\sum_{i=1}^{k} P_\theta\left(\left|\overline{X}_{(i)} - \mu_{(i)}(\theta)\right| > \frac{\delta}{2k}\right)d\theta\\
&\quad \le \frac{K}{n} + n^{(k+d+1/2)/2}\sup_{\theta,i} P_\theta\left(\left|\overline{X}_{(i)} - \mu_{(i)}(\theta)\right| > \frac{\delta}{2k}\right).
\end{aligned}$$

By Markov's inequality and a well known result bounding the moments of sums of independent random variables [see Ibragimov and Hasminskii (1981), page 186], we have that the last term is bounded by $Kn^{(k+d+1/2)/2}/n^{(k+d+1)/2}\sup_{i,\theta}E_\theta|X_{(i)} - \mu_{(i)}(\theta)|$. Now that (2.45) is established we use it to prove (2.44). Let $\varepsilon > 0$. By intersecting with the event in (2.45) and its complement we have

$$P_{\theta_0}\left(\left(\frac{\int_{C^c} w(\theta) p_\theta(\overline{X})\, d\theta}{p_{\theta_0}(\overline{X})}\right)\frac{p_{\theta_0}(\overline{X})}{\int_C w(\theta) p_\theta(\overline{X})\, d\theta} > \varepsilon\right)$$

$$\leq P_{\theta_0}\left(\frac{m_{C^c}(\overline{X})}{p_{\theta_0}(\overline{X})} > \frac{K}{n^{(k+d+1)/2}}\right) + P_{\theta_0}\left(\frac{p_{\theta_0}(\overline{X})}{m_C(\overline{X})} > Kn^{(k+d+1)/2}\right).$$

The first term is controlled by (2.45). Intersecting the second term with U^c and U gives two terms. The term with U^c goes to zero since $P_{\theta_0}(U^c) \to 0$. On the term with U, we use $m_C(\overline{X}) > 1/n^{(k+d)/2}$ and note $P_{\theta_0}(p_{\theta_0}(\overline{X}) > K\sqrt{n}\,) \to 0$.

The first term in (2.43) is bounded between zero and one, and dominated by the ratio in (2.44) which goes to zero. The other term in (2.43) is bounded by

$$KE_{\theta_0}\, \chi_{\{|\mu(\hat{\theta}) - \mu(\theta_0)| < \delta\}} \int_{C^c} n^{d/2} \exp\Big(-n(\theta - \hat{\theta})^t J_\mu^t(\theta_0)\Sigma^{-1}(\theta_0) J_\mu(\theta_0)(\theta - \hat{\theta})\Big)\, d\theta$$
$$+ KE_{\theta_0}\, \chi_{\{|\mu(\hat{\theta}) - \mu(\theta_0)| > \delta/2\}}.$$

The second term goes to zero by consistency of $\overline{X}$ for $\mu(\theta_0)$. The first term is the same as (2.33) and so goes to zero also. $\square$

If $d > k$, then there is a problem of identifiability for θ. If $d < k$ the desired result can be proved by centering at the estimator obtained in the following way. Let

$$\tilde{\theta} = \arg\min_{\theta'} \| \overline{X} - \mu(\theta') \|, \tag{2.46a}$$

where $\| \cdot \|$ is the Euclidean norm, and then set

$$\hat{\theta} = \arg\min_{\theta'} \| \overline{X} - \mu(\theta') \|_{\Sigma(\tilde{\theta})}, \tag{2.46b}$$

where the norm in (2.46b) is defined from the inner product induced by $\Sigma(\tilde{\theta})$. When $d = k$, $\hat{\theta}$ reduces to $\mu^{-1}(\overline{X})$. Our result is the following.

THEOREM 2.3. *Assume the hypotheses of Theorems* 2.1 *and* 2.2, *but that* $k > d$. *Then, if the assumptions in Proposition* 2.1 *hold for* $r \geq d$, (2.40) *holds for the estimator* $\hat{\theta}$ *in* (2.46b).

PROOF. We indicate how to modify the proof in the compact case for $d = 1$ and general $k > 1$; parts which are the same as before are ignored. Extensions to larger values of d can be established by straightforward modifications of this proof. The case of noncompact parameter spaces can be handled by using the technique of proof of Theorem 2.2.

Step 1, *part* 1: Note that by adding and subtracting $\mu(\hat{\theta})$ in the exponent we obtain

$$
\begin{aligned}
&m_r(\overline{X}) \\
(2.47)\quad &\geq \frac{k}{n^{k/2}} \int \exp\biggl(-\left(\frac{n}{2}\right) \| \overline{X} - \mu(\hat{\theta}) \|_{\theta}^{2} - \left(\frac{n}{2}\right) \| \mu(\hat{\theta}) - \mu(\theta) \|_{\theta}^{2} \\
&\qquad - \left(\frac{n}{2}\right) \bigl(\overline{X} - \mu(\hat{\theta})\bigr) \Sigma^{-1}(\theta) \bigl(\mu(\hat{\theta}) - \mu(\theta) \bigr) \biggr) \, d\theta,
\end{aligned}
$$

where $\|\cdot\|_\theta$ indicates the inner product with respect to $\Sigma(\theta)^{-1}$. On U we have that $\|\overline{X} - \mu(\hat{\theta})\|_{\theta_0} \leq k_n/\sqrt{n}$ and if we use the implicit function theorem we can assert the existence of a solution h to the equation $L(\hat{\theta}) = \Sigma(\overline{X}_i - \mu_i(\hat{\theta}))\mu_j(\hat{\theta})\sigma^{ij}(\theta_0) = 0$, where $\hat{\theta} = h(\overline{X})$, $\theta_0 = h(\mu(\theta_0))$ and $\sigma^{ij}(\theta_0)$ are the entries of $\Sigma^{-1}(\theta_0)$. As a result $\|\hat{\theta} - \theta_0\| \leq K\|\overline{X} - \mu(\theta_0)\|_{\theta_0} \leq K\sqrt{(\ln n)/n}$. If we cut the domain of integration down to $\|\theta - \hat{\theta}\| \leq k'_n/\sqrt{n}$, then by the triangle inequality $\|\theta - \theta_0\| \leq K\sqrt{(\ln n)/n}$. By Taylor expanding we then obtain that

$$
\Sigma^{-1}(\theta) \cong (1 + \varepsilon_n)\Sigma^{-1}(\theta_0), \tag{2.48}
$$

where $\varepsilon_n = O(\sqrt{(\ln n)/n})$ and $\cong$ means the left-hand side is bounded above and below by expressions of the form of the right-hand side.

Next we note that the third term in the exponent of (2.47) is negligible compared to the other two, at least when restricted to U: from (2.48) it is enough to examine $n(\overline{X} - \mu(\hat{\theta}))\Sigma(\theta_0)^{-1}(\mu(\hat{\theta}) - \mu(\theta))$. Taylor expanding μ at $\hat{\theta}$ and using $L(\hat{\theta}) = 0$ gives that the third term is $Kn(\theta - \hat{\theta})^2\|\overline{X} - \mu(\hat{\theta})\|_{\theta_0}$, which is seen to be $O((\ln n)/\sqrt{n})$. As a result we have on U that $m_r(\overline{X}) \geq K\exp(-(n/2)(1 + \varepsilon_n)\|\overline{X} - \mu(\hat{\theta})\|_{\theta_0}^2)/n^{(k+d)/2}$.

