Statistica Sinica 2(1992), 91-111

INTEGRABLE EXPANSIONS FOR POSTERIOR DISTRIBUTIONS FOR ONE-PARAMETER EXPONENTIAL FAMILIES

Michael Woodroofe

University of Michigan

Abstract: The main results provide asymptotic expansions for posterior distributions which may be integrated termwise with respect to the marginal distribution of the data. The proof uses a data dependent transformation which converts the likelihood function to exact normality and then applies a version of Stein's Identity to the posterior distributions. Applications to sequential confidence intervals are described briefly.

Key words and phrases: Posterior distributions, parameter transformations, Stein's Identity, martingale convergence theorem, stopping times, sequential confidence intervals.

1. Introduction

Asymptotic expansions for posterior distributions may be traced to the time of Laplace, but only recently have researchers investigated conditions under which these expansions may be integrated with respect to the marginal distribution of the data. See Johnson (1970) for a rigorous account of the pointwise expansions and Bickel, Goetze and Van Zwet (1985) and Ghosh, Sinha and Joshi (1983) for recent work on integrating them.

In this article, an alternative approach to the expansions is presented which makes the question of integrability more transparent. First, instead of renormalized estimation error, a data dependent transformation (3) of the parameter is considered, which converts the likelihood to exact normality. Then a version of Stein's Identity is applied to the posterior distributions to isolate the remainder terms. The alternative approach avoids the use of messy Taylor series expansions and leaves the renormalized remainder terms in the form of conditional expectations, so that the martingale theory may be brought to bear on the integrability question.

Integrable expansions for posterior distributions are needed in design problems where the overall Bayes risk must be computed in order to see the effect of design parameters. Surprisingly, they also have applications to the problem of setting confidence intervals after sequential testing. Such applications are described by Woodroofe (1986, 1989) and briefly in Section 7, below.

The model is detailed in Section 2. Conditions for mean convergence of functions of the maximum likelihood estimator in a Bayesian model are derived in Section 3. This material may be of independent interest. Stein's Identity is reviewed in Section 4 and applied to obtain second order expansions in Sections 5 and 6. Applications to sequential confidence intervals are discussed briefly in Section 7, and higher order expansions in Section 8.

2. The Model

Let Λ denote a nondegenerate, sigma-finite measure on the Borel sets of $\mathbf{R}$; let $\Omega = (\underline{\theta}, \overline{\theta})$ denote a nondegenerate, open interval (so that, $-\infty \le \underline{\theta} < \overline{\theta} \le \infty$); and let

$$f_\theta(x) = e^{\theta x - \psi(\theta)}, \qquad x \in \mathbf{R},\ \theta \in \Omega,$$

denote a one parameter exponential family of densities (w.r.t. Λ). The reader is assumed to be familiar with exponential families, as in Brown (1986, Ch.s 1 and 2), for example.

Let $X_1, X_2, \ldots$ denote random variables which are i.i.d. with common density f_θ under a probability measure P_θ for each $\theta \in \Omega$. Then the log likelihood function given $X_1, \ldots, X_n$ is

$$L_n(\theta) = \theta S_n - n\psi(\theta), \qquad \theta \in \Omega,$$

where $S_n = X_1 + \cdots + X_n$, $n \ge 1$. Let $\overline{X}_n = S_n/n$, for $n \ge 1$. If $\overline{X}_n \in \psi'(\Omega)$, then $L_n(\theta)$ attains its maximum at $\hat{\theta}_n$, where $\psi'(\hat{\theta}_n) = \overline{X}_n$; and then

$$L_n(\theta) - L_n(\hat{\theta}_n) = \theta S_n - n\psi(\theta) - [\hat{\theta}_n S_n - n\psi(\hat{\theta}_n)] = -nI(\hat{\theta}_n, \theta), \tag{2}$$

where $I(\omega, \theta) = \psi(\theta) - \psi(\omega) - \psi'(\omega)(\theta - \omega)$, $\omega, \theta \in \Omega$. Derivatives of I are needed below: for $\omega, \theta \in \Omega$, $I_{01}(\omega, \theta) = \psi'(\theta) - \psi'(\omega)$, $I_{10}(\omega, \theta) = -\psi''(\omega)(\theta - \omega)$, $I_{02}(\omega, \theta) = \psi''(\theta)$ and $I_{11}(\omega, \theta) = -\psi''(\omega)$, where $I_{jk}(\omega, \theta) = \partial^{j+k} I(\omega, \theta)/\partial\omega^j \partial\theta^k$ for $j, k = 0, 1, 2, \ldots$ and $\omega, \theta \in \Omega$.

The data dependent transformation mentioned above is

$$z = \sqrt{\{2nI(\hat{\theta}_n, \theta)\}}\,\mathrm{sign}(\theta - \hat{\theta}_n). \tag{3}$$

Observe that z is increasing in θ, since $dz/d\theta = \sqrt{n}|I_{01}|/\sqrt{2I} > 0$ and that the likelihood function is exactly normal in z; that is, $L_n(\theta) = L_n(\hat{\theta}_n) - \frac{1}{2}z^2$ for all $\theta \in \Omega$ and $n \ge 1$.

Now consider a Bayesian model in which there is a random variable Θ with density ξ and $X_1, X_2, \ldots$ are conditionally i.i.d. with common density f_θ given $\Theta = \theta$ for every $\theta \in \Omega$. Probability and expectation in the Bayesian model are denoted by P_ξ and E_ξ; P_θ and E_θ denote conditional probability and expectation given $\Theta = \theta$; and conditional expectation given $X_1, \ldots, X_n$ is denoted by E_ξ^n.

If $\overline{X}_n \in \psi'(\Omega)$, then the conditional density of Θ given $X_1, \ldots, X_n$ is $\xi_n(\theta) \propto \exp\{-nI(\hat{\theta}_n, \theta)\}\xi(\theta)$, $\theta \in \Omega$, by (2). Of course, n may be replaced by any stopping time t in the likelihood function and posterior density, since these are unaffected by optional stopping. See, for example, Berger and Walpole (1984, Section 4.2).

