The Annals of Statistics
1977, Vol. 5, No. 3, 522–529

ASYMPTOTIC EFFICIENCY OF MINIMUM VARIANCE UNBIASED ESTIMATORS[1]

By Stephen Portnoy

University of Illinois at Champaign-Urbana

Consider a regular p-dimensional exponential family such that either the distributions are concentrated on a lattice or they have a component whose k-fold convolution has a bounded density with respect to Lebesgue measure. Then, if a parametric function has an unbiased estimator, the minimum variance unbiased estimators are asymptotically equivalent to the maximum likelihood estimators; and, hence, are asymptotically efficient. Examples are given to show that a condition like the above is needed to obtain the asymptotic equivalence.

1. Introduction and preparatory lemma. Although examples where minimum variance unbiased (mvu) estimators are very poor are well known, these examples seem to require small sample size. In fact, in all common examples, mvu estimators appear to be asymptotically efficient. (Here a sequence of estimators, $\{T_n\}$, of a parameter, θ, is defined to be asymptotically efficient if $n^{\frac{1}{2}}(T_n - \theta) \to \mathscr{N}(0, v(\theta))$ where $v(\theta)$ is the Cramér–Rao lower bound.) However, the only general result along these lines is given by Sharma (1973) who proves asymptotic efficiency of the mvu estimators of a positive integral power of the natural parameter in certain regular one-dimensional exponential families. This paper considers the case of p-dimensional exponential families satisfying the following:

Condition A. Either the distributions are all lattice distributions (on the same lattice) or they have a component such that for some k (independent of the parameter value) the k-fold convolution has a bounded density with respect to Lebesgue measure (equivalently, for some j the jth power of the characteristic function of the component is absolutely integrable).

The basic result presented here is the following: in a regular exponential family satisfying Condition A, if a parametric function has an unbiased estimator then the sequence of mvu estimators is asymptotically equivalent to the sequence of maximum likelihood estimators (in the sense that the difference is $\mathscr{O}_p(1/n)$); and, hence, mvu estimators are asymptotically efficient. Furthermore, examples are presented to show that some hypothesis like Condition A is necessary. It is interesting to note that asymptotic efficiency of mvu estimators also seems to hold in nonregular cases (for example, in estimating parameters of a uniform distribution). However, the methods of this paper would appear to be inapplicable to the nonregular case.

Received April 1976; revised October 1976.

[1] Research partially supported by NSF grant MPS 75-07978.

AMS 1970 *subject classifications.* Primary 62F20; Secondary 62F10.

Key words and phrases. Asymptotic efficiency, exponential families, minimum variance unbiased estimators.

The results will first be proven under the following:

CONDITION A′. Either the distributions are all lattice, or for some j the jth power of the characteristic function is absolutely integrable.

The result here depends basically on the following fact: if $X_1, X_2, \cdots$ are i.i.d. random variables with finite third moment, mean μ, and satisfying Condition A′, then for any function u with $E|X_1|^2|u(X_1)| < +\infty$, $E[u(X_1) \mid \bar{X}_n = \mu] = Eu(X_1) + \mathscr{O}(1/n)$ (where $\bar{X}_n$ is the sample mean). This result was first formulated in this form by Professor Sternberg at Harvard University and is given together with various generalizations and applications to statistical mechanics and probability theory in Zabell (1974 and 1976). The result in this paper requires a bound of the form $c(\theta)/n$ where c is continuous in the parameter θ in p dimensions. Although this extension is straightforward, the proof is sketched here since the result is probably not well known by statisticians.

Consider an exponential family with densities

$$p_\theta(x) = e^{-\psi(\theta)+\theta x'} \qquad x \in R^p,\ \theta \in \Theta \subset R^p \tag{1.1}$$

with respect to a measure ν on R^p, where $\psi(\theta) = \log \int e^{\theta x'}\, d\nu(x)$ and $\Theta = \{\theta : \psi(\theta) < \infty\}$ is the natural parameter space. Let $\mu = \psi'(\theta)$ denote the mean parameter, where ψ' denotes the gradient of ψ; and let

$$D = \{\mu \in R^p : \mu = \psi'(\theta) \text{ for some } \theta \in \Theta^0\} \tag{1.2}$$

where Θ^0 is the interior of Θ (note: D is diffeomorphic to Θ^0).