Step 2, *part* 1: We use the modified bound of Step 1, part 1, to obtain

$$
\begin{aligned}
\chi_U m(\overline{X}) \geq \frac{K\chi_U}{n^{(k+d)/2}} &\exp\left(-\left(\frac{n}{2}\right)(1 + \varepsilon_n) \| \overline{X} - \mu(\hat{\theta}) \|_{\theta_0}^{2} \right) \\
&\times \left(1 - \frac{(k'_n)^d \exp\bigl((n/2)(1 - \varepsilon_n) \| \overline{X} - \mu(\hat{\theta}) \|_{\theta_0}^{2}\bigr)}{n^{(r+1-d)/2}} \right).
\end{aligned}
$$

Since $n\|\overline{X} - \mu(\hat{\theta})\|_{\theta_0} \leq c^2 \ln n$, r can be chosen large enough to ensure that the second term in parentheses goes to zero. Indeed, it is enough for r to be greater than $(d - 1) + c^2(1 - \varepsilon_n)$. That is, if c is small enough, then $r \geq d$ will suffice.

Step 2, *part* 2: Expression (2.17) is no problem and it is seen that (2.18) goes to zero by noting that

$$\chi_U \int \frac{w(\theta)\left|p_\theta(\bar{X}) - q_{\theta r}(\bar{X})\right|}{m(\bar{X})}\,d\theta + \chi_U \int \frac{w(\theta) q_{\theta r}(\bar{X})}{m_r(\bar{X})}\,\frac{\left|m(\bar{X}) - m_r(\bar{X})\right|}{m(\bar{X})}\,d\theta$$

$$\leq K\chi_U\left(\frac{k_n^d}{n^{(k+r+d)/2}}\cdot\frac{n^{(k+d)/2}}{\exp\left(-(n/2)(1+\varepsilon_n)\left\|\bar{X} - \mu(\hat{\theta})\right\|_{\theta_0}^2\right)}\right.$$

$$\left.+\frac{k_n'^d}{n^{(k+r+1)}}\,\frac{n^{(k+d)/2}}{\exp\left(-(n/2)(1+\varepsilon_n)\left\|\bar{X} - \mu(\hat{\theta})\right\|_{\theta_0}^2\right)}\right),$$

which goes to zero for $r > (d-1) + c^2(1-\varepsilon_n)$, that is, $r \geq d$ for c small.

Step 3, *part* 2: Showing that analogs of (2.21) and (2.22) go to zero can be readily done. It is enough to show that

$$E_{\theta_0}\chi_U \int_{U'} q_{\theta r}(\bar{X})\,\frac{n^{(k+d)/2}}{\exp\left(-(n/2)(1+\varepsilon_n)\left\|\bar{X} - \mu(\hat{\theta})\right\|_{\theta_0}^2\right)}\,d\theta \to 0, \tag{2.49a}$$

$$E_{\theta_0}\chi_U \int_{U'} q_{\theta_0 r}(\bar{X})\,\frac{n^{(k+d)/2}}{\exp\left(-(n/2)\left\|\bar{X} - \mu(\hat{\theta})\right\|_{\theta_0}^2\right)}\,d\theta \to 0, \tag{2.49b}$$

since the analog to (2.24), $m_{\theta_0 r}(\bar{X}) \geq K\exp(-(n/2)\|\bar{X} - \mu(\hat{\theta})\|_{\theta_0}^2)/n^{(k+d)/2}$ can be derived by the same technique as in the modified Step 1, part 1.

Now, for both cases it is enough to note that on U, $n\|\bar{X} - \mu(\hat{\theta})\|_{\theta_0}^2 \leq c^2 \ln n$, and one obtains from the other part of either of the integrands bounds of the form $n^{-K(c'-c)^2}$. It is enough to choose $c' - c$ large enough.

Step 3, *part* 3: It is enough to show

$$\begin{aligned} Kn^{(k+d)/2}\Bigg(&2E_{\theta_0}\chi_U\int_{U'} w(\theta)\left|q_{\theta r}(\bar{X})q_{\theta\theta_0 r}(\bar{X})\right|d\theta \\ &\times\exp\left(\left(\frac{n}{2}\right)(1+\varepsilon_n)\left\|\bar{X} - \mu(\hat{\theta})\right\|_{\theta_0}^2\right) \\ &+E_{\theta_0}\chi_U\int_{U'^c} w(\theta)\left|q_{\theta r}(\bar{X}) - q_{\theta\theta_0 r}(\bar{X})\right|d\theta \\ &\times\exp\left(\left(\frac{n}{2}(1+\varepsilon_n)\left\|\bar{X} - \mu(\hat{\theta})\right\|_{\theta_0}^2\right)\right)\Bigg)\end{aligned} \tag{2.50}$$

goes to zero. By earlier reasoning in Step 3, part 3, (2.49a) and (2.49b) can be used to control the second term in (2.50). For the first term we observe that the extra exponential factor is bounded above by $\exp((1+\varepsilon_n)c^2 \ln n) \leq n^{(1+\varepsilon_n)c^2} \leq n^{3c^2/2} \leq n^{1/4}$, for n large enough and c small enough. The earlier proof of this part gave a bound of the form $K(\ln n)^{3r}/\sqrt{n}$ so the extra $n^{1/4}$ does not alter the convergence to zero.

Step 4, *part* 2: Use the result from the modified version of Step 3, part 2.

Step 4, *part* 4: It is enough to show that

$$\frac{K(k_n + k'_n)^{3(i-1)}}{n^{(i-1)/2}} \chi_U \frac{\int_{U'} w(\theta)\, \varphi_{\Sigma(\theta_0)}\big(\sqrt{n}\,(\bar{X} - \mu(\theta))\big)\, d\theta}{\int w(\theta)\, \varphi_{\Sigma(\theta_0)}\big(\sqrt{n}\,(\bar{X} - \mu(\theta))\big)\, d\theta}$$
$$+ \frac{Kn^{(k+d)/2}}{n^{(i-1)/2}} \chi_U \int_{U'^c} w(\theta) f_i\big(\sqrt{n}\,(\bar{X} - \mu(\theta))\big) \varphi_{\Sigma(\theta_0)}\big(\sqrt{n}\,(\bar{X} - \mu(\theta))\big)\, d\theta$$
$$\times \exp\left(\left(\frac{n}{2}\right) \|\bar{X} - \mu(\hat{\theta})\|_{\theta_0}\right)$$

goes to zero. This is obvious for the first term. Since $n\|\bar{X} - \mu(\hat{\theta})\|^2 \le c^2 \ln n$, the second term can be controlled by choosing c' large enough.