In the Bayesian model, the transformation (3) may be written as the random variable

$$Z_n = \sqrt{\{2nI(\hat{\theta}_n, \Theta)\}}\operatorname{sign}(\Theta - \hat{\theta}_n)$$

on $\{\overline{X}_n \in \psi'(\Omega)\}$. ($Z_n$ is undefined off of this event.) Let $\mathbf{R}_n$ denote the range of Z_n; that is $\mathbf{R}_n = (-\sqrt{\{2nI(\hat{\theta}_n, \underline{\theta})\}}, \sqrt{\{2nI(\hat{\theta}_n, \overline{\theta})\}})$, where $I(\hat{\theta}_n, \underline{\theta})$ and $I(\hat{\theta}_n, \overline{\theta})$ denote the limits of $I(\hat{\theta}_n, \theta)$ as $\theta \to \underline{\theta}$ and $\overline{\theta}$ (finite or infinite), still assuming that $\overline{X}_n \in \psi'(\Omega)$. Then the conditional density of Z_n given $X_1, \ldots, X_n$ is

$$\zeta_n(z) \propto \begin{cases} \xi(\theta)J(\hat{\theta}_n, \theta)e^{-\frac{1}{2}z^2} : & \text{if } z \in \mathbf{R}_n, \\ 0 : & \text{otherwise,} \end{cases}$$

where

$$J(\omega, \tau) = \frac{\sqrt{\{2I(\omega, \tau)\}}}{|I_{01}(\omega, \tau)|}, \quad \omega, \tau \in \Omega, \tag{4}$$

and z and θ are related by (3), since $dz/d\theta \propto 1/J(\hat{\theta}_n, \theta)$.

The partial derivatives of J are also needed below: for $\omega, \theta \in \Omega$ and $\theta \neq \omega$, they are

$$J_{01} = \frac{1}{\sqrt{2I}}\{1 - I_{02}J^2\}\operatorname{sign}(\theta - \omega),$$

and

$$J_{02} = -\left\{\frac{I_{01}}{2I}J_{01} + \frac{1}{\sqrt{2I}}\left[I_{03}J^2 + 2I_{02}JJ_{01}\right]\right\}\operatorname{sign}(\theta - \omega). \tag{5}$$

The values of J and its partial derivatives on the line $\omega = \theta$ may be obtained from L'Hospital's rule as

$$J(\omega,\omega)=\frac{1}{\sqrt{\psi''(\omega)}},\quad J_{01}(\omega,\omega)=-\frac{\psi'''(\omega)}{3\psi''(\omega)^{3/2}},$$

$$\text{and}\qquad J_{02}(\omega,\omega)=\frac{11}{36}\cdot\frac{\psi'''(\omega)^2}{\psi''(\omega)^{5/2}}-\frac{1}{4}\cdot\frac{\psi^{(4)}(\omega)}{\psi''(\omega)^{3/2}},\tag{6}$$

where ψ''' and $\psi^{(4)}$ denote the third and fourth derivatives of ψ.

3. Inequalities

At one point in the development, it is necessary to compute the limit of the expectation of a function of the M.L.E. in the Bayesian model. This is a question of independent interest. In its discussion below, let

$$B_n=\{\overline{X}_n\in\psi'(\Omega)\},\qquad n\geq 1.$$

Proposition 1. *For all $\theta,\omega\in\Omega$ and $m=1,2,\ldots,$*

$$P_\theta\{B_n \text{ and } \hat{\theta}_n>\omega,\exists\, n\geq m\}\leq e^{-mI(\omega,\theta)},\quad \text{if } \omega>\theta,$$
$$P_\theta\{B_n \text{ and } \hat{\theta}_n<\omega,\exists\, n\geq m\}\leq e^{-mI(\omega,\theta)},\quad \text{if } \omega<\theta.$$

Proof. If B_n occurs and $\hat{\theta}_n>\omega>\theta$ for some $n\geq 1$, then $\overline{X}_n>\psi'(\omega)$ for the same n. Let $t=m(\omega-\theta)$. Then, by the submartingale inequality (applied to the reverse submartingale $\exp(t\overline{X}_n)$, $n\geq m$),

$$P_\theta\Big\{\sup_{n\geq m}\overline{X}_n>\psi'(\omega)\Big\}\leq e^{-t\psi'(\omega)}E_\theta\Big\{e^{t\overline{X}_m}\Big\}$$
$$\leq\exp\Big\{m\big[\psi(\theta+\frac{t}{m})-\psi(\theta)\big]-t\psi'(\omega)\Big\}=\exp[-mI(\omega,\theta)].$$

This establishes the first assertion of the Proposition; and the second may be established similarly.

Corollary 1. *For all $\theta\in\Omega$, $x,z>0$, and $m=1,2,\ldots,$*

$$P_\theta\{B_n \text{ and } I(\hat{\theta}_n,\theta)>x,\exists\, n\geq m\}\leq 2e^{-mx}$$

and

$$P_\theta\{B_m \text{ and } |Z_m|>z\}\leq 2\exp\Big(-\frac{1}{2}z^2\Big).$$

In particular, all powers of Z_n, $n \geq 1$, are uniformly integrable under P_ξ for any ξ.

Proof. For a fixed θ, $I(\omega,\theta)$ is increasing in $\omega > \theta$, since $I_{10}(\omega,\theta) = \psi''(\omega)(\omega-\theta)$. For fixed $m \geq 1$ and $x > 0$, let A^+ be the event that B_n occurs and $I(\hat{\theta}_n,\theta) > x$ for some $n \geq m$. Then A^+ is empty if $x \geq I(\overline{\theta},\theta)$; and if $x < I(\overline{\theta},\theta)$, then $A^+ = \{B_n \text{ and } \hat{\theta}_n > \omega, \exists\, n \geq m\}$, where $I(\omega,\theta) = x$. So, $P_\theta(A^+) \leq e^{-mx}$, by the Proposition. The first assertion of the corollary follows easily from this and an obvious dual; and the others are immediate consequences.

