The Annals of Statistics
1984, Vol. 12, No. 4, 1380–1387

ASYMPTOTIC EFFICIENCY OF ESTIMATORS OF FUNCTIONALS OF MIXED DISTRIBUTIONS

BY LUKE TIERNEY AND DIANE LAMBERT[1]

University of Minnesota and Carnegie-Mellon University

Suppose F is a mixture of a known parametric family of distributions with an unknown nonparametric mixing distribution. Consider the problem of estimating the value $T(dF)$ of a smooth functional T at the mixed distribution F. One nonparametric estimator of $T(dF)$ is $T(dF_n)$ where F_n denotes the empirical distribution. Under a differentiability condition on T given in this note, $T(dF_n)$ is shown to be fully efficient asymptotically in the sense that the limiting distribution of any other estimator of $T(dF)$ that is regular in the sense of Hajek and Beran can be expressed as a convolution of the limiting normal distribution of $T(dF_n)$ with another distribution. The differentiability condition is verified for the case that the parametric family being mixed is a one-parameter exponential family and the support of the mixing distribution is an infinite set with a finite accumulation point. Regularity is verified for the maximum likelihood estimator of the probability of zero in a mixed Poisson distribution under certain conditions on the mixing distribution. Moreover, the maximum likelihood estimator and the sample proportion of zeroes are both shown to be fully efficient in this example.

1. Introduction. Consider a distribution F with density f that is a mixture of a parametric family of densities $\{f(\cdot, \theta)\}$ with respect to a σ-finite measure μ. That is, $f(x) = \int f(x, \theta)dG(\theta)$ for some mixing distribution G. The form of the family $\{f(\cdot, \theta)\}$ is assumed to be known, but the distribution G is assumed to be unknown. Suppose the value of a known smooth functional $T(dF)$ of the mixed distribution is to be estimated using a random sample $X_1, \cdots, X_n$ from F. For example, $T(dF)$ might be the probability of zero in a mixed Poisson distribution.

Several approaches to estimating $T(dF)$ are available. One is to use the nonparametric estimator $T(dF_n)$, where F_n denotes the empirical distribution of $X_1, \cdots, X_n$. Another, following Simar (1976), Lindsay (1983a, 1983c), and Jewell (1982), among others, is to compute the maximum likelihood estimator (mle) $\hat{G}_n$ of the mixing distribution G (when G is identifiable) and to use $T((\int f(\cdot, \theta)d\hat{G}_n(\theta))d\mu)$. Both the nonparametric estimator and the partially parametric mle of $T(dF)$ are consistent in general, but little else is known about their statistical properties.

The limiting distribution of the mle of $T(dF)$ is investigated in Lambert and Tierney (1984) for the particular case that (i) $\{f(\cdot, \theta)\}$ is a family of Poisson

Received October 1982; revised May 1984.

[1] Research for this paper was done while both authors were at the Department of Statistics, Carnegie-Mellon University and was partially funded by NSF Grant MCS-8301692 at Carnegie-Mellon University.

AMS 1980 *subject classifications.* 62G20, 62G05.

Key words and phrases. Representation theorem, regular estimator, asymptotic variance bound.

densities, (ii) the support of the mixing distribution is an infinite set with a finite accumulation point, and (iii) $T(dF) = f(x_0)$ for $x_0 = 0, 1, 2, \cdots$. It is shown there that the mle has the same limiting distribution as the nonparametric estimator, which is the corresponding sample proportion, under an additional assumption on the behavior of the mixing distribution near zero. In this case, then, the nonparametric estimator has relative efficiency one with respect to the mle.