Step 4, *final part*: To control the analog of (2.29) we add and subtract

$$\frac{w(\theta)\exp\big(-(n/2)\|\bar{X} - \mu(\theta)\|^2_{\theta_0}\big)\exp\big((n/2)\|\bar{X} - \mu(\hat{\theta})\|^2_{\theta_0}\big)}{w(\theta_0)(2\pi)^{d/2}\big|nJ(\theta_0)^t \Sigma^{-1}(\theta_0) J(\theta_0)\big|^{-1/2}}$$

and

$$\frac{\exp\big(-(n/2)\|\bar{X} - \mu(\theta)\|^2_{\theta_0}\big)\exp\big((n/2)\|\bar{X} - \mu(\hat{\theta})\|^2_{\theta_0}\big)}{(2\pi)^{d/2}\big|nJ(\theta_0)^t \Sigma^{-1}(\theta_0) J(\theta_0)\big|^{-1/2}}$$

so that we must control

$$E_{\theta_0} \chi_U \left| 1 - \frac{e^{(n/2)\|\bar{X} - \mu(\hat{\theta})\|^2_{\theta_0}} \int w(\theta) e^{-(n/2)\|\bar{X} - \mu(\theta)\|^2_{\theta_0}}\, d\theta}{(2\pi)^{d/2} w(\theta_0) \big|nJ^t(\theta_0)\Sigma^{-1}(\theta_0) J(\theta_0)\big|^{-1/2}} \right| \tag{2.51}$$

$$+ E_{\theta_0} \chi_U \int_{U'} \left| \frac{w(\theta)}{w(\theta_0)} - 1 \right| \frac{e^{-(n/2)\|\bar{X} - \mu(\theta)\|^2_{\theta_0}} e^{(n/2)\|\bar{X} - \mu(\hat{\theta})\|^2_{\theta_0}}}{(2\pi)^{d/2} |nJ^t(\theta_0)\Sigma^{-1}(\theta_0) J(\theta_0)|^{-1/2}}\, d\theta \tag{2.52}$$

(2.53)
$$+ E_{\theta_0} \chi_U \int_{U'} \frac{e^{(n/2)\|\bar{X} - \mu(\hat{\theta})\|^2_{\theta_0}} e^{-(n/2)\|\bar{X} - \mu(\theta)\|^2_{\theta_0}} - e^{-(n/2)(\theta - \hat{\theta})^t J^t(\theta_0)\Sigma^{-1}(\theta_0) J(\theta_0)(\theta - \hat{\theta})}}{(2\pi)^{d/2} \big|nJ^t(\theta_0)\Sigma^{-1}(\theta_0) J(\theta_0)\big|^{-1/2}}\, d\theta,$$

the analogs of (2.36), (2.37) and (2.38). For (2.52) we use the fact that

$$\begin{aligned} n\|\bar{X} - \mu(\theta)\|^2_{\theta_0} &= n\|\bar{X} - \mu(\hat{\theta})\|^2 + n\|\mu(\hat{\theta}) - \mu(\theta)\|^2_{\theta_0} \\ &\quad + O\left(\frac{(\ln n)^\alpha}{\sqrt{n}}\right) \end{aligned} \tag{2.54}$$

so as to obtain the upper bound $o(1)E_{\theta_0}\chi_U\int_{U'}n^{d/2}\exp(-(n/2)\|\mu(\hat{\theta})-\mu(\theta)\|_{\theta_0}^2)\,d\theta$, which goes to zero since the integral gives a constant. [The modulus of continuity goes the $o(1)$.] For (2.53) we use (2.54) so as to reduce it to the analog of (2.29), as in the final step of Part 4 before. (Choose $c > 0$ small enough.)

For the last term (2.51), Laplace integration gives the desired convergence to zero, by use of (2.54) again. □

3. Nonidentically distributed random variables. To introduce our approximations, we require some notation. We denote the sum of the first n outcomes by $S^n(X) = \sum_{i=1}^n X_j$, with mean $\mu^n(\theta) = E_\theta S^n(X) = \sum_{j=1}^n \mu_j(\theta)$, where $\mu_j(\theta) = E_\theta X_j$. Analogously, we write $\Sigma_{j=1}^n(\theta) = \sum_{j=1}^n \Sigma_j(\theta)$, where $\Sigma_j(\theta) = \mathrm{Var}_\theta X_j$. The average mean is $\overline{\mu} = \overline{\mu}^n(\theta) = (1/n)\mu^n(\theta)$; the average variance is $\overline{\Sigma} = \overline{\Sigma}^n(\theta) = (1/n)\Sigma^n(\theta)$. We write $\overline{J}_{\mu,n}(\theta) = \nabla\overline{\mu}_n(\theta)$ to mean the $k \times d$ Jacobian matrix of first derivatives of $\overline{\mu}$. [The jth column is $((\partial/\partial\theta_j)\overline{\mu}_1(\theta), \dots, (\partial/\partial\theta_j)\overline{\mu}_k(\theta))$, where $\overline{\mu}_i$ is the ith component of $\overline{\mu}$.] To define the location of the limiting normal, we require the following.

DEFINITION 3.1. A sequence of functions $\langle f_n(\theta)\rangle|_{n=1}^\infty$ is locally invertible at θ_0 if and only if there is a neighborhood N_{θ_0} of θ_0 so that for all n, $f_n|_{N_{\theta_0}}\colon N_{\theta_0} \to f_n(N_{\theta_0})$ is invertible, for $\theta \in N_{\theta_0}^c$ we have that $f_n(\theta) \in f_n(N_{\theta_0})^c$ and the set $\bigcap_{n=1}^\infty f_n(N_{\theta_0})$ contains an open set around $\lim_{n\to\infty} f_n(\theta_0)$, assumed to exist.

Now, the target normal is

$$n(\theta;\theta_0,\hat{\theta}) = \frac{\left|n\overline{J}_{\mu,n}(\theta_0)\overline{\Sigma}^{-1}(\theta_0)\overline{J}_{\mu,n}(\theta_0)\right|^{1/2}}{(2\pi)^{d/2}} \times\exp\left(-\left(\frac{n}{2}\right)(\theta-\hat{\theta})\overline{J}_{\mu,n}(\theta_0)\overline{\Sigma}^{-1}(\theta_0)\overline{J}_{\mu,n}^{-1}(\theta-\hat{\theta})\right), \tag{3.1}$$

where $\hat{\theta} = (\overline{\mu}^n)^{-1}(\overline{X})$ near θ_0 since $\overline{\mu}^n$ is assumed to be locally invertible at θ_0. Note that $(\overline{\mu}^n)^{-1}$ is well defined only when k and d are equal.

As in Section 2, we continue to write

$$q_{\theta r}(\overline{X}) = \frac{l}{n^{k/2}}\sum_{i=1}^r \frac{f_i\left(\sqrt{n}\left(\overline{X}-\overline{\mu}^n(\theta)\right)\right)}{n^{(i-1)/2}}\varphi_{\Sigma(\theta)}\left(\sqrt{n}\left(\overline{X}-\overline{\mu}^n(\theta)\right)\right), \tag{3.2}$$

for the r term approximation to $p_\theta(\overline{X})$, where $f_1 \equiv 1$ and for $i \geq 2, f_i$ is a polynomial of degree at most $3r$ in k variables and $\varphi_{\overline{\Sigma}(\theta)}$ is the Normal$(0, \overline{\Sigma}(\theta))$ density. A variant on (3.2) is

$$q_{\theta\theta_0 rn}(\overline{X}) = \frac{l}{n^{k/2}}\sum_{i=1}^r \frac{f_i\left(\sqrt{n}\left(\overline{X}-\mu^n(\theta)\right)\right)}{n^{(i-1)/2}}\varphi_{\overline{\Sigma}(\theta_0)}\left(\sqrt{n}\left(\overline{X}-\mu^n(\theta)\right)\right) \tag{3.3}$$

in which the variance matrix is evaluated at θ_0. Mixtures of the densities in (3.2) and (3.3) with respect to θ are denoted $m_r(\overline{X})$ and $m_{\theta_0 r}(\overline{X})$, respectively.

Our first result is an inid version of Proposition 2.1; a proof is in the Appendix.