LEMMA 1. *Let $X_1, X_2, \cdots$ be a sequence of* i.i.d. *random variables with distribution in an exponential family (as described above) satisfying Condition* A′. *Let $g: D \to R$ and $u: R^p \to R$ be such that for some k*

$$E_\mu u(X_1 + \cdots + X_k) = g(\mu) \qquad \textit{for} \quad \mu \in D \tag{1.3}$$

with

$$E_\mu|u(X_1 + \cdots + X_k)| < +\infty \qquad \textit{for} \quad \mu \in D\,. \tag{1.4}$$

Then there is a continuous function c on D such that

$$|E\{u(X_1 + \cdots + X_k) \mid \bar{X}_n = \mu\} - g(\mu)| \leqq \frac{c(\mu)}{n} \tag{1.5}$$

for all $\mu \in D$ and for $n \geqq j + k$.

PROOF. The proof will be sketched for the case of densities in Condition A. The lattice case (involving integrals of characteristic functions over compact sets) would be even simpler. So let $f_m(x)$ denote the density of $S_m = X_1 + \cdots + X_m - m\mu$ and (by change of variable) write

$$E[u(S_k) \mid S_n = n\mu] = \frac{\int f_k(-z)u(k\mu - z)f_{n-k}(z)\, dz}{f_n(0)}\,. \tag{1.6}$$

By Parseval's relation, the numerator integral equals

$$(1.7)\qquad \left(\frac{1}{2\pi}\right)^p \int h(t)\varphi^{n-k}(t)\,dt = \left(\frac{1}{2\pi}\right)^p \int \tilde{h}(t)\varphi^n(t)\,dt$$

where $\tilde{h}(t) = h(t)\varphi^{-k}(t)$ and

$$(1.8)\qquad h(t) = \int u(k\mu + z)f_k(z)e^{-itz'}\,dz$$

and where $\varphi(t) = \varphi_\mu(t)$ is the characteristic function of the distribution in the exponential family with mean, μ, subtracted.

Let $\varepsilon(\mu)$ be the largest value $\varepsilon_0 \leqq 1$ such that $\operatorname{Re}\varphi_\mu(t) \geqq \frac{1}{2}$ for $||t|| \leqq \varepsilon(\mu)$. Since $\varphi_\mu(t)$ is uniformly continuous in μ, $\varepsilon(\mu)$ is continuous in μ on D. Thus, since X has finite third moments, we have the Taylor series expansion

$$n \log \varphi_\mu(sn^{-\frac{1}{2}}) = -\tfrac{1}{2}s\Sigma s' + n^{-\frac{1}{2}}P_0(s)$$

for $||s|| \leqq n^{\frac{1}{2}}\varepsilon(\mu)$. Therefore, for $||s|| \leqq n^{\frac{1}{2}}\varepsilon(\mu)$

$$(1.9)\qquad \varphi_\mu^{\,n}(sn^{-\frac{1}{2}}) = e^{-\frac{1}{2}s\Sigma s'}(1 + n^{-\frac{1}{2}}P(s))$$

where Σ is the covariance matrix of X and $P(s)$ is a linear combination of powers of coordinates of s with coefficients uniformly bounded in terms of $E_\mu||X||^3$.

Next, break the integral in (1.7) into the integral over $\{||t|| \leqq \varepsilon\}$ and $\{||t|| > \varepsilon\}$. The latter integral can be bounded by

$$(1.10)\qquad \int_{||t||>\varepsilon} |h(t)\varphi^j(t)|\varphi^{n-k-j}(t)|\,dt \leqq \int |h(t)|\,|\varphi^j(t)|\,dt \cdot e^{-(n-k-j)c_1(\mu)}$$

where $c_1(\mu) = -\log \sup_{||t||\geqq\varepsilon(\mu)} |\varphi_\mu(t)| > 0$; $c_1(\mu)$ is continuous in μ since $\varphi_\mu(t) \to \varphi_{\mu_0}(t)$ as $\mu \to \mu_0$ uniformly in t.