In this note the nonparametric estimator $T(dF_n)$ is shown to be asymptotically efficient among all "regular" estimators in the mixture context under more general conditions than those studied in Lambert and Tierney (1983). In Section 3 it is shown that under a given condition, called Condition D, the limiting distribution of any "regular" estimator $\hat{T}_n$ of $T(dF)$ can be expressed as the convolution of the limiting normal distribution of $T(dF_n)$ with another distribution. Thus, under Condition D any regular estimator is asymptotically at least as dispersed as the nonparametric estimator $T(dF_n)$. Roughly, Condition D, which is defined precisely and interpreted more fully in Section 2, might be described as a two-sided differentiability requirement on T. Our concept of regularity, which is also defined precisely in Section 2, is analogous to that of Hajek (1970) and Beran (1977). Condition D is verified in Section 3 for the case that $\{f(\cdot, \theta)\}$ is a one-parameter exponential family of densities and the support of the mixing distribution G is an infinite set with a finite accumulation point. Condition D fails typically if the support of G is a finite set. Regularity is verified in Section 4 for the mle of the probability of zero in a mixed Poisson distribution under the condition that the support of G is an infinite set with a finite accumulation point and $G(\theta)$ satisfies a certain condition on its behavior near zero. The implications of these results are discussed further in Section 5.

2. The lower bound on the variance.

The proof that the nonparametric estimator $T(dF_n)$ is asymptotically fully efficient has two parts. In this section, a lower bound on the variance of the asymptotic distribution of any "regular" estimator of $T(dF)$ is obtained. In Section 3, the nonparametric estimator $T(dF_n)$ is shown to achieve the lower bound.

The lower bound is most easily understood by drawing an analogy with a simpler version of our problem. Consider a k-dimensional parameter θ for a family of densities $\{f(\cdot, \theta)\}$ and a real-valued function $g(\cdot)$ of θ with a vector $D(\cdot)$ of first order partial derivatives. Let $\{\hat{T}_n\}$ be a sequence of asymptotically normal estimators of $g(\theta)$; that is, under θ the asymptotic distribution of $n^{1/2}(\hat{T}_n - g(\theta))$ is normal with variance $\sigma^2(\theta)$. When the behavior of $\hat{T}_n$ is "regular", $\sigma^2(\theta)$ is no smaller than $D(\theta)^t \mathscr{I}^{-1}(\theta) D(\theta) = \| D(\theta) \|^2$ where $\mathscr{I}(\theta)$ is the Fisher information matrix and $\| \cdot \|$ represents what might be called the information norm rather than the usual Euclidean norm (Walker, 1963). The analogy to be drawn in the mixture setting is that, under regularity conditions, the variance of the asymptotic distribution of an estimator $\hat{T}_n$ of a functional $T(dF)$ of the mixed distribution is no smaller than $\| \rho \|^2$ where ρ is the derivative of T at f and $\| \cdot \|$ is a non-Euclidean norm. In contrast to the typical k-dimensional problem, however, in the mixture problem the nonparametric estimator $T(dF_n)$

achieves the lower bound $\| \rho \|^2$. To extend the k-dimensional parameter lower bound to the mixture setting, we work with a generalized χ^2 norm that can be interpreted as an information type norm, a notion of score functions developed by Lindsay (1983b), and a definition of regularity that is similar to that used by Beran (1977a) and Begun et al. (1982). With these choices, the correctness of the lower bound analogy between the k-dimensional parameter problem and the mixture problem follows directly from Theorem 6 of Beran (1977a) on the representation of limiting distributions of regular estimators.

The specific framework used throughout this note is the following. Let $(\mathbb{R}, \mathscr{R})$ denote the real line with its Borel σ-field, and let $(\Theta, \mathscr{F})$ be a measurable space with the point sets $\{\theta\}$ belonging to $\mathscr{F}$ for each $\theta \in \Theta$. Also let $\{f(\cdot, \theta): \theta \in \Theta\}$ be a family of densities on $\mathbb{R}$ with respect to the σ-finite measure μ. Assume that $f(\cdot, \cdot)$ is $\mathscr{R} \times \mathscr{F}$ measurable. Let $\mathscr{P}$ denote the collection of all probability measures on $\mathscr{F}$ and let $\mathscr{M} = \{f(\cdot) = \int f(\cdot, \theta)P(d\theta): P \in \mathscr{P}\}$ be the set of densities for measures of $\{f(\cdot, \theta)\}$. Also, let $\mathscr{S}$ be the set of all densities on $\mathbb{R}$ with respect to μ.