PROPOSITION 3.1. *Suppose the characteristic functions for the X_i's are jointly continuous in (t, θ) uniformly in i. Then,*

$$\sup_{\theta \in C} \sup_{\alpha \in L} \left(1 + \left\| \frac{\alpha - \overline{\mu}^n(\theta)}{\sqrt{n}} \right\|^{r+1}\right) \left| p_\theta\left(\frac{\alpha}{n}\right) - q_{\theta r}\left(\frac{\alpha}{n}\right) \right| = O\left(\frac{1}{n^{(k+r)/2}}\right). \tag{3.4}$$

Now write $U_n = \{X^n | \, \|\overline{\mu}_n(\hat{\theta}) - \overline{\mu}_n(\theta_0)\| \le k_n/\sqrt{n}\}$ and $U_n' = \{\theta | \, \|\overline{\mu}_n(\theta) - \overline{\mu}_n(\theta_0)\| \le k_n'/\sqrt{n}\}$, where $k_n/\sqrt{n}, k_n'\sqrt{n} \to 0$ and $\|\cdot\|$ is a norm on L embedded in $\mathbb{R}^k$. To permit Taylor expansions we make the following definition.

DEFINITION 3.2. A sequence of functions $\langle g_n(\theta) \rangle|_{n=1}^\infty$ is uniformly Taylor expandable at θ_0 if and only if (1) each g_n is continuously differentiable on an open set N_{θ_0} containing θ_0; (2) there are $\alpha, \beta > 0$ so that for all n and all $\theta \in N_{\theta_0}$, $\beta > \|\nabla g_n(\theta)\| > \alpha$, where ∇g_n is the Jacobian matrix of first derivatives of the components of g_n with respect to the components of θ; (3) on N_{θ_0}, ∇g_n has maximal rank.

The defining conditions in U_n and U_n' can be expressed as $\|\hat{\theta} - \theta_0\| \le k_n/\alpha\sqrt{n}$ and $\|\theta - \theta_0\| \le k_n'/\alpha\sqrt{n}$. We set $k_n = c\sqrt{\ln n}$ and $k_n' = c'\sqrt{\ln n}$, where $c', c > 0$ and $c' - c > 0$.

THEOREM 3.1. *Assume the hypotheses of Proposition* 3.1 *are satisfied with* $r + 1$, *where* $r > \max(0, d/2) - 1, (2/3)d - 4/3)$, *and that w is positive at θ_0. Assume also that $\langle \overline{\mu}_n(\theta) \rangle|_{n=1}^\infty$, $\langle \overline{\Sigma}^{-1}(\theta) \rangle|_{n=1}^\infty$ and $\langle \overline{J}_{\mu,n}^t(\theta)\Sigma^{-1}(\theta_0)\overline{J}_{\mu,n}(\theta) \rangle$ are uniformly Taylor expandable and that $\langle \overline{\mu}_n(\theta) \rangle|_{n=1}^\infty$ is locally invertible at θ_0. Finally, suppose there is a neighborhood N_{θ_0} of θ_0 and $\alpha, \beta > 0$ so that for $\theta, \theta' \in N_{\theta_0}$ we have that*

$$\beta I_d \ge \overline{J}_{\mu,n}(\theta)\overline{\Sigma}^{-1}(\theta')J_{\mu,n}(\theta) \ge \alpha I_d, \tag{3.5}$$

uniformly in n. Then if Ω is compact we have that

$$E_{\theta_0} \int_\Omega \left| w(\theta | \overline{X}) - n\left(\theta; \theta_0, \hat{\theta}\right) \right| d\theta \to 0, \tag{3.6}$$

as $n \to \infty$, where $n(\theta; \theta_0, \hat{\theta})$ is as in (3.1).

PROOF. In reviewing the proof of Theorem 2.1, it can be seen that most of the steps go through with only cosmetic changes. For instance, we use the inid forms of $m_r(\overline{X})$ and $m_{\theta_0 r}$ as defined in this section rather than their iid analogs. Also, we replace $\mu(\theta)$, $\Sigma(\theta)$ and $J_\mu(\theta)$ by $\overline{\mu}_n(\theta)$, $\overline{\Sigma}(\theta)$ and $\overline{J}_{\mu,n}(\theta)$.

There are, however, steps where the modifications are not solely a matter of notation. They are Step 1, part 1, Step 3, part 3 and Step 4, parts 3 and 5. It will be seen that they follow largely by the uniform Taylor expandability and local invertibility assumptions on sequences of functions.

For Step 1, part 1, (3.5) ensures that the last inequality in proving the extension to (3.6) in Section 2 continues to hold. Part 2 relies on the properties of (3.2), U_n and U_n' as before. The product $|f_i(\sqrt{n}(\overline{X} - \overline{\mu}_n(\theta)))|\varphi_{\overline{\Sigma}(\theta)}(\sqrt{n}(\overline{X} - \overline{\mu}_n(\theta)))$ remains bounded by a constant, for n large enough and θ in a compact set. Part 3 only requires cosmetic changes.

Step 2 continues to hold, subject to cosmetic changes, once Step 1 is extended. Part 1 is obvious. Part 2 only requires that one observes $P_{\theta_0}(U^c)$ tends to zero.

Step 3 uses the assumptions on $\langle \overline{\Sigma}^{-1}(\theta)\rangle|_{n=1}^{\infty}$. Part 1 is unchanged and part 2 follows by the same tchniques as before. The main difference occurs in Part 3: the uniform Taylor expandability of $\langle \overline{\Sigma}^{-1}(\theta)\rangle|_{n=1}^{\infty}$ gives the appropriate analog of (2.20).

Step 4 requires a bit more. While parts 1 and 2 continue to hold, part 3 requires the local invertibility and uniform Taylor expandability of $\langle \overline{\mu}_n(\theta)\rangle|_{n=1}^{\infty}$ to ensure the inid analog of (2.31) goes to zero by straightforward modifications of the earlier technique. Part 4 is again cosmetic. Part 5, the last one, requires that the Laplace integration in (2.36) and the bounding of the difference in the exponents in (2.38) be generalized. The latter is covered by the uniform Taylor expandability of $\langle \overline{J}_{\mu,n}^{t}(\theta)\overline{\Sigma}^{-1}(\theta_0)\overline{J}_{\mu,n}(\theta)\rangle|_{n=1}^{\infty}$. The former follows as before. [One observes that (3.5) controls the analog to (2.37).] So, the earlier proof has been adapted to give a proof of Theorem 3.1. □

It is of interest to generalize one step further so as to obtain a result in the case of noncompact parameter spaces. Our technique of proof will be to reduce the result to the compact case. Thus we define two mixtures, one over a compact set C, the other over its complement. They are

$$m_C(\overline{X}) = \int_C \frac{w(\theta)}{W(C)} p_\theta(\overline{X})\,d\theta \quad \text{and} \quad m_{C^c}(\overline{X}) = \int_{C^c} \frac{w(\theta)}{W(C^c)} p_\theta(\overline{X})\,d\theta,$$

where W is the probability with density w. Our result is Theorem 3.2.

THEOREM 3.2. *Assume the hypotheses of Theorem* 3.1. *In addition, assume that for some* $\delta > 0$, $\bigcap_{n=1}^{\infty} \overline{\mu}_n^{-1}(\beta(\mu(\theta_0), \delta))$ *contains a nonvoid open set around* θ_0 *and that for each* X_i *and for all components* $X_{i(j)}$, $j = 1, \dots, k$, *of* X_i *the central moments of order* $k + d + 1$ *are uniformly bounded in* θ, *that is*, $\sup_{\theta, i} \sup_{j=1}^{k} E_\theta|X_{i(j)} - \mu_{i(j)}(\theta)| < \infty$. *Then we have that* $E_{\theta_0}\int|w(\theta|\overline{X}) - n(\theta; \theta_0, \hat{\theta})|\,d\theta \to 0$, *where* $n(\theta; \theta_0, \hat{\theta})$ *is as in* (3.1).