Corollary 2. *For all $z \geq 1$, $\theta \in \Omega$, and $n = 1,2,\ldots$,*

$$P_\theta\Big\{\max_{k\leq n}|Z_k|I_{B_k} > z\Big\} \leq 2[1+\log_2 n]\exp\Big(-\frac{1}{4}z^2\Big),$$

where $\log_2$ *denotes logarithm to the base two.*

Proof. For each $k \geq 1$, there is a unique $m \geq 1$ for which $2^{m-1} \leq k < 2^m$. Moreover, if B_k occurs and $|Z_k| > z \geq 1$ for some $k \in [2^{m-1}, 2^m]$, then $I(\hat{\theta}_k,\theta) > z^2/2^{m+1}$ for some $k \geq 2^{m-1}$. Let M_n be the least integer which exceeds $\log_2 n$. Then, for all $\theta \in \Omega$ and $n \geq 1$,

$$\begin{aligned} P_\theta\Big\{\max_{k\leq n}|Z_k|I_{B_k} > z\Big\} &\leq \sum_{m=1}^{M_n} P_\theta\{B_k \text{ and } |Z_k| > z, \exists\, k \in [2^{m-1},2^m]\} \\ &\leq \sum_{m=1}^{M_n} P_\theta\{B_k \text{ and } I(\hat{\theta}_k,\theta) > \frac{z^2}{2^{m+1}}, \exists\, k \geq 2^{m-1}\} \\ &\leq 2M_n \exp\Big(-\frac{1}{4}z^2\Big). \end{aligned}$$

For a fixed density ξ, let G_ξ denote the class of all functions $G : \Omega\times\Omega \to \mathbf{R}$ for which

$$\int \sup_{n\geq m}|G(\hat{\theta}_n,\Theta)|I_{B_n}dP_\xi < \infty \tag{7}$$

for some $m = 1,2,\ldots$. Observe that G_ξ is a linear space which contains all constant functions; moreover, if $G \in G_\xi$ and $|H| \leq |G|$, then $H \in G_\xi$.

Proposition 2. *Let ξ be a density. If $G : \Omega\times\Omega \to \mathbf{R}$ is a continuously differentiable function for which*

$$\int_\Omega |G(\theta,\theta)|\xi(\theta)d\theta < \infty$$

and

$$\int_{\Omega\times\Omega} |G_{10}(\omega,\theta)| e^{-mI(\omega,\theta)} \xi(\theta) d\omega d\theta < \infty$$

for some $m \geq 1$, then $G \in G_\xi$.

Proof. There is no loss of generality in supposing that $G(\theta,\theta) = 0$ for all $\theta \in \Omega$. Let m be as in the statement of the Proposition. Then

$$\sup_{n\geq m} |G(\hat{\theta}_n, \Theta)| I_{B_n} \leq \int_{\Theta}^{Y^+} |G_{10}(\omega,\Theta)| d\omega + \int_{Y^-}^{\Theta} |G_{10}(\omega,\Theta)| d\omega, \qquad (8)$$

where Y^+ is the supremum of $\hat{\theta}_n$ over $n \geq m$ for which B_n occurs and Y^- is the infimum of $\hat{\theta}_n$ over $n \geq m$ for which B_n occurs. Now,

$$\begin{aligned} E_\xi\left\{\int_{\Theta}^{Y^+} |G_{10}(\omega,\Theta)| d\omega\right\} &= \int_\Omega \left\{\int_\theta^{\bar{\theta}} |G_{10}(\omega,\theta)| P_\theta\{Y^+ > \omega\} d\omega\right\} \xi(\theta) d\theta \\ &\leq \int_\Omega \left\{\int_\theta^{\bar{\theta}} |G_{10}(\omega,\theta)| e^{-mI(\omega,\theta)} d\omega\right\} \xi(\theta) d\theta, \end{aligned}$$

by Proposition 1; and the latter integral is finite, by assumption. The other integral in (8) may be analyzed similarly to complete the proof.

Corollary 3. *If ξ is a density, $g : \Omega \to \mathbf{R}$, and $|g|\xi$ is integrable, then $G(\omega,\theta) = g(\theta)e^{I(\omega,\theta)}$, $\omega, \theta \in \Omega$, defines an element of G_ξ.*

Proof. In this case, $G(\theta,\theta) = g(\theta)$ for all $\theta \in \Omega$; $\int_\Omega |g|\xi d\theta < \infty$, by assumption; and

$$\begin{aligned} &\iint_{\Omega\times\Omega} |G_{10}(\omega,\theta)| e^{-2I(\omega,\theta)} \xi(\theta) d\omega d\theta \\ &= \int_\Omega \left\{\int_\Omega |I_{10}(\omega,\theta)| e^{-I(\omega,\theta)} d\omega\right\} |g(\theta)| \xi(\theta) d\theta, \\ &\leq \int_\Omega 2|g(\theta)|\xi(\theta) d\theta < \infty, \end{aligned}$$

since $\int_\Omega |I_{10}| e^{-I} d\omega \leq \int_{-\infty}^{\infty} |x| e^{-x} dx = 2$ for all $\theta \in \Omega$.

4. Stein's Identity

Let H denote the collection of measurable functions $h : \mathbf{R} \to \mathbf{R}$ of polynomial growth; let Φ denote the standard normal distribution; and let

$$\Phi h = \int_{-\infty}^{\infty} h d\Phi$$

and

$$Uh(z) = e^{\frac{1}{2}z^2} \int_z^\infty [h(y) - \Phi h] e^{-\frac{1}{2}y^2} dy, \quad z \in \mathbf{R},$$

for $h \in H$. Then U is a linear transformation from H back into itself (See Lemma 1 below). For example, if $h_1(z) = z$ and $h_2(z) = z^2$ for $z \in \mathbf{R}$, then $Uh_1(z) = 1$ and $Uh_2(z) = z$ for all $z \in \mathbf{R}$.

Let H_p denote the collection of all $h \in H$ for which $|h(z)| \leq 1 + |z|^p$ for all $z \in \mathbf{R}$; and let $\tilde{H}_p$ denote the class of all $h \in H$ for which $h/c \in H_p$ for some $0 < c < \infty$.

Lemma 1. *There are (finite) positive constants $c_0, c_1, c_2, \ldots$ for which $UH_0 \subseteq c_0 H_0$ and $UH_p \subseteq c_p H_{p-1}$ for all $p = 1, 2, \ldots$.*

Proof. For $p = 0$, the assertion is proved by Stein (1987). For $p \geq 1$, it follows from the identity,

$$e^{\frac{1}{2}z^2} \int_z^\infty y^p e^{-\frac{1}{2}y^2} dy = z^{p-1} + (p-1) e^{\frac{1}{2}z^2} \int_z^\infty y^{p-2} e^{-\frac{1}{2}y^2} dy, \tag{9}$$

for $z > 0$ and a simple induction.