Choose and fix a mixed density f from $\mathscr{M}$; in the following we assume that T is to be estimated at f. For the chosen f, define the corresponding space of $f^{-1/2}$-weighted square integrable densities by $\mathscr{L}_f^2 = \{g \in \mathscr{S}: \int (g^2/f)d\mu < \infty\}$. Define an inner product $\langle \cdot , \cdot \rangle$ and its associated norm $\| \cdot \|$ on $\mathscr{L}_f^2$ by $\langle g, h \rangle = \int (gh/f)d\mu$ and $\| g \| = \langle g, g \rangle^{1/2}$. If the measure μ is discrete and uniform then $\| g - f \|$ reduces to what might be called the χ^2-distance from g to f. Thus, this norm can be thought of as a generalized χ^2-distance. Next, let $\mathscr{G}$ be the set of mixed densities that are a finite distance from f; i.e., $\mathscr{G} = \mathscr{M} \cap \mathscr{L}_f^2$.

In the k-dimensional problem, the score functions are the derivatives of the log-likelihood in the coordinate directions of the k-dimensional space. Here the analogy to a score function is defined in terms of the set $\mathscr{B}$ of directions from which the mixed density f can be approached through sequences in $\mathscr{G}$. To be precise, let

$$\mathscr{B} = \{\beta \in L_f^2: \| n^{1/2}(f_n - f) - \beta \| \to 0 \text{ for some } \{f_n\} \subset \mathscr{G}\}.$$

As Lindsay (1983b) has shown, $\mathscr{B}$ essentially consists of those densities in $\mathscr{L}_f^2$ that are of the form $cf \cdot \mathsf{U}$ where c is a positive constant and U is a "score function", i.e. U is a directional derivative of the log-likelihood in a direction leading into $\mathscr{G}$ along a smooth one-dimensional subfamily of $\mathscr{G}$. In this sense, $\mathscr{B}$ can be thought of as the "score space" of $\mathscr{G}$ at f. Further, if $n^{1/2}(f_n - f)$ tends to β, then $n \| f_n - f \|^2$ tends to $\| \beta \|^2 = \int (\beta^2/f)d\mu$, the Fisher information at f in the direction β. Note that $\mathscr{B}$ is a closed cone, and that if $\mathscr{G}$ is convex then $\mathscr{B}$ is a closed, convex cone.

Beran (1977a) calls an estimator $\hat{T}_n$ regular at a density f if the asymptotic distribution of $n^{1/2}(\hat{T}_n - T(f))$ is unchanged whenever f is perturbed to a density f_n for each n and the sequence $\{f_n\}$ converges to f at rate $n^{-1/2}$. As in Begun et al. (1982), we require regularity only for some perturbations f_n of f. Here, for each n, the perturbation f_n must belong to the set $\mathscr{G}$ of mixed densities within a finite distance of f rather than to the set of all densities. Specifically, an estimator $\hat{T}_n$

of $T(f)$ based on a sample of size n from f is said to be $\mathscr{G}$-regular at f if the limiting distribution of $n^{1/2}(\hat{T}_n - T(f_n))$ under the sequence $\{f\}$ is the same for any sequence $\{f_n\}$ in $\mathscr{G}$ such that $n^{1/2}(f_n - f)$ converges to some $\beta \in \mathscr{B}$.

Finally, a functional T on $\mathscr{L}_f^2$ is said to be differentiable at f if there is a function $\rho \in \mathscr{L}_f^2$ such that for any sequence $\{f_n\}$ of densities in $\mathscr{L}_f^2$ that converge to f,

$$\lim_n \| f_n - f \|^{-1}[T(f_n) - T(f) - \langle \rho, f_n - f \rangle] = 0.$$

The derivative ρ is determined uniquely if it is taken to be orthogonal to f. Since regularity and differentiability of T alone are not strong enough for our proof of the lower bound to hold, we also require that the derivative of T meets the following condition.