PROOF. The structure and techniques of the proof of Theorem 2.2 continue to be valid. It is enough to deal with the inid analogs of (2.37), (2.38) and (2.39) in Section 2. The inid analog of expression (2.37) goes to zero by

Theorem 3.1. The remaining analogous quantities (2.38) and (2.39) go to zero provided that

$$\frac{\int_{C^c} w(\theta) p_\theta(\overline{X})\, d\theta}{\int_C w(\theta) p_\theta(\overline{X})\, d\theta} \to_{P_{\theta_0}} 0, \tag{3.7}$$

in the inid case. Again, it is enough to show that

$$P_{\theta_0}\Big(m_{C^c}(\overline{X}) n^{(k+d+(1/2))/2} > p_{\theta_0}(\overline{X})\Big) = o(1). \tag{3.8}$$

We can then multiply and divide the left-hand side of (3.7) by p_{θ_0} and intersect with the event in (3.8) and its complement, as in the proof of Theorem 2.2.

Choose C to be compact with nonvoid interior, contained in $\bigcap_{n=1}^{\infty}\{\theta: \|\overline{\mu}_n(\theta) - \overline{\mu}_n(\theta_0)\| < \delta\}$. On C we have that $\|\overline{\mu}_n(\theta) - \overline{\mu}(\theta_0)\| < \delta$, so we may upper bound (3.8) by

$$\begin{aligned} &P_{\theta_0}\Big(\big\|\overline{X} - \overline{\mu}_n(\theta_0)\big\| > \delta/2\Big) \\ &\quad + P_{\theta_0}\Big(\big\|\overline{X} - \overline{\mu}_n(\theta_0)\big\| < \delta/2, \\ &\qquad\qquad m_{C^c}(\overline{X}) \exp(n^{(k+d+(1/2)/2)}) \geq p_{\theta_0}(\overline{X})\Big). \end{aligned} \tag{3.9}$$

The first term in (3.9) is of $O(1/n)$. The second term tends to zero by the same technique as in the proof of Theorem 2.2. □

We remark that under a somewhat messy list of assumptions these results can be extended to the case that $d < k$.

4. Implications for testing independence of test items. We use Proposition 3.1 and Theorem 3.1 to obtain a result which has implications for educational testing. We give conditions under which educational tests have a property called asymptotic covariance given the sum is negative (ACSN). ACSN is a variant of covariance given the sum is negative (CSN) used by Junker (1993). Both ACSN and CSN express the idea that conditional on an examinee's score, the examinee's performance on different test questions should be uncorrelated. Specifically, in Junker (1993), the CSN condition $\mathrm{Cov}(X_i, X_j|\overline{X}) \leq 0$, for $i \neq j$, is studied as a verifiable condition that can be used to imply unidimensionality and local asymptotic discrimination—two main hypotheses of educational testing.

ACSN is useful for two reasons. The first is that one can base a test of the independence of items i and j on the convergence of $\mathrm{Cov}(X_i, X_j|\overline{X})$ to a nonpositive number. The other is that it can be used to obtain a partial converse to a characterization result for tests which satisfy strict unidimensionality and are locally asymptotically discriminating; for definitions, see

Junker (1993). Stating what exactly the test is and proving the characterization are of a specialized nature which we do elsewhere.

We begin with a lemma to control the difference between $p_\theta(X_i|S^n)$ and $p_\theta(X_i)$. In the proof we use Proposition 3.1 for the density of S^n and for the density of $S^n - X_i$. We denote their one-term normal approximations by $q = q_{\theta_n}$ and $q_i = q_{\theta n_i}$. For brevity we write $\overline{\Sigma}_i = (1/(n-1))\Sigma_{j \neq i}\Sigma_j(\theta)$. In addition, we assume that the X_i's take values in a finite range, that their variances are uniformly bounded above and below by constant multiples of the $d \times d$ identity matrix and that the set U_n is reexpressed as $U_{n,s}(\theta) = \{s\colon \sqrt{n}\,|s/n - \overline{\mu}(\theta)| \le c\sqrt{\ln n}\}$. Letting x denote a fixed value of X_i we have the following.

LEMMA 4.1. *Assume the hypotheses of Proposition* 3.1 *hold on a compact set C for* $r = 1$. *Then there is an* $\xi > 0$ *so that*

$$\sup_{\theta \in C} \sup_{s} \left| \frac{P_\theta(S^n - X_i = s - x)}{P_\theta(S_n = s)} - 1 \right| \chi_{U_{n,s}(\theta)} = O\left(\frac{1}{n^{\xi}}\right). \tag{4.1}$$

PROOF. By Proposition 3.1 we have that $P_\theta(S_n - X_i = s - x) = q_i + T_i$ and $P_\theta(S_n = s) = q + T$, where the T_i and T are error terms from the $r = 1$ term normal approximation satisfying $\sup_{s-x}|T_i|, \sup_s|T| = O(1/n^{(k+1)/2})$, uniformly for $\theta \in C$.

Now, consider the left-hand side of (4.1) for fixed θ. Add and subtract q_i/q and use the triangle inequality to obtain the upper bound

$$\left| \frac{q_i + T_i}{q + T} - \frac{q_i}{q} \right| \chi_{U_{n,s}(\theta)} + \left| \frac{q_i}{q} - 1 \right| \chi_{U_{n,s}(\theta)}. \tag{4.2a, b}$$

Apart from $\chi U_{n,s(\theta)}$, expression (4.2a) is, after adding and subtracting qT, bounded from above by

$$\begin{aligned} \left| \frac{(q_i + T_i)q - q_i(q + T)}{q(q + T)} \right| &= O\left(\frac{1}{n^{(k+1)/2}}\right) \frac{1}{q + T} \\ &\quad + O\left(\frac{1}{n^{(k+1)/2}}\right) \frac{(q_i - q)}{q(q + T)}. \end{aligned} \tag{4.3}$$

On $U_{n,s}(\theta)$ we have that there is an $\varepsilon > 0$ so that $q, q + T \ge O(1/n^{(k+\varepsilon)/2})$. In fact, ε may be chosen as small as desired by using small enough c in the definition of $U_{n,s}$. Now, we upper bound (4.3) by $O(n^{(\varepsilon-1)/2}) + O(n^{(k-1+2\varepsilon)/2})|q_i - q|$. Apart from $\chi_{U_{n,s(\theta)}}$ expression (4.2b) is $|q_i - q|/q$, which is bounded from above by $O(n^{(k+\varepsilon)/2})|q_i - q|$. So, (4.2) is bounded from above by the sum of the last two upper bounds. Since $|q_i - q|$, appears in both of these bounds, we bound it. Let $m_i = \sum_{j \neq i} \mu_j(\theta)$. By adding and subtracting

$$|\overline{\Sigma}|^{-1/2} \exp\left(-(n-1)\left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \overline{\Sigma}_i^{-1} \left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \right),$$

we have that

$$\begin{aligned} |q - q_i| \le {} & \frac{K}{n^{k/2}} |\overline{\Sigma}|^{-1/2} \left| \exp\left(-n\left(\frac{s}{n} - \overline{\mu}^n(\theta) \right) \overline{\Sigma}^{-1} \left(\frac{s}{n} - \overline{\mu}^n(\theta) \right) \right) \right. \\ & \qquad \left. - \exp\left(-(n-1)\left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \overline{\Sigma}_i^{-1} \left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \right) \right| \\ & + \frac{K}{n^{k/2}} \left| \, |\overline{\Sigma}|^{-1/2} - |\overline{\Sigma}_i|^{-1/2} \right| \\ & \times \left| \exp\left(-(n-1)\left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \overline{\Sigma}_i^{-1} \left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \right) \right|. \end{aligned} \tag{4.4}$$

Expression (4.4) is clearly $O(n^{-k/2})$, so $|q_i - q| O(n^{(k-1+2\varepsilon)/2})$ tends to zero, provided that we choose ε small enough. It remains to show that $|q_i - q| O(n^{(k+\varepsilon)/2})$ goes to zero. This requires that we obtain a faster rate of convergence to zero for (4.4).