Lemma 2.

$$\Phi U h = \int_{-\infty}^\infty z h(z) \Phi(dz)$$

and

$$\Phi U^2 h = \frac{1}{2} \int_{-\infty}^\infty (z^2 - 1) h(z) \Phi(dz), \quad \forall h \in H.$$

The simple proof of Lemma 2 is omitted.

In the next result, Γ denotes a finite signed measure of the form $d\Gamma = f d\Phi$, where f is a real valued measurable function which is integrable with respect to Φ; and $\Gamma h = \int_{\mathbf{R}} h d\Gamma$, whenever $h \in H$ and the integral exists.

The following result is similar to one exploited by Stein (1987).

Stein's Identity. *Let $p \geq 0$ be an integer. If $d\Gamma = f d\Phi$, where f is absolutely continuous on (every compact subinterval of) $\mathbf{R}$ and*

$$\int_{-\infty}^\infty |z^p f'(z)| \Phi(dz) < \infty, \tag{10}$$

then

$$\Gamma h - \Gamma 1 \times \Phi h = \int_{-\infty}^\infty U h \times f' d\Phi, \quad \forall h \in \tilde{H}_{p+1}.$$

Proof. There is no loss of generality in supposing that $\Phi h = 0$ and $h \in H_{p+1}$, since otherwise h may be replaced by $(h-\Phi h)/c$ for an appropriate c. If $h \in H_{p+1}$, then $Uh \in c_{p+1}H_p$, so that $|Uh \times f'|$ is integrable (w.r.t. Φ), by (10). Now,

$$\begin{aligned}\int_0^\infty Uh \times f' d\Phi &= \frac{1}{\sqrt{2\pi}} \int_0^\infty \left[\int_z^\infty h(y) e^{-\frac{1}{2}y^2} dy\right] f'(z) dz \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty \left[\int_0^y f'(z) dz\right] h(y) e^{-\frac{1}{2}y^2} dy \\ &= \int_0^\infty h(y)[f(y) - f(0)]\Phi(dy),\end{aligned}$$

where the interchange of integration is easily justified using (9) and (10). A similar argument may be applied to the integral over $(-\infty, 0]$; and it follows that

$$\int_{-\infty}^\infty Uh \times f' d\Phi = \int_{-\infty}^\infty h(y)[f(y) - f(0)]\Phi(dy) = \int_{-\infty}^\infty hf d\Phi = \Gamma h.$$

5. Consequences

Observe that if B_n occurs, then the posterior distribution, Γ_n say, of Z_n given $X_1, \ldots, X_n$ is of the form $d\Gamma_n = f_n d\Phi$, where

$$f_n(z) = \frac{1}{c_n} \xi(\theta) J(\hat{\theta}_n, \theta) I_{\mathbf{R}_n}(z), \quad z \in \mathbf{R}, \tag{11}$$

where $0 < c_n = c_n(X_1, \ldots, X_n) < \infty$, $\mathbf{R}_n$ denotes the range of Z_n, and z and θ are related by (3).

Some restrictions on the prior density are needed to exploit Stein's Identity. Let Ξ denote the collection of all continuous densities ξ on Ω for which $\xi(\theta) \to 0$ as either $\theta \to \underline{\theta}$ or $\theta \to \overline{\theta}$; let Ξ_0 denote the class of all $\xi \in \Xi$ with compact support in Ω; let AC denote the collection of all absolutely continuous functions on Ω; and let

$$\Xi_1 = \left\{\xi \in \Xi \mid \xi \in AC \text{ and } \int_\Omega \left|\frac{\xi'}{\xi}\right|^\alpha \xi d\theta < \infty, \exists\, \alpha > 1\right\},$$

and

$$\Xi_2 = \left\{\xi \in \Xi \mid \xi' \in AC \text{ and } \int_\Omega \left|\frac{\xi''}{\xi}\right|^\alpha \xi d\theta < \infty, \exists\, \alpha > 1\right\}.$$

To simplify some of the formulas below, it is convenient to let

$$K_1^\xi(\omega, \theta) = \left(\frac{\xi'}{\xi}\right)(\theta) J(\omega, \theta) + J_{01}(\omega, \theta)$$

and

$$K_2^\xi(\omega,\theta) = \Big(\frac{\xi''}{\xi}\Big)(\theta)J^2(\omega,\theta) + 3\Big(\frac{\xi'}{\xi}\Big)(\theta)J(\omega,\theta)J_{01}(\omega,\theta) + J(\omega,\theta)J_{02}(\omega,\theta) + J_{01}^2(\omega,\theta)$$

for $\omega,\theta \in \Omega$, $\xi \in \Xi_1$ and $\xi \in \Xi_2$, where $0/0 = 0$. The values of K_1^ξ and K_2^ξ on the line $\omega = \theta$ are especially important; they are

$$\tilde{K}_1^\xi(\theta) = K_1^\xi(\theta,\theta) = \frac{1}{\sqrt{\psi''(\theta)}} \cdot \frac{\xi'}{\xi}(\theta) - \frac{\psi'''(\theta)}{3\psi''(\theta)^{3/2}}$$

and

$$\tilde{K}_2^\xi(\theta) = K_2^\xi(\theta,\theta) = \frac{1}{\psi''(\theta)} \cdot \frac{\xi''}{\xi}(\theta) - \frac{\psi'''(\theta)}{\psi''(\theta)^2} \cdot \frac{\xi'}{\xi}(\theta) + \frac{15}{36} \cdot \frac{\psi'''(\theta)^2}{\psi''(\theta)^3} - \frac{1}{4} \cdot \frac{\psi^{(4)}(\theta)}{\psi''(\theta)^2}, \quad \theta \in \Omega.$$

Lemma 3. *If $\xi \in \Xi_1$ and B_n occurs, then the function f_n defined in* (11) *is absolutely continuous with*

$$f_n'(z) = \frac{1}{\sqrt{n}} K_1^\xi(\hat{\theta}_n,\theta) f_n(z), \quad z \in \mathbf{R}.$$