CONDITION D. The derivative ρ of T at f satisfies $a\rho \in \mathscr{B}$ for any real constant a.

In words, Condition D requires that the score space $\mathscr{B}$ (or directions from which f can be approached through mixtures) be large enough to contain the linear space generated by the derivative ρ of T. Therefore, Condition D can be satisfied only if the set of mixed densities is sufficiently rich. Condition D is very restrictive when the mixing distribution G has a finite number N of support points. In this case, the largest linear space that can be contained in the set $\mathscr{B}$ of score directions is typically $2N - 1$ dimensional, since there are N score functions for the locations of the support points of the mixing distribution and $N - 1$ score functions for their weights. Hence, there are at most $2N - 1$ asymptotically linearly independent functionals of f that can satisfy Condition D if the mixing distribution has only N points. On the other hand, if the support of G is an infinite set, then it is plausible that every functional $T(dF)$ satisfies Condition D. A restricted version of this conjecture is proved for one-parameter exponential families in Section 3. Finally, another rough interpretation of Condition D is that if ρ is thought of as the gradient of T at f, then f must be an interior point of the segment of the line through f in the direction ρ that intersects the set $\mathscr{G}$. In this sense, Condition D is a two-sided differentiability requirement.

It is now straightfoward to recast and reprove Theorem 6 of Beran (1977a) in the mixture framework. The revised theorem is given next as Theorem 1; its proof is omitted. Theorem 1 differs from Beran's result in two ways. First, the χ^2-distance $\| \cdot \|$ defined above is used instead of the Hellinger distance, and, second, regularity is required only with respect to the subset $\mathscr{G}$ of mixed densities, not with respect to all densities in $\mathscr{L}_f^2$.

THEOREM 1. *Let T be a functional on $\mathscr{L}_f^2$ that is differentiable at f with derivative ρ. Suppose that ρ is orthogonal to f and satisfies Condition D. Let $\hat{T}_n$ be an estimator of $T(f)$ that is $\mathscr{G}$-regular at f. Then the limiting distribution of $n^{1/2}(\hat{T}_n - T(f))$ can be written as the convolution of a normal$(0, \| \rho \|^2)$ distribution with another distribution depending on f.*

3. The asymptotic efficiency of $T(dF_n)$. Given a functional T to be estimated at a mixed density f, let S denote the restriction of T to distributions with densities with respect to the measure μ. The functional T is called smooth at f if (i) S is differentiable at f with derivative ρ and (ii) $[T(dF_n) - T(dF)] - \int \psi(dF_n - dF) = o_p(n^{-1/2})$, where $\psi = \rho/f$ and F_n denotes the empirical distribution of a random sample of size n from f. Note that if T is smooth at f then $n^{-1/2}(T(dF_n) - T(dF))$ has a limiting normal distribution with mean zero and variance $\int \psi^2 dF = \| \rho \|^2$. Also note that, as in Beran (1977b), a contiguity argument implies that $T(dF_n)$ is regular if T is smooth. Also, if T is smooth and satisfies Condition D, then Theorem 1 implies that no $\mathscr{G}$-regular estimator $\hat{T}_n$ can be more efficient than $T(dF_n)$. Moreover, if $\mathscr{B}$, the score space, is sufficiently rich that it contains every g in $\mathscr{L}_f^2$ orthogonal to f, then Condition D is satisfied by every smooth functional T and every smooth functional $T(f)$ is estimated asymptotically efficiently by $T(dF_n)$. This argument proves Theorem 2 below.