Since $E|u(S_k)| < +\infty$, differentiating a Laplace transform shows that $ES_k^{\,2}|u(S_k)| < +\infty$. Hence, $\tilde{h}(t)$ has uniformly bounded second partial derivatives; and the remaining integral in (1.7) can be written

$$(1.11)\qquad \left(\frac{1}{2\pi}\right)^p \int_{||t||\leqq\varepsilon} (\tilde{h}(0) + t\tilde{h}'(0) + \tfrac{1}{2}t\tilde{h}''(\xi)t')\varphi^n(t)\,dt\,.$$

By replacing the integral over $\{||t|| \leqq \varepsilon\}$ of the first term by the integral over R^p (as above), this first term contributes

$$h(0)\left(\frac{1}{2\pi}\right)^p \int \varphi^n(t)\,dt = Eu(S_k)f_n(0)\,.$$

For the remaining terms, let $t = sn^{-\frac{1}{2}}$. The second term becomes

$$(1.12)\qquad n^{-p/2}\int_{||s||\leqq\varepsilon n^{\frac{1}{2}}} n^{-\frac{1}{2}}sh'(0)\varphi^n(sn^{-\frac{1}{2}})\,ds$$
$$= n^{-p/2}\left\{\int_{||s||\leqq\varepsilon n^{\frac{1}{2}}} n^{-\frac{1}{2}}sh'(0)e^{-\frac{1}{2}s\Sigma s'}\,ds + \frac{1}{n}\int_{||s||\leqq\varepsilon n^{\frac{1}{2}}} sh'(0)P(s)e^{-\frac{1}{2}s\Sigma s'}\,ds\right\}$$

(where (1.9) is used). Since $\int s\exp\{-\frac{1}{2}s\Sigma s'\}\,ds = 0$, the first term in (1.12) contributes an error like (1.10). The other term is bounded by $n^{-p/2}\cdot n^{-1}c_2(\mu)$, where $c_2(\mu)$ depends only on moments and, hence, is continuous in μ. Since

$n^{p/2}f_n(0)$ converges to a $(2\pi)^{-p/2}$, the term (1.12) divided by $f_n(0)$ is bounded by

$$\frac{c_3(\mu)}{n} + c_4(\mu)n^{p/2}e^{-nc_1(\mu)} \leqq \frac{c_5(\mu)}{n}. \tag{1.13}$$

Equation (1.9) can similarly be used in the last term in (1.11) to obtain a bound of form (1.13); and this completes the proof. □

REMARK. It should be noted that exponentiality is not actually required in the lemma. Indeed, for fixed values of the parameter, the result is given by Zabell (1974 and 1976) and only requires (essentially) Condition A′ and finite third moments. It would be possible to prove the continuity of the bound $c(\mu)$ under appropriate regularity conditions, but these conditions would be somewhat complicated. More important, the only application would be to families of distributions where the mean is sufficient, and this would essentially imply that the family is an exponential one.

2. The basic results and counterexamples.

THEOREM 1. *Let* $X_1, X_2, \cdots$ *be a sequence of* i.i.d. *random variables from a member of an exponential family in* R^p *(see* (1.1)) *satisfying Condition* A′. *Let* $g\colon D \to R$ *be a parametric function for which there is an unbiased estimator* $t(X_1, \cdots, X_k)$ *satisfying* $E_\mu|t(X_1, \cdots, X_k)| < +\infty$ *for all* $\mu \in D$. *Let* $T_1, T_2, \cdots$ *be the sequence of* mvu *estimators,* $T_n(\bar{X}_n) = E[t(X_1, \cdots, X_k) \mid \bar{X}_n]$ *(where* $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$*).*

Let V_n *be the maximum likelihood estimator* (mle) *of* $g(\mu)$ *if it exists. Then if* $\mu \in D$, V_n *exists with probability tending to one,* $T_n(\bar{X}_n) = g(\bar{X}_n) + \mathscr{O}_p(n^{-1})$, *and* $n^{\frac{1}{2}}(T_n(\bar{X}_n) - g(\mu)) \to \mathscr{N}(0, v(\mu))$ where $v(\mu)$ *is the Cramér–Rao lower bound (and, thus, the sequence* $\{T_n\}$ *is asymptotically efficient).*

PROOF. Since $\psi(\theta)$ is convex, if $\bar{X}_n \in D$ the likelihood is maximized uniquely at $\hat{\theta}$ satisfying $\psi'(\hat{\theta}) = \bar{X}_n$; and, hence, the mle of μ is $\bar{X}_n$. Thus, if $\bar{X}_n \in D$, the mle of $g(\mu)$ is $g(\bar{X}_n)$. So, if $\mu_0 \in D$ is fixed, with probability tending to one, $\bar{X}_n \in D$ and $V_n = g(\bar{X}_n)$ (since D is open and $\bar{X}_n \to^P \mu_0$). Straightforward calculations show that (under μ_0) $n^{\frac{1}{2}}(V_n - g(\mu_0)) \to^D \mathscr{N}(0, v(\mu_0))$ where $v(\mu_0)$ is the Cramér–Rao lower bound. Thus $\{V_n\}$ is efficient, and it remains to show that $T_n(\bar{X}_n) = V_n + \mathscr{O}_p(n^{-1})$.