If $\xi \in \Xi_0 \cap \Xi_2$ and B_n occurs, then f_n' is absolutely continuous with

$$f_n''(z) = \frac{1}{n} K_2^\xi(\hat{\theta}_n,\theta) f_n(z), \quad z \in \mathbf{R}.$$

Proof. Suppose first that $\xi \in \Xi_1$ and B_n occurs. Then it is clear that f_n is absolutely continuous on every compact interval of $\mathbf{R}_n$ with the claimed derivative. So, it suffices to show that $f_n(z)$ approaches zero as z approaches an endpoint of $\mathbf{R}_n$. If $I(\hat{\theta}_n,\overline{\theta}) = \infty$, then there is no right endpoint; and if $I(\hat{\theta}_n,\overline{\theta}) < \infty$, then $J(\hat{\theta}_n,\theta)$ remains bounded as $\theta \to \overline{\theta}$ and, therefore, $f_n(z) \to 0$ as z approaches the right endpoint. The left endpoint, if any, may be handled similarly.

The proof of the second assertion is similar and simpler.

Recall that G_ξ denotes the linear space of all functions $G : \Omega \times \Omega \to \mathbf{R}$ for which (7) holds.

Proposition 3. *If $\xi \in \Xi_0 \cap \Xi_1$ then $|K_1^\xi|^\alpha \in G_\xi$ for some $1 < \alpha < \infty$; and if $\xi \in \Xi_0 \cap \Xi_2$, then $|K_2^\xi|^\alpha \in G_\xi$ for some $1 < \alpha < \infty$.*

Proof. Let Ω_ξ denote the compact support of $\xi \in \Xi_0 \cap \Xi_1$; and let Ω_0 denote another compact set for which $\Omega_\xi \subseteq \Omega_0^o \subseteq \Omega_0 \subseteq \Omega$, where Ω_0^o denotes the interior of Ω_0. Then J and $|J_{01}|$, are bounded on $\Omega_0 \times \Omega_\xi$, by continuity and compactness; I_{02} is bounded on $\Omega_0' \times \Omega_\xi$, by compactness; and $1/I$ and $1/|I_{01}|$ are bounded on $\Omega_0' \times \Omega_{\xi'}$ by convexity and monotonicity. So, there are constants C' and C'' for which

$$J \leq C'(1+I) \quad \text{and} \quad |J_{01}| \leq C''(1+I) \tag{12}$$

on $\Omega \times \Omega_\xi$, by (4) and (5). So, $|(\xi'/\xi)J|^\alpha \in G_\xi$ and $|J_{01}|^\alpha \in G_\xi$ for some $1 < \alpha < \infty$, by Corollary 3 and, therefore, $|K_1^\xi|^\alpha \in G_\xi$ for some $1 < \alpha < \infty$. Similarly, if $\xi \in \Xi_0 \cap \Xi_2$ then $|(\xi''/\xi)J^2|^\alpha \in G_\xi$, $|(\xi'/\xi)JJ_{01}|^\alpha \in G_{\xi'}$ and $J_{01}^2 \in G_\xi$ for some $1 < \alpha < \infty$, by (12). The term JJ_{02} requires more care. Since JJ_{02} is bounded on $\Omega_0 \times \Omega_\xi$, there is a constant C for which

$$\begin{aligned}|JJ_{02}| &\leq J\left|\frac{I_{01}}{2I}J_{01} + \frac{1}{\sqrt{2I}}(I_{03}J^2 + 2I_{02}JJ_{01})\right| \\ &\leq C\{1 + |J_{01}| + J^3 + J^2|J_{01}|\}\end{aligned}$$

on $\Omega \times \Omega_\xi$. So, $|K_2^\xi|^\alpha \in G_\xi$ for some $1 < \alpha < \infty$, by (12) and Corollary 3.

Proposition 4. *Suppose that $\xi \in \Xi_1$ and that $K_1^\xi \in G_\xi$. Then*

$$E_\xi^n\{h(Z_n)\} = \Phi h + \frac{1}{\sqrt{n}}E_\xi^n\{K_1^\xi(\hat{\theta}_n, \Theta)Uh(Z_n)\} \tag{13}$$

(simultaneously) for all $h \in H_1$ a.e. (P_ξ) on B_n for all $n \geq m$ for some $m \geq 1$; and if $|K_1^\xi|^\alpha \in G_\xi$ for some $1 < \alpha < \infty$, then (13) holds for all $h \in H$. If $\xi \in \Xi_0 \cap \Xi_2$, then

$$E_\xi^n\{h(Z_n)\} = \Phi h + \frac{1}{\sqrt{n}}(\Phi Uh)E_\xi^n\{K_1^\xi(\hat{\theta}_n, \Theta)\} + \frac{1}{n}E_\xi^n\{K_2^\xi(\hat{\theta}_n, \Theta)U^2h(Z_n)\} \tag{14}$$

for all $h \in H$ a.e. (P_ξ) on B_n for $n \geq m$, for some $m \geq 1$, where $U^2 = U \circ U$.

Proof. For (13), let m be as in (7) with $G = K_1^\xi$, or $G = |K_1^\xi|^\alpha$, and let f_n be as in (11). If $n \geq m$ and B_n occurs, then

$$\begin{aligned}\sqrt{n}\int_{-\infty}^{\infty} |z|^p|f_n'(z)|\Phi(dz) &= \int_{-\infty}^{\infty} |z|^p|K_1^\xi(\hat{\theta}_n, \theta)|\Gamma_n(dz) \\ &= E_\xi^n\{|Z_n|^p|K_1^\xi(\hat{\theta}_n, \Theta)|\},\end{aligned}$$

where θ and z are related by (3), for all $0 \leq p < \infty$. If $p = 0$, the last line is finite w.p.1, since $K_1^\xi \in G_\xi$; and if $|K_1^\xi|^\alpha \in G_\xi$ for some $\alpha > 1$, then it is finite w.p.1 for all $1 \le p < \infty$, by Hölder's Inequality and Corollary 2. In either case,

$$\begin{aligned} E_\xi^n\{h(Z_n) - \Phi h\} &= \int_{-\infty}^{\infty} h(z)\Gamma_n(dz) - \Phi h \\ &= \int_{-\infty}^{\infty} Uh(z) f_n'(z)\Phi(dz) \qquad (15) \\ &= \frac{1}{\sqrt{n}} E_\xi^n\{K_1^\xi(\hat{\theta}_n, \Theta) Uh(Z_n)\} \end{aligned}$$

for all $h \in \tilde{H}_{p+1}$, a.e. on B_n by Stein's Identity, establishing (13).