THEOREM 2. *Let T be any smooth functional at f and suppose $\hat{T}_n$ is a $\mathscr{G}$-regular estimator of T at f. If the orthogonal complement of the linear space generated by f is contained in $\mathscr{B}$, then the limiting distribution of $n^{1/2}(\hat{T}_n - T(dF))$ is the convolution of a normal$(0, \| \rho \|^2)$ distribution with another distribution depending on f.*

To summarize, the asymptotic efficiency of $T(dF_n)$ depends on the richness of the set $\mathscr{B}$ of directions from which f can be approached. If the support of the mixing distribution G is finite with N points, say, then the largest linear space that can be contained in $\mathscr{B}$ has $2N - 1$ dimensions typically. In this case, therefore, there are usually $2N - 1$ asymptotically linearly independent smooth functionals T for which the associated nonparametric estimators are asymptotically efficient. In the same spirit, if the support of G is infinite, then one might expect every smooth T to be estimated asymptotically efficiently by $T(dF_n)$. What is needed is that $\mathscr{B}$ contain the orthogonal complement of the space generated by f.

If the support of G is an infinite set, then in some cases a proof that $\mathscr{B}$ contains the orthogonal complement of the space generated by f can be constructed as follows. Suppose not. Then since $\mathscr{B}$ is a closed convex cone, the supporting hyperplane theorem implies that there is a $\gamma \neq 0$ in $\mathscr{L}_f^2$ such that γ is orthogonal to f and $\langle \gamma, \beta \rangle \leq 0$ for all β in $\mathscr{B}$. Taking $f_n(\cdot) = f(\cdot) + n^{-1/2}[f(\cdot, \theta) - f(\cdot)]$, which corresponds to the mixing distribution $G_n = G + n^{-1/2}[\delta_\theta - G]$, where δ_θ is the probability distribution that assigns mass one to θ, shows that $f(\cdot, \theta) - f(\cdot) \in \mathscr{B}$ for all $\theta \in \Theta$. Therefore, since γ is orthogonal to f,

$$\int [(f(\cdot, \theta) - f(\cdot))\gamma(\cdot)/f(\cdot)] \, d\mu = \int [f(\cdot, \theta)\gamma(\cdot)/f(\cdot)] d\mu \leq 0 \quad \text{for all} \quad \theta \in \Theta.$$

Integrating the left side of this inequality with respect to G gives zero, since γ is orthogonal to f, so that the inequality is an equality for G-almost all θ. If the support of G is large, it usually follows that $\gamma = 0$ μ-a.e., which is a contradiction.

For a case in which the preceding proof by contradiction is valid, take the one-parameter exponential family $f(x, \theta) = \exp(\theta x - \kappa(\theta))$ with $\Theta \subset (-\infty, \infty)$ and $\mathscr{F}$ the Borel σ-field on Θ. Suppose the support of the mixing distribution G is an infinite set with a finite accumulation point. Also suppose that the orthogonal complement of the linear space spanned by f is not contained in $\mathscr{B}$. Then there is a $\gamma \in \mathscr{L}_f^2$ such that $\gamma \neq 0$ μ-a.e. and $\langle \gamma, f(\cdot, \theta) \rangle = 0$ for G-almost all θ. Hence, there is a sequence of distinct θ_j converging to a finite limit such that $\int [\exp(\theta_j x)\gamma(x)/f(x)]d\mu = 0$ for all j. Since the support of G is infinite, the trivial case $G(\{-\infty, \infty\}) = 1$ is excluded; hence $f(x) > 0$ for all x, and for all θ

$$\left| \int \left[\frac{e^{\theta x}\gamma(x)}{f(x)} \right] d\mu \right|^2 \leq \| \gamma \|^2 \int \left[\frac{e^{2\theta x}}{f(x)} \right] d\mu,$$

which is finite. Thus, $\int [\exp(\theta x)\gamma(x)/f(x)]d\mu(x)$ is an analytic function of θ, and by an elementary property of power series $\langle \gamma, f(\cdot, \theta) \rangle = 0$ holds for all complex θ. Hence $\gamma(x) = 0$ μ-a.e., which is a contradiction. The following corollary to Theorem 2 is now proved.