So let $S_k = X_1 + \cdots + X_k$ and let $u(S_k) = E[t(X_1, \cdots, X_k) \mid S_k]$. Then $u(S_k)$ is an unbiased estimate of g with $E|u(S_k)| < +\infty$, and $T_n(\bar{X}_n) = E[u(S_k) \mid \bar{X}_n)]$. So by Lemma 1,

$$|T_n(\mu) - g(\mu)| \leqq \frac{c(\mu)}{n} \qquad \text{for} \quad \mu \in D.$$

Now fix $\mu_0 \in D$. Since $P_{\mu_0}\{\bar{X}_n \notin D\} \to 0$ and c is continuous, there is a constant b such that $P_{\mu_0}\{\bar{X}_n \in D$ and $c(\bar{X}_n) \geqq b\} \to 0$. Therefore $P_{\mu_0}\{|T_n(\bar{X}_n) - g(\bar{X}_n)| \geqq b/n\} \leqq P_{\mu_0}\{\bar{X}_n \notin D$ or $(\bar{X}_n \in D$ and $c(\bar{X}_n) \geqq b)\} \to 0$, and the result follows. □

THEOREM 2. *The conclusion of Theorem* 1 *holds if Condition* A′ *is replaced by Condition* A *in the hypotheses.*

PROOF. It suffices to consider the case where the dominating measure ν may be written $\nu = \nu_0 + \nu_1$ and where, with

$$e^{\psi_l(\theta)} \equiv \int e^{\theta x'}\, d\nu_l(x) \qquad l = 0, 1\,,$$

the characteristic function of $e^{-\psi_0(\theta)+\theta x'}\, d\nu_0(x)$ has its jth power absolutely integrable. In this case, the distribution of the observations $(X_1, X_2, \cdots)$ may be represented as follows: let $(X_1^{(l)}, X_2^{(l)}, \cdots)$ be i.i.d., according to distribution $e^{-\psi_l(\theta)+\theta x'}\, d\nu_l(x)$ $(l = 0, 1)$, and define

$$q(\theta) = e^{\psi_0(\theta)}/(e^{\psi_0(\theta)+\psi_1(\theta)})\,.$$

Then $X_i = X_i^{(0)}$ with probability $q(\theta)$ and $X_i = X_i^{(1)}$ with probability $(1 - q(\theta))$ (for $i = 1, 2, \cdots$). Now fix n and let J be the random variable defined to be the number of indices, i, $k < i \leqq n$ such that $X_i = X_i^{(0)}$. Then (with μ the mean parameter and θ the corresponding natural parameter)

$$E[u(S_k) \mid S_n] = E_\mu[u(S_k) \mid S_n, J \geqq j] + P_\mu[J < j \mid S_n] w(S_n) \tag{2.1}$$

where j comes from Condition A and

$$w(S_n) = E_\mu[u(S_k) \mid S_n, J < j] - E_\mu[u(S_k) \mid S_n, J \geqq j]\,. \tag{2.2}$$

Let $p_\mu(i) = P_\mu(J = i) = \binom{n-k}{i} q^i(\theta)(1 - q(\theta))^{n-k-i}$. Then (as in Lemma 1, equation (1.6))

$$E_\mu[u(S_k) \mid S_n = n\mu, J \geqq j] = \frac{\sum_{i=j}^{n-k} p_\mu(i) \int u(k\mu - z) f_{n-k}(z)\, dF_{k\mu - S_k}(z)}{\sum_{i=j}^{n-k} p_\mu(i) f_n(0)}\,. \tag{2.3}$$