First note that $\| \overline{\Sigma}|^{-1/2} - |\overline{\Sigma}_i|^{-1/2}| = O(1/n)$. This follows by noting that $|\overline{\Sigma}|$ and $|\overline{\Sigma}_i|$ are controlled by the hypotheses on the variances of the X_i's. Indeed, take a common denominator, multiply and divide by $|\overline{\Sigma}|^{1/2} + |\overline{\Sigma}_i|^{1/2}$, bound the denominator from below and remove and bound the common factor $|\overline{\Sigma}_i|$ to obtain the bound $K \| \overline{\Sigma}_i^{-1} \overline{\Sigma}| - 1|$. Apply the identity $\overline{\Sigma} = ((n-1)/n)\overline{\Sigma}_i + (1/n)\Sigma_i$, add and subtract $|((n-1)/n) I_d|$ and use the triangle inequality. One term is $O(1/n)$ immediately; the other term is seen to be $O\ (1/n)$ by Taylor expanding the determinant function at the identity.

Now if we use (4.4) to bound $|q - q_i|$, we can note that $e^{-x} \le 1$, for $x \ge 0$ so that one of the resulting terms goes to zero at rate $O(1/n^{1-\varepsilon/2})$. The other term is bounded above (on $U_{n,s}$) by

$$\begin{aligned} O(n^{\varepsilon/2}) & \left| \exp\left(-n\left(\frac{s}{n} - \overline{\mu}^n(\theta) \right) \Sigma^{-1} \left(\frac{s}{n} - \overline{\mu}^n(\theta) \right) \right) \right. \\ & \qquad \left. - \exp\left(-(n-1)\left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \Sigma_i^{-1} \left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \right) \right|. \end{aligned}$$

Since ε can be made arbitrarily small and $|\overline{\Sigma}|^{-1/2}$ is bounded by assumption, we can use the fact that $|e^{-x} - e^{-y}| \le |x - y|$ to see that, on $U_{n,s}(\theta)$, it is enough to show

$$\begin{aligned} & \left| n\left(\frac{s}{n} - \overline{\mu}^n(\theta) \right) \overline{\Sigma}^{-1} \left(\frac{s}{n} - \overline{\mu}^n(\theta) \right) \right. \\ & \qquad \left. - (n-1)\left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \overline{\Sigma}_i^{-1} \left(\frac{s-x}{n-1} - \frac{m_i}{n-1} \right) \right| = o\left(\frac{1}{n^{\xi}} \right), \end{aligned} \tag{4.5}$$

for some $\xi > 0$. By straightforward but tedious manipulations we have that, on $U_{n,s}(\theta)$,

$$O\left(\frac{\ln n}{n}\right) + |\mu_i - x| O\left(\frac{\ln n}{n}\right) + |\mu_i - x|^2 O\left(\frac{1}{n}\right) + O(\ln n)$$
$$(4.6)\qquad + \frac{|\mu_i - x|^2}{n} \|\overline{\Sigma}^{-1} - \overline{\Sigma}_i^{-1}\| + O\left(\frac{\ln n}{n} + \frac{|\mu_i - x|^2}{n^2}\right)$$

is an upper bound on the right-hand side of (4.5). [Derive that $|(s-x)/(n-1) - m_i/(n-1)|^2$ is less than $K\ ((\ln n)/n + |\mu_i - x|^2/n^2)$ on $U_{n,s}(\theta)$.] The matrix norm in the fifth term in (4.6) is seen to be $O(1/n)$. Consequently, rearranging gives

$$O\left(\frac{\ln n}{n}\right) + |\mu_i - x| O\left(\sqrt{\frac{\ln n}{n}}\right) + |\mu_i - x|^2 O\left(\frac{1}{n}\right)$$
$$(4.7)\qquad + (\mu_i - x)^2 O\left(\frac{1}{n^2}\right),$$

as an upper bound on the left-hand side of (4.5), on $U_{n,s}(\theta)$. Now, expression (4.5) holds. The uniformity over C is clear. □

We use the technical result in Lemma 4.1 to prove Proposition 4.1, to see that $E(X_i|\overline{X}, \theta)$ is close to the full expectation when the X_i's assume finitely many values.

PROPOSITION 4.1. *Assume the hypotheses of Lemma* 4.1. *Let $\overline{X} = S^n/n$ be an element of $U_{n,s}(\theta)$. Then there is an $\eta > 0$ so that as n increases*,

$$(4.8)\qquad \sup_{\theta \in C} \sup_{S^n} \left| E\left(X_i|\overline{X}, \theta\right) - E\left(X_i|\theta\right)\right| \chi_{U_n(\theta)} = O\left(\frac{1}{n^\eta}\right).$$

PROOF. Note that the left-hand side of (4.8) is

$$\left|\Sigma x_i P_\theta\left(x_i|S^n\right) - \Sigma x_i P_\theta\left(X_i\right)\right|$$
$$(4.9)\qquad \le \Sigma x_i P\left(x_i|\theta\right)\left|\frac{P_\theta\left(S^n - X_i = s - x_i\right)}{P_\theta\left(S^n = s\right)} - 1\right|.$$

For each of the finitely many values x_i, the quantity in absolute value bars on the right-hand side of (4.9) is controlled by the Lemma 4.1, so the proposition is proved. □

Finally, we state the main result of this section.

THEOREM 4.1. *If the hypotheses of Proposition* 4.1 *and Theorem* 3.2 *are satisfied*, *then*

$$(4.10)\qquad \operatorname{Cov}\left(E\left(X_i|\overline{X}, \theta\right), E\left(X_j|\overline{X}, \theta\right)|\overline{X}\right) \to_{P_{\theta_0}} 0.$$

PROOF. Note that

$$\mathrm{Cov}\Big(E\big(X_i|\overline{X},\theta\big),\,E\big(X_i|\overline{X},\theta\big)|\overline{X}\Big)$$

$$= \int_{B(\theta_0,\,\varepsilon)} \chi_U E\big(X_i|\theta,\overline{X}\big)E\big(X_i|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta \tag{4.11a}$$

$$+ \int_{B(\theta_0,\,\varepsilon)^c} \chi_U E\big(X_i|\theta,\overline{X}\big)E\big(X_j|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta \tag{4.11b}$$

$$+ \int \chi_{U^c} E\big(X_i|\theta,\overline{X}\big)E\big(X_j|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta \tag{4.11c}$$

$$-\bigg(\int_{B(\theta_0,\,\varepsilon)} \chi_U E\big(X_i|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta \tag{4.12a}$$

$$+ \int_{B(\theta_0,\,\varepsilon)^c} \chi_U E\big(X_i|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta \tag{4.12b}$$

$$+ \int \chi_{U^c} E\big(X_i|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta\bigg) \tag{4.12c}$$

$$\times\bigg(\int_{B(\theta_0,\,\varepsilon)} \chi_U E\big(X_j|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta \tag{4.13a}$$

$$+ \int_{B(\theta_0,\,\varepsilon)^c} \chi_U E\big(X_i|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta \tag{4.13b}$$

$$+ \int \chi_{U^c} E\big(X_j|\theta,\overline{X}\big)w(\theta|\overline{X})\,d\theta\bigg). \tag{4.13c}$$

For terms (4.11a), (4.12a) and (4.13a) we use Proposition 4.1 to approximate the integrands with vanishing error. For terms (4.11b), (4.12b) and (4.13b) we use the fact that X_i, X_j and χ_U are bounded. Thus their conditional expectations are bounded so the concentration of the posterior forces them to zero.