For (14), one first verifies that $\int_{-\infty}^{\infty} |z|^p |f_n''(z)|\Phi(dz) < \infty$ a.e. on B_n for all $p = 1, 2, \dots$, as above. Then

$$\begin{aligned} &\frac{1}{\sqrt{n}} E_\xi^n\{K_1^\xi(\hat{\theta}_n, \Theta) Uh(Z_n)\} - \frac{1}{\sqrt{n}}(\Phi Uh) E_\xi^n\{K_1^\xi(\hat{\theta}_n, \Theta)\} \\ &= \int_{-\infty}^{\infty} Uh(z) f_n'(z)\Phi(dz) - (\Phi Uh)\int_{-\infty}^{\infty} f_n'(z)\Phi(dz) \\ &= \int_{-\infty}^{\infty} U^2h(z) f_n''(z)\Phi(dz) = \frac{1}{n} E_\xi^n\{K_2^\xi(\hat{\theta}_n, \Theta) U^2 h(Z_n)\} \end{aligned}$$

for all $h \in H_p$ a.e. on B_n for $p = 1, 2, \dots$.

A definition is needed for the last result of this section. Let $Y_1, Y_2, \dots$ denote a sequence of random variables adapted to a family of sigma-algebras, $A_1, A_2, \dots$, on a probability space (X, A, P); and let T denote the collection of all finite stopping times τ w.r.t. $A_1, A_2, \dots$; then $Y_1, Y_2, \dots$ is said to be *nearly dominated* iff $\{Y_\tau : \tau \in T\}$ is uniformly integrable. Observe that then $E|Y_\tau|$, $\tau \in T$, is bounded and $\sup_{n \ge 1} |Y_n| < \infty$ w.p.1. If $W_1, W_2, \dots$ are random variables for which $E\{\sup_{n\ge 1}|W_n|\} < \infty$, then $Y_n = E(W_n|A_n)$, $n \ge 1$, is nearly dominated, since $|Y_n| \le E(\sup_{k \ge 1}|W_k||A_n)$ for all $n = 1, 2, \dots$. Also, if $1 < \alpha, \beta < \infty$ are conjugate values $(1/\alpha + 1/\beta = 1)$ and if $|W_n|^\alpha$, $n \ge 1$, and $|Y_n|^\beta$, $n \ge 1$, are both nearly dominated, then so is $W_n Y_n$, since $\int_A |W_\tau Y_\tau| dP \le \{\int_A |W_\tau|^\alpha dP\}^{1/\alpha} \times \{\int_A |Y_\tau|^\beta dP\}^{1/\beta}$ for all stopping times τ and events A.

This definition is used below with $A_n = \sigma\{X_1, \dots, X_n\}$, $n = 1, 2, \dots$ and $P = P_\xi$. For the definition of the essential supremum of a family of random variables, needed below, see, for example, Chow, Robbins and Siegmund (1971, p.8).

Proposition 5. *If* $1 \le p < \infty$ *and* $|K_1^\xi|^p \in G_\xi$, *then the sequence*

$$\operatorname*{ess\,sup}_{h \in H_p} \sqrt{n}|E_\xi^n[h(Z_n)] - \Phi h| I_{B_n}, \quad n \ge m,$$

is nearly dominated for some $m \geq 1$.

Proof. The following inequality is used: if $0 < b,\ c < \infty$, $1 \leq p < \infty$, and $0 \leq x \leq b + cx^{(p-1)/p}$, then $x \leq pb + c^p$. It is easily verified by constructing a tangent to the curve $y = b + cx^{(p-1)/p}$ at $x = c^p$. Let m be as in (7) with $G = |K_1^\xi|^p$. If $h \in H_p$, then, by Proposition 4,

$$\begin{aligned}\sqrt{n}|E_\xi^n[h(Z_n) - \Phi h]| &= |E_\xi^n\{K_1^\xi(\hat{\theta}_n, \Theta)Uh(Z_n)\}| \\ &\leq E_\xi^n\{|K_1^\xi(\hat{\theta}_n, \Theta)|^p\}^{\frac{1}{p}} E_\xi^n\left\{|Uh(Z_n)|^{\frac{p}{p-1}}\right\}^{\frac{p-1}{p}},\end{aligned}$$

a.e. on B_n by Proposition 4 and Hölder's Inequality. If $p = 1$, then the last factor is to be interpreted as the maximum of $|Uh(z)|$; and since this is at most $2c_1$, by Lemma 1, the Proposition follows easily. For $1 < p < \infty$, let $g_p(z) = 1 + |z|^p$, for $z \in \mathbf{R}$. Then

$$|Uh(z)|^{\frac{p}{p-1}} \leq (2c_p)^{\frac{p}{p-1}} g_p(z)$$

for all $z \in \mathbf{R}$ and all $h \in H_p$, by a simple application of Lemma 1. It follows that, a.e. on B_n,

$$\operatorname*{ess\,sup}_{h \in H_p} |E_\xi^n[h(Z_n) - \Phi h| \leq \frac{1}{\sqrt{n}} 2c_p E_\xi^n\{|K_1^\xi(\hat{\theta}_n, \Theta)|^p\}^{\frac{1}{p}} E_\xi^n\{g_p(Z_n)\}^{\frac{p-1}{p}}.$$

Let $b = \Phi g_p$, $c = (2c_p/\sqrt{n})E_\xi^n\{|K_1^\xi(\hat{\theta}_n, \Theta)|^p\}^{1/p}$, and $x = E_\xi^n[g_p(Z_n)]$. Then $x \leq b + cx^{(p-1)/p}$ a.e. on B_n. So, a.e. on B_n,

$$E_\xi^n[g_p(Z_n)] \leq pb + c^p = p\Phi g_p + \left(\frac{2}{\sqrt{n}} c_{p-1}\right)^p E_\xi^n\{|K_1^\xi(\hat{\theta}_n, \Theta)|^p\};$$

and, therefore, $E_\xi^n[g_p(Z_n)]I_{B_n}$, $n \geq m$, are nearly dominated. So,

$$\begin{aligned}&\operatorname*{ess\,sup}_{h \in H_p} \sqrt{n}|E_\xi^n[h(Z_n)] - \Phi h| I_{B_n} \\ &\leq 2c_p E_\xi^n\{|K_1^\xi(\hat{\theta}_n, \Theta)|^p\}^{\frac{1}{p}} E_\xi^n\{[g_p(Z_n)]\}^{\frac{p-1}{p}} I_{B_n}, \quad n \geq m,\end{aligned}$$

are nearly dominated (by Hölder's Inequality, as above).