From Lemma 1, the numerator of (2.3) can be written

$$\sum_{i=j}^{n-k} p_\mu(j)\{f_n(0) E_\mu u(S_k) + \varepsilon(n, i, \mu)\}$$

where from the proof of Lemma 1 (see (1.13))

$$\varepsilon(n, i, \mu) \leqq \frac{c_1(\mu)}{n} + c_2(\mu) n^{p/2} e^{-ic_3(\mu)}$$

with $c_m(\mu)$ continuous in μ on D. Therefore

$$\begin{aligned}
&|E_\mu[u(S_k) \mid S_n = n\mu, J \geqq j] - E_\mu u(S_k)| \\
&\quad \leqq \frac{c_1(\mu)}{n} + \frac{\sum_{i=j}^{n-k} p_\mu(i) c_2(\mu) n^{p/2} e^{-ic_3(\mu)}}{\sum_{i=j}^{n-k} p_\mu(i)} \\
&\quad \leqq \frac{c_1(\mu)}{n} + \frac{c_2(\mu) n^{p/2}(1 - q(\theta)(1 - e^{-c_3(\mu)}))^{n-k}}{1 - \sum_{i=0}^{j-1} p_\mu(i)} \leqq \frac{c_4(\mu)}{n}
\end{aligned}$$

where $c_4(\mu)$ is continuous in μ. Hence by the proof of Theorem 1

$$E_\mu[u(S_k) \mid \bar{X}_n, J \geqq j] = g(\bar{X}_n) + \mathcal{O}_p\left(\frac{1}{n}\right).$$

To complete the proof, consider the last term in (2.1). Since $E_\mu P_\mu[J < j \mid S_n] = P_\mu[J < j] = \sum_{i=0}^{j-1} p_\mu(i) = \mathcal{O}(n^j(1 - q(\theta))^n)$, $P_\mu[J < j \mid S_n] = \mathcal{O}_p(n^{-1})$. Similarly, since $E_\mu|w(S_n)| \leqq 2E_\mu|u(S_k)|$ (see (2.2)) $w(S_n) = \mathcal{O}_p(1)$; and the proof is complete. □

A remark about the theorems should be made: the condition that the unbiased estimator, t, be absolutely integrable for all $\mu \in D$ is not necessary. If

$$D_0 \equiv \{\mu \in D: E_\mu|t(X_1 + \cdots + X_k)| < +\infty\}$$

has nonempty interior, then the same proofs will hold at least on the interior of D_0.

Counterexamples are now given showing that the results of the theorems need not hold when Condition A fails. First, let ν be concentrated on the set $\{-1, 0, \pi\}$ with $\nu(-1) = \frac{1}{2}$, $\nu(\pi) = 1/2\pi$, and $\nu(0) = p_0 = 1 - \frac{1}{2} - 1/2\pi$. Let $D(\theta) = \frac{1}{2}e^{-\theta} + p_0 + (1/2\pi)e^{\pi\theta}$ so that $\psi(\theta) = \log D(\theta)$. This gives an exponential family on R not satisfying Condition A. The problem is to estimate $g(\theta) = P_\theta(X = 0) = p_0/D(\theta)$. The maximum likelihood estimator has asymptotic variance $(g'(\theta))^2/\psi''(\theta)$. But, if $N(0) = \#\{i: X_i = 0\}$ then $N(0)$ is a function of $\bar{X}$. Hence, the mvu estimator is

$$T_n(\bar{X}_n) = P\{X_1 = 0 \mid \bar{X}_n\} = P\{X_1 = 0 \mid N(0)\} = \frac{N(0)}{n} .$$

Therefore

$$n \operatorname{Var}_\theta T_n(\bar{X}_n) = g(\theta)(1 - g(\theta)) .$$

These asymptotic variances are different, as the following table clearly shows:

θ	-2.5	-1.0	$-.5$	0	$.5$	1.0	2.0
$(g'(\theta))^2/\psi''(\theta)$	.0500	.1103	.0488	.0000	.0388	.0353	.0029
$g(\theta)(1 - g(\theta))$	.0502	.1598	.2035	.2247	.1833	.0744	.0040

This counterexample can be generalized as follows: Let ν_1 be the distribution of the sum of independent random variables taking on the values $\pm r_n$ each with probability $\frac{1}{2}$. That is, ν_1 has characteristic function $\varphi(t) = \prod_{n=1}^{\infty} \cos(tr_n)$. Suppose $\{r_n\}$ are chosen so that the convolutions $\nu_1{}^{*k}$ are singular with respect to $\nu_1{}^{*j}$ for $j > k$. Let $\nu = \sum_{n=1}^{\infty} (\frac{1}{2})^n \nu_1{}^{*n}$. Then the exponential family generated by ν has

$$\psi(\theta) = \log\{\textstyle\sum_{m=1}^{\infty} (\frac{1}{2} \prod_{n=1}^{\infty} \cosh \theta r_n)^m\} .$$

Let U_k be the support $\nu_1{}^{*k}$ $(k = 1, 2, \cdots)$, and assume that $\{U_k\}$ are mutually disjoint (this can be done since $\{\nu_1{}^{*k}\}$ are mutually singular). If $\theta = 0$, the observations $\{X_i\}$ have distribution ν; so if (for $i = 1, 2, \cdots$) M_i is defined to be that index, m, such that $X_i \in U_m$, then $(M_i - 1)$ has a geometric distribution with parameter $\frac{1}{2}$ (under $\theta = 0$). Furthermore, defining L to be that index l such that $S_n \equiv \sum_{i=1}^{n} X_i \in U_l$, $L \sim \sum_{i=1}^{n} M_i$.