It remains to deal with terms (4.11c), (4.12c) and (4.13c). We use the local invertibility of $\overline{\mu}^n(\theta)$: Since $\overline{\Sigma}(\theta)$ is bounded above and below, we have that there is a M' so that on U^c, $\|\theta - \hat{\theta}\| \geq M'\sqrt{(\ln n)/n}$. Also, by the central limit theorem we have that, under θ_0, the probability of the set $\|\hat{\theta} - \theta_0\| \leq M\sqrt{(\ln n)/n}$ tends to unity for any $M > 0$. By the boundedness of the integrands and the fact that the inequalities go in opposite directions we can control (4.11c), (4.12c) and (4.13c).

For instance, (4.13c) is controlled in L^1 by

$$E_{\theta_0}\int \chi_{U^c}\Big|E\big(X_j|\theta,\overline{X}\big)\Big|w(\theta|\overline{X})\,d\theta$$

$$\leq K\bigg(E_{\theta_0}\,\chi_{\{\|\hat{\theta}-\theta_0\|\leq M\sqrt{(\ln n)/n}\}}\int \chi_{U^c} w(\theta|\overline{X})\,d\theta$$

$$+E_{\theta_0}\,\chi_{\{\|\hat{\theta}-\theta_0\|>M\sqrt{(\ln n)/n}\}}\int \chi_{U^c} w(\theta|\overline{X})\,d\theta\Bigg)$$

$$\leq K\left(E_{\theta_0}\,\chi_{\{\|\hat{\theta}-\theta_0\|\leq M\sqrt{(\ln n)/n}\}}\int \chi_{\{\|\theta-\theta_0\|\geq (M-M')\sqrt{(\ln n)/n}\}} w(\theta|\overline{X})\,d\theta\right)+o(1)$$

in which the integral in the last expression goes to zero by the L^1 asymptotic normality of the posterior, provided $M - M'$ is large enough. Thus (4.13c) goes to zero in P_{θ_0} probability. Terms (4.11c) and (4.12c) are similar. □

COROLLARY TO THEOREM 4.1. *Assume the hypotheses of Theorem* 3.1. *If, in addition, the densities of the X_i's are log-concave, then we have, for any fixed θ_0 and $\varepsilon > 0$, that*

$$P_{\theta_0}\Big(\operatorname{Cov}\big(X_i, X_j|\overline{X}\big) \geq \varepsilon\Big) \to 0. \tag{4.14}$$

PROOF. By Junker's identity, see Junker [(1993), Section 4] we have that

$$\begin{aligned}\operatorname{Cov}\big(X_i, X_j|\overline{X}\big) &= E\Big(\operatorname{Cov}\big(X_i, X_j|\overline{X}, \theta\big)|\overline{X}\Big)\\ &\quad+\operatorname{Cov}\Big(E\big(X_i|\overline{X},\theta\big), E\big(X_j|\overline{X},\theta\big)|\overline{X}\Big).\end{aligned}\tag{4.15a, b}$$

By Theorem 32.8 in Joag-dev and Proschan (1983) [see also Theorem 4.1 in Junker (1993)], (4.15a) is nonpositive. By Theorem 4.1, expression (4.15b) converges to zero in P_{θ_0}-probability. Thus, (4.14) follows. □

Appendix. To obtain Proposition 3.1, we use characteristic function (cf) arguments. Write the cf of X_j as $\phi_j(\theta, t) = E_\theta \exp(i(t, X_j))$. Since the X_j's take values in a common lattice, these cf's have a common fundamental domain $\mathscr{F}^*$. Central to the statement and proof of the result is a proper subset E_1 of $\mathscr{F}^*$, defined by $E_1 = \{t \in \mathbb{R}^k\colon \|t\| \leq \xi\}$, where ξ is a constant. Let C be a compact set in the parameter space. We require that ξ satisfies the following:

ASSUMPTION 1. (i) On $\sqrt{n}\,E_1$ we can use the expansion given in Theorem 9.9 of BR modified in the same way as Theorem 9.12 of BR.

(ii) For $t \in \sqrt{n}\,E_1$ we have that

$$\sup_{j=1}^{n}\left|\phi_j\left(\theta, \frac{\overline{\Sigma}^n(\theta)^{-1/2}}{\sqrt{n}}\right)-1\right| \leq 1/2.$$

(iii) For $\delta(\theta) = \sup_{j\in\mathbb{N}} \sup_{t\in\mathscr{F}^*-E_1(\theta)}|\phi_j(\theta, t)|$, we have that $\delta_C = \sup_{\theta\in C}\delta(\theta) < 1$.

ASSUMPTION 2. For $r \geq 1$ suppose that on C,

$$g(\theta) = sup_n (1/n)\Sigma_{j=1}^n E_\theta \|X_j\|^{r+2}$$

exists and is bounded.

ASSUMPTION 3. We have that, on C, $\eta_1 I_d \leq \overline{\Sigma}(\theta) \leq \eta_2 I_d$ for some $\eta_1, \eta_2 > 0$ and all n.

Assumptions 1–3 hold if the ϕ_j's are jointly continuous in (t, θ), uniformly in j. This is the case in the Rasch model and in the generalisation of that model considered by Tsutakawa and Johnson (1990) and Tsutakawa and Soltys (1990). More generally, suppose X_i is distributed according to a probability function $p(x_i, \theta, \alpha_i)$, where the dependence on i is only in the third argument. Then, Assumptions 1–3 hold if (i) $p(x_i, \theta, \alpha)$ is a continuous function of (θ, α), which ranges over a fixed compact set; (ii) the moments $E_{(\theta,\alpha)}\|X\|^{r+2}$ are continuous and finite for (θ, α) in the compact set; and (iii) for some positive constants η_1 and η_2, the variance matrix $\Sigma(\theta, \alpha)$ is continuous and satisfies $\eta_1 I_d \leq \Sigma(\theta, \alpha) \leq \eta_2 I_d$ on the compact set. For generality, we use Assumptions 1–3.