Corollary 4. $\operatorname*{ess\,sup}_{h \in H_p} |E_\xi^n[h(Z_n)]| I_{B_n}$, *$n \geq m$, are nearly dominated.*

6. Second Order Expansions

In the Theorem below, ξ denotes a fixed member of Ξ_1; and $t = t_a$, $a \geq 1$, denotes an increasing family of stopping times, w.r.t. $A_n = \sigma\{X_1, \ldots, X_n\}$,

$n = 1, 2, \ldots$, for which there exist $0 < \eta < \infty$ and a positive function ρ for which

$$P_\theta\{\overline{X}_t \in \psi'(\Omega)\} = 1, \quad \forall a \geq 1,\ \theta \in \Omega, \tag{16}$$

$$\frac{a}{t_a} \to \rho^2(\theta) \text{ in } P_\theta\text{-probability} \quad a.e.\ \theta, \tag{17}$$

and

$$P_\xi\{t_a \leq \eta a\} = o(a^{-q}) \tag{18}$$

as $a \to \infty$ for some $q > 1/2$. Observe that (18) holds for every $\xi \in \Xi_0$ if for every compact $\Omega_0 \subseteq \Omega$, there is an $\eta = \eta(\Omega_0)$ for which

$$\int_{\Omega_0} P_\theta\{t_a \leq \eta\} d\theta = o(a^{-q}), \quad \text{as } a \to \infty. \tag{19}$$

Also observe that (17) and (18) require ρ to be essentially bounded on the support of ξ.

For the theorem below, observe that $E_\xi|Z_t|^p < \infty$ for any stopping time t, if $\xi \in \Xi_1$ and $|K_1^\xi|^p \in G_\xi$, by Corollary 4 and let

$$R_{1,a}(\xi; h) = \sqrt{a}\{E_\xi^t[h(Z_t)] - \Phi h\} - (\Phi U h) E_\xi^t[\rho(\Theta)\tilde{K}_1^\xi(\Theta)]$$

for $h \in H_p$ and $a \geq 1$.

Theorem 1. *Suppose that $\xi \in \Xi_1$ and $|K_1^\xi|^p \in G_\xi$, where $1 \leq p < \infty$. If $t = t_a$ are stopping times for which* (16), (17) *and* (18) *hold, then*

$$\lim_{a \to \infty} E_\xi\{\operatorname{ess\ sup}_{h \in H_p} |R_{1,a}(\xi; h)|\} = 0.$$

Proof. First observe that there is a constant C for which

$$\begin{aligned}
&\int_{\{t \leq \eta a\}} \operatorname{ess\ sup}_{h \in H_p} |R_{1,a}(\xi; h)| dP_\xi \\
&\leq C\sqrt{a} \int_{\{t \leq \eta a\}} (1 + |Z_t|^p) dP_\xi + C \int_{\{t \leq \eta a\}} \rho(\Theta)|\tilde{K}_1^\xi(\Theta)| dP_\xi
\end{aligned}$$

for all $a \geq 1$. The final integral here approaches zero as $a \to \infty$, by (17) and (18). For the first, let η and q be as in (18); let $1 < \alpha < 2q$; and let $\beta = \alpha/(\alpha - 1)$ denote the conjugate value. Then

$$\sqrt{a} \int_{\{t \leq \eta a\}} (1 + |Z_t|^p) dP_\xi \leq \sqrt{a} P_\xi\{t \leq \eta a\}^{1/\alpha} E_\xi\Big\{\max_{k \leq \eta a} (1 + |Z_k|^p)^\beta I_{B_k}\Big\}^{1/\beta}$$

which is of order $\sqrt{a} \cdot a^{-q/\alpha} \cdot \log^p a = o(1)$, by (18) and Corollary 2.

To estimate the integral over $\{t > \eta a\}$, it is convenient to write (using Proposition 4)

$$R_{1,a}(\xi; h) = \mathrm{I}_a + \mathrm{II}_a + \mathrm{III}_a,$$

where

$$\mathrm{I}_a = \sqrt{(\frac{a}{t})}E_\xi^t\Big\{[K_1^\xi(\hat\theta_t, \Theta) - E_\xi^t[\tilde K_1^\xi(\Theta)]]Uh(Z_t)\Big\}$$

$$\mathrm{II}_a = \sqrt{(\frac{a}{t})}E_\xi^t\{\tilde K_1^\xi(\Theta)\}E_\xi^t\{Uh(Z_t) - (\Phi Uh)\}$$

$$\mathrm{III}_a = (\Phi Uh)E_\xi^t\Big\{\tilde K_1^\xi(\Theta)\Big[\sqrt{(\frac{a}{t})} - \rho(\Theta)\Big]\Big\}.$$