COROLLARY. *Suppose f is a mixture of densities from a one-parameter exponential family and the support of the mixing distribution is an infinite set with a finite accumulation point. Let T be smooth at f with corresponding derivative ρ and let $\hat{T}_n$ be $\mathscr{G}$-regular at f. Then the limiting distribution $n^{1/2}(T(dF_n) - T(f))$ is normal$(0, \| \rho \|^2)$ and the limiting distribution of $n^{1/2}(\hat{T}_n - T(f))$ is the convolution of the normal$(0, \| \rho \|^2)$ distribution with another distribution depending on f.*

4. An example. Take F to be a mixed Poisson distribution and $T(dF)$ to be the probability of zero under F, i.e. $T(dF) = f(0) = \int e^{-\lambda} dG(\lambda)$ for some mixing distribution G. The derivative ρ of T at F is described by the vector $(f(0)(1 - f(0)), -f(0)f(1), -f(0)f(2), \cdots)$. Note that ρ is orthogonal to f, and that $\| \rho \|^2 = f(0)(1 - f(0))$. Suppose the support of G is an infinite set with a finite accumulation point. Then by the corollary, the sample proportion of zeroes, which is $\mathscr{G}$-regular, is an asymptotically fully efficient estimator for $T(dF)$ among $\mathscr{G}$-regular estimators, and its asymptotic variance is $\| \rho \|^2$.

As pointed out by a referee, this example can be generalized considerably. If f is a mixture of densities from any one-parameter exponential family and the support of the mixing distribution is an infinite set with a finite accumulation point, then for any function h the statistic $T(dF_n) = \int h\, dF_n = n^{-1} \sum h(X_i)$ is an efficient estimator (among $\mathscr{G}$-regular estimators) of its expectation $T(dF) = \int h\, dF$. The derivative of T is $\rho(x) = f(x)(h(x) - \int h\, dF)$, and the asymptotic variance of $T(dF_n)$ is $\| \rho \|^2 = \int h^2 dF - (\int h\, dF)^2$.

It is not clear whether the mle $\hat{T}_n$ is also $\mathscr{G}$-regular in the Poisson example, but it is $\mathscr{G}_0$-regular for a restricted class $\mathscr{G}_0$ of mixed densities. The argument is as follows. Given positive constants C, d, b, ε, let $\mathscr{G}_0$ be the class of mixtures with respect to mixing distributions G satisfying (i) the support of G is an infinite set bounded above by C and (ii) $G(x + y) - G(x) > dy^b$ for all $x, y, \in (0, \varepsilon)$. The requirements (i) and (ii) are given as Condition 5.1 in Lambert and Tierney (1984), and are shown to imply that $n^{1/2}(\hat{T}_n - T(dF))$ is asymptotically normal$(0, f(0)(1 - f(0)))$. Since the limiting arguments of that paper hold uniformly over $\mathscr{G}_0$ and the variances of the limiting distributions vary continuously over $\mathscr{G}_0$, the mle $\hat{T}_n$ is $\mathscr{G}_0$-regular. Further, for any $G \in \mathscr{G}_0$, the mixtures corresponding to $G_n = G + n^{-1/2}(\delta_\theta - G)$ are in $\mathscr{G}_0$ for any $\theta \le C$ once n is large enough. Hence, the argument preceding the Corollary shows that for any f in $\mathscr{G}_0$ the set of directions $\mathscr{B}_0$ from which f can be approached through $\mathscr{G}_0$ is equal to the orthogonal complement of f. Hence, Condition D holds for $\mathscr{B}_0$, and the mle $\hat{T}_n$ is asymptotically efficient among $\mathscr{G}_0$-regular estimators of $T(dF)$.