First we show that

$$\sup_{\theta\in K}\sup_{\alpha\in L}\left(1+\left\|\frac{\alpha-\overline{\mu}^n(\theta)}{\sqrt{n}}\right\|^{r+1}\right)\left|p_\theta\left(\frac{\alpha}{n}\right)-q_{\theta r}\left(\frac{\alpha}{n}\right)\right| = o\left(\frac{1}{n^{(k+r-1)/2}}\right).$$

The cf of $S^n(X)$ is $\tilde{\phi}^n(\theta,t) = \Pi_{j=1}^n \phi_j(\theta,t)$ and the cf of $Y_n = (S^n(X) - \mu^n(\theta))/\sqrt{n}$ is $\phi^n(\theta,t) = \tilde{\phi}^n(\theta, t/\sqrt{n})\exp(-i(t/\sqrt{n}, \mu^n(\theta)))$. By the inversion formula we have $P(S^n(X) = \xi) = (l/(2\pi)^k)\int_{\mathscr{F}^*}\tilde{\phi}(\theta,t)e^{-i(t,\xi)}\,dt$. Using $t = t'/\sqrt{n}$ we obtain

$$P_\theta(S^n(X) = \xi) = \frac{l}{(2\pi)^k}\frac{1}{n^{k/2}}\int_{\sqrt{n}\mathscr{F}^*}\phi(\theta,t)\exp\left(-i\left(t, \frac{\xi-\mu^n(\theta)}{\sqrt{n}}\right)\right)dt,$$

from which we see for $Y_n = (1/\sqrt{n})(S^n(X) - n\overline{\mu}^n)$ that

$$Y_{n\xi}^\beta P_\theta(Y_{n\xi}) = \frac{l(-i)^{|\beta|}}{(2\pi)^k n^{k/2}}\int_{\sqrt{n}\mathscr{F}^*}\left[D^\beta\phi^n(\theta,t)\right]\exp\left(-i(t, Y_{n\xi})\right)dt,$$

for vectors $\beta = (\beta_1, \dots, \beta_k)$, where $\beta_i \geq 0$ are integers summing to $|\beta| \leq r + 2$ and D^β denotes the differentiation operator $(D_{t_1})^{\beta_1}, \dots, (D_{t_k})^{\beta_k}$. Vectors raised to powers β mean that each entry in the vector is raised to the corresponding entry in β.

Denote the Fourier transform of $q_{\theta r}(\overline{X})$ by $\tilde{q}_{\theta r}(t)$. Then,

$$Y_{n\xi}^\beta q_{\theta r}(Y_{n\xi}) = \frac{l(-i)^{|\beta|}}{(2\pi)^k n^{k/2}}\int_{\mathbb{R}^k}\left[D^\beta\tilde{q}_{\theta r}(t)\right]\exp\left(-i(t, Y_{n\xi})\right)dt,$$

where $\tilde{q}_{\theta r}(t) = \sum_{j=1}^{r} n^{-(j-1)/2}\tilde{P}_j(it\colon \{\chi_\nu\})e^{-\|t\|^2/2}$. The $\tilde{P}_j(it\colon \{\chi_\nu\})$'s are polynomials with coefficients depending on cumulants χ_ν. Now we have the upper bound

$$\begin{aligned}
&\left|Y_{n\xi}^{\beta}\left(p_\theta(Y_{n\xi}) - q_{\theta r}(Y_{n\xi})\right)\right| \\
\text{(A.1)}\qquad &\leq \frac{K}{n^{k/2}}\left[\int_{\sqrt{n}E_1}\left|D^{\beta}(\phi^n(\theta,t) - \tilde{q}_{\theta r}(t))\right|dt\right. \\
&\qquad\left. + \int_{\sqrt{n}\mathscr{F}^* - \sqrt{n}E_1}\left|D^{\beta}\phi^n(\theta,t)\right|dt + \int_{\mathbb{R}^k - \sqrt{n}E_1}\left|D^{\beta}\tilde{q}_{\theta r}(t)\right|dt\right].
\end{aligned}$$

Since the domain of integration excludes a ball with radius increasing as $\sqrt{n}$, the presence of the exponential factor implies that the last integral tends to zero at rate $O(e^{-nr'})$ for some $r' \geq 0$. The middle integral tends to zero at an exponential rate also: After differentiating $\phi^n(\theta,t)$ and observing that the exponential factor has norm 1, one can transform back to $\mathscr{F}^* - E_1$. The product $\tilde{\phi}(\theta,t)$ can be bounded from above by $O(\delta_K^n)$, in which $\delta_K < 1$.

The first integral in (A.1) requires Theorem 9.12 in BR, which is based on Theorems 9.9 and 9.10, also in BR. Examination of the proofs of those theorems shows that our assumptions give an upper bound for the integral of order $o(1/n^{r/2})$ uniformly in θ. Now (A.1) gives (3.4) by the same triangle inequality argument as was used in the proof of Proposition 2.1. $\square$

We can dispense with Assumption 1(i) by making use of the other assumptions with Theorem 9.11 (modified as in Theorem 9.12 in BR) and Lemma 14.3 in BR.

REFERENCES

ACKERMAN, T. A. (1991). An examination of the effect of multidimensionality on parallel forms construction. Presented at the Annual Meeting of the National Council on Measurement in Education, Chicago.

BARTHOLOMEW, D. J. (1987). *Latent Variable Models and Factor Analysis*. Oxford Univ. Press.

BHATTACHARYA, R. N. and RAO, R. R. (1986). *Normal Approximation and Asymptotic Expansions*. Krieger, Malabar.

BICKEL, P. and YAHAV, J. (1969). Some contributions to the asymptotic theory of Bayes' solutions. *Z. Wahrsch. Verw. Gebeite* **11** 257–276.

DOKSUM, K. A. and LO, A. Y. (1990). Consistent and robust Bayes procedures for location based on partial information. *Ann. Statist.* **18** 443–453.

HAMBLETON, R. K. (1989). Principles and selected applications of item response theory. In *Educational Measurement*, 3rd ed. (R. L. Linn, ed.) 147–200. MacMillan, New York.

HOLLAND, P. W. (1991). Private communication.

IBRAGIMOV, I. A. and HASMINSKII, R. Z. (1981). *Statistical Estimation*. Springer, New York.

JOAG-DEV, K. and PROSCHAN, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11** 286–295.

JUNKER, B. W. (1993). Conditional association, essential independence, and monotone unidimensional item response models. *Ann. Statist.* **21** 1359–1378.

LE CAM, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *University of California Publications in Statistics* **1** 277–330.

LE CAM, L. (1958). Les proprietes asymptotiques des solution de Bayes. *Publ. Inst. Statist. Univ. Paris* **7** 18–35.

LINDSAY, B., CLOGG, C. and GREGO, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *J. Amer. Statist. Assoc.* **86** 96–107.

LORD, F. M. (1980). *Application of Item Response Theory to Practical Testing Problems*. Erlbaum, Hillsdale, NJ.

TSUTAKAWA, R. K. and JOHNSON, J. C. (1990). The effect of uncertainty of item parameter estimation on ability estimates. *Psychometrika* **55** 371–390.

TSUTAKAWA, R. K. and SOLTYS, M. J. (1990). Approximation for Bayesian ability estimation. *Journal of Educational Statistics* **13** 117–130.

WALKER, A. M. (1969). On the asymptotic behavior of posterior distributions. *J. Roy. Statist. Soc. Ser. B* **31** 80–88.

YEN, W. M. (1984). Obtaining maximum likelihood estimates from number correct scores for the three parameter logistic model. *Journal of Educational Measurement* **21** 93–112.

DEPARTMENT OF STATISTICS
UNIVERSITY OF BRITISH COLUMBIA
6356 AGRICULTURAL ROAD, ROOM 333
VANCOUVER, BRITISH COLUMBIA
CANADA V6T 1Z2

INDIAN STATISTICAL INSTITUTE
203 BARRACKPORE TRUNK ROAD
CALCUTTA 700035
INDIA