Now consider the problem of estimating $g(\theta) = P_\theta(X_1 \in U_1) = P_\theta(M_1 = 1)$. Since L is a function of S_n, the mvu estimator, t_n, satisfies (for any θ)

$$t_n(S_n) = P_\theta[M_1 = 1 \mid S_n] = P_\theta[M_1 = 1 \mid L] .$$

But, under $\theta = 0$, sums of $\{M_i - 1\}$ have negative binomial distributions; and hence,

$$P[M_1 = 1 \mid L = l] = \frac{P\{M_1 = 1, \sum_{i=2}^{n} M_i = l - 1\}}{P\{L = l\}} = \frac{n-1}{l-1}.$$

Therefore, $t_n(S_n) = (n-1)/(L-1)$ which is asymptotically equivalent to $1/(\bar{Z}_n + 1)$ where $\bar{Z}_n$ is the average of n i.i.d. geometric random variables. Thus, if $\theta = 0$ the asymptotic variance of the mvu estimator is $\frac{1}{8}$.

However,

$$g(\theta) = \int_{U_1} e^{-\psi(\theta)+\theta x}\, d\nu(x) = \tfrac{1}{2}e^{-\psi(\theta)}(\prod_{n=1}^{\infty} \cosh \theta r_n)$$

which is symmetric about zero. Thus, $g'(0) = 0$, and, again, the maximum likelihood estimator has asymptotic variance zero at $\theta = 0$.

These examples clarify the question of exactly what conditions may be placed on the distributions so that the results of the theorems hold. For example, one might hope for a sufficient condition like Cramér's condition that the characteristic function remain bounded away from zero. However, values for $\{r_n\}$ can be chosen so that the convolutions of ν_1 are singular and at the same time the characteristic function of ν_1 has arbitrary asymptotic behavior. For example, results in Salem (1942) and Kahane (1968, page 169) show how to construct $\{r_n\}$ for which $\nu_1{}^{*n}$ are appropriately singular and the characteristic function tends to zero. Also, the symmetrized Cantor distribution (with $r_n = 3^{-n}$) has $0 < \limsup_{t\to\infty} \varphi_{\nu_1}(t) < 1$. Since $\varphi_\nu(t) = \frac{1}{2}\varphi_{\nu_1}(t)/(1 - \frac{1}{2}\varphi_{\nu_1}(t))$, a rather general class of counterexamples can be constructed. It is interesting that for distributions, ν_1, generated by $\{r_n\}$ as above, Brown and Moran (1973) have shown that either the convolutions are singular (so a counterexample can be constructed) or some k-fold convolution is absolutely continuous (so Theorem 2 would hold). However, it is not known whether the exponential family generated by ν itself (for example, if ν is the symmetrized Cantor distribution) has an inefficient mvu estimator of a parametric function.

Acknowledgment. I wish to thank my colleague Robert Kaufman for his valuable assistance in finding the measures used to construct the last class of counterexamples. I also want to thank the referee for a number of helpful comments.

REFERENCES

Brown, G. and Moran, W. (1973). A dichotomy for infinite convolutions of discrete measures. *Proc. Cambridge Philos. Soc.* **73** 307–316.

Kahane, Jean-Pierre (1968). *Some Random Series of Functions.* Heath Mathematical Monograph, Lexington, Mass.

Salem, R. (1942). On sets of multiplicity for trigonometric series. *Amer. J. Math.* **64** 531–538.

Sharma, Divakar (1973). Asymptotic equivalence for two estimators for an exponential family. *Ann. Statist.* **1** 973–980.

ZABELL, SANDY (1974). A limit theorem for conditional expectations with applications to probability theory and statistical mechanics. Ph. D. thesis, Harvard Univ.

ZABELL, SANDY (1976). Rates of convergence for conditional expectations, I. Large deviation case; and II. General theorems. Unpublished.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801