5. Discussion. The condition of regularity was developed by Hajek (1970) and Beran (1977) to exclude superefficient estimators. As a continuity requirement, it has some intuitive appeal, at least if the parameter is in the interior of the parameter space. In finite dimensional problems, mle's are generally regular on the interior of the parameter space. When the parameter is on the boundary of the parameter space, however, regularity may be an unreasonable requirement. For example, if the mean θ of a Poisson distribution is estimated on the basis of a random sample of size n, and θ is known to be less than or equal to 5, then the mle $\hat{\theta}$ is $\min\{\bar{X}_n, 5\}$. If θ is perturbed from 5 to $\theta_n = 5 - n^{-1/2}\varepsilon$ for some positive ε, then the limiting distribution of $n^{1/2}(\hat{\theta}_n - \theta_n)$ is the same as the distribution of $\min(5^{1/2}Z - \varepsilon, 0)$, where Z has a normal(0, 1) distribution. Thus, the mle is not regular. For the best regular estimator T_n, $n^{1/2}(T_n - 5)$ is asymptotically normal (0, 5) when $\theta = 5$, but an estimator with smaller mean squared error can be obtained by truncating this T_n.

Since a finite mixture is on the boundary of the space of all mixtures, regularity may not be an appropriate requirement to impose and the mle's cannot be expected to be regular if the mixing distribution is known to be finite. On the other hand, if the support of the mixing distribution is an infinite set, then the mixture is, in a sense, in the interior of the set of all mixtures, and the condition of regularity may not be too restrictive. This informal argument is supported by the results of the preceding section.

Finally, we note that if, as is reasonable to expect, the mle's are asymptotically efficient for general mixtures when the mixing distribution is an infinite set with a finite accumulation point, then the results of this paper imply that their limiting distribution is normal$(0, \|\rho\|^2)$. The nonparametric estimator based on the empirical distribution has the same distribution, but this observation should not be construed as a suggestion to use the nonparametric estimator instead of the mle. In finite samples the mle is preferable since it produces smoother estimates than the estimator based on the empirical distribution. Moreover, functionals of the mixing distribution itself can generally be estimated only by means of the mle; simple estimators based on the empirical distribution are not available.

Acknowledgement. We would like to thank the associate editor and the referees for their comments that led to significant improvements in the presentation of this paper.

REFERENCES

BEGUN, J. M., HALL, W. J., HWANG, W-M. and WELLNER, J. A. (1982). Information and asymptotic efficiency in parametric-nonparametric models. *Ann. Statist.* **11** 432–452.

BERAN, R. (1977a). Robust location estimates. *Ann. Statist.* **5** 431–444.

BERAN, R. (1977b). Estimating a distribution function. *Ann. Statist.* **5** 400–404.

HAJEK, J. (1970). A characterization of limiting distributions of regular estimates. *Z. Wahrsch. verw. Gebiete* **14** 323–330.

JEWELL, N. (1982). Mixtures of exponential distributions. *Ann. Statist.* **10** 479–484.

LAMBERT, D. and TIERNEY, L. (1984). Asymptotic properties of maximum likelihood estimates in the mixed Poisson model. *Ann. Statist.* **12** 1388–1399.

LINDSAY, B. G. (1983a). The geometry of mixture likelihoods: A general theory. *Ann. Statist.* **11** 86–94.

LINDSAY, B. G. (1983b). Efficiency of the conditional score in a mixture setting. *Ann. Statist.* **11** 486–497,

LINDSAY, B. G. (1983c). The geometry of mixture likelihoods, part II: The exponential family. *Ann. Statist.* **11** 783–792.

SIMAR, L. (1976). Maximum likelihood estimation of a compound Poisson process. *Ann. Statist.* **4** 1200–1209.

WALKER, A. M. (1963). A note on the asymptotic efficiency of an asymptotically normal estimator sequence. *J. Roy. Statist. Soc. B* **25** 193–200.

DEPARTMENT OF THEORETICAL STATISTICS
UNIVERSITY OF MINNESOTA
206 CHURCH STREET, S.E.
MINNEAPOLIS, MINNESOTA 55455

DEPARTMENT OF STATISTICS
CARNEGIE-MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213