In the analysis of these terms write $\mathrm{I}_a^\# = \text{ess}\sup_{h\in H_p} \mathrm{I}_a \times I_{\{t>\eta a\}}$, etc.. It is clear from (17) and (18) that $[\sqrt{(\frac{a}{t})} - \rho(\Theta)]I_{\{t>\eta a\}} \to 0$ in probability as $a \to \infty$ and that $|\sqrt{(\frac{a}{t})} - \rho(\Theta)|I_{\{t>\eta a\}} \le 1/\sqrt{\eta} + \rho(\Theta)$, which is essentially bounded. So,

$$\int \mathrm{III}_a^\# dP_\xi \le C\int_{\{t>\eta a\}} |\tilde K_1^\xi(\Theta)|\Big|\sqrt{(\frac{a}{t})} - \rho(\Theta)\Big|dP_\xi \to 0$$

as $a \to \infty$, by the dominated convergence theorem. It is clear from Proposition 5 that $\mathrm{II}_a^\# \to 0$ in P_ξ-probability as $a \to \infty$. If $p = 1$, then $|E_\xi^t[Uh(Z_t) - \Phi Uh]|$, $a \ge 1$, are bounded, by Lemma 1, so that $\mathrm{II}_a^\#$, $a \ge 1$, are uniformly integrable; and if $1 < p < \infty$, then $\mathrm{II}_a^\#$, $a \ge 1$, are uniformly integrable by Hölder's Inequality and Corollary 4. In either case, $\lim_{a\to\infty} E_\xi(\mathrm{II}_a^\#) \to 0$. For the first term, it is easily seen that

$$\int \mathrm{I}_a^\# dP_\xi \le \frac{c_p}{\sqrt{\eta}}E_\xi\{|K_1^\xi(\hat\theta_t, \Theta) - E_\xi^t[\tilde K_1(\Theta)]|^p\}^{\frac{1}{p}} \times E_\xi\{(1 + |Z_t|^{p-1})^{\frac{p}{p-1}}\}^{\frac{p-1}{p}},$$

where the right most factor is to be interpreted as 2 if $p = 1$. Then the right most factor remains bounded as $a \to \infty$, by Corollary 4; and

$$\begin{aligned}&E_\xi\{|K_1^\xi(\hat\theta_t, \Theta) - E_\xi^t[\tilde K_1(\Theta)]|^p\}^{\frac{1}{p}}\\ \le &E_\xi\{\sup_{n\ge\eta a} |K_1^\xi(\hat\theta_n, \Theta) - \tilde K_1^\xi(\Theta)|^p \mathrm{I}_{B_n}\}^{\frac{1}{p}} + E_\xi\{|E_\xi^t[\tilde K_1^\xi(\Theta)] - \tilde K_1^\xi(\Theta)|^p\}^{\frac{1}{p}},\end{aligned}$$

which approach zero as $a \to \infty$, by the dominated and martingale convergence theorems.

For the corollaries below, suppose that t_a, $a \ge 1$, satisfy (16), (17) and (19) for some $q > 1/2$ and that ρ is absolutely continuous on (all compact subsets of) Ω. Then the conditions of Theorem 1 are satisfied with $p = 1$ for every $\xi \in \Xi_0 \cap \Xi_1$. Let

$$\kappa_1 = \frac{1}{6} \times \frac{\psi'''}{\psi''^{3/2}} - \frac{1}{\sqrt{\psi''}} \times \frac{\rho'}{\rho}$$

and

$$\overline{\kappa}_1(\xi) = \int_\Omega \kappa_1(\theta)\rho(\theta)\xi(\theta)d\theta.$$

Observe that κ_1 does not depend on ξ.

Corollary 5. *For all $\xi \in \Xi_0 \cap \Xi_1$,*

$$\lim_{a\to\infty} E_\xi\Big\{ \operatorname*{ess\,sup}_{h\in H_1} \sqrt{a}\big|E_\xi^t[h(Z_t) - \Phi h] - \frac{1}{\sqrt{a}}(\Phi U h)\overline{\kappa}_1(\xi)\big|\Big\} = 0. \tag{20}$$

Proof. It is clear from Theorem 1 and the martingale convergence theorem that (20) holds with $\overline{\kappa}_1(\xi)$ replaced by $E_\xi\{\rho(\Theta)\tilde{K}_1^\xi(\Theta)\}$; and

$$\begin{aligned} E_\xi\{\rho(\Theta)\tilde{K}_1^\xi(\Theta)\} &= \int_\Omega \rho(\theta)\Big[\frac{1}{\sqrt{\psi''(\theta)}}\cdot\xi'(\theta) - \frac{1}{3}\cdot\frac{\psi'''(\theta)}{\psi''(\theta)^{3/2}}\xi(\theta)\Big]d\theta \\ &= -\int_\Omega \Big[\Big(\frac{\rho}{\sqrt{\psi''}}\Big)'(\theta) + \frac{1}{3}\cdot\frac{\psi'''(\theta)}{\psi''(\theta)^{3/2}}\rho(\theta)\Big]\xi(\theta)d\theta \\ &= \int_\Omega \kappa_1(\theta)\rho(\theta)\xi(\theta)d\theta = \overline{\kappa}_1(\xi) \end{aligned}$$

by a simple integration by parts.

One nice feature of the transformation (3) is that the correction terms in the asymptotic expansion may be described by rescaling. For $z \in \mathbf{R}$, $\theta \in \Omega$ and $a \geq 1$, let

$$F_{a,\theta}(z) = P_\theta\{Z_t \leq z\}$$

and

$$\Phi_{a,\theta}^{(1)}(z) = \Phi\Big[z - \frac{1}{\sqrt{a}}\rho(\theta)\kappa_1(\theta)\Big].$$

Corollary 6. *For all $\xi \in \Xi_0 \cap \Xi_1$,*

$$\lim_{a\to\infty} \sup_{h\in H_1} \Big| \int_\Omega \sqrt{a}\Big[F_{a,\theta}h - \Phi_{a,\theta}^{(1)}h\Big]\xi(\theta)d\theta\Big| = 0. \tag{21}$$

Proof. Since $\int_\Omega F_{a,\theta}h\xi(\theta)d\theta = E_\xi[h(Z_t)]$, the integral on the left side of (21) is the sum of

$$\mathrm{I}_a = \sqrt{a}E_\xi[h(Z_t) - \Phi h] - (\Phi U h)\overline{\kappa}_1(\xi)$$

and

$$\mathrm{II}_a = -\int_\Omega \left\{\sqrt{a}[\Phi^{(1)}_{a,\theta}h - \Phi h] - (\Phi U h)\kappa_1(\theta)\rho(\theta)\right\}\xi(\theta)d\theta$$

for each $h \in H_1$ and $a \geq 1$. Here $\mathrm{I}_a \to 0$ uniformly in $h \in H_1$ as $a \to \infty$, by Corollary 5. Moreover, the integrand in II_a is at most

$$\left|\int_{\mathbf{R}} \sqrt{a}\left\{\varphi\left[z - \frac{1}{\sqrt{a}}\rho(\theta)\kappa_1(\theta)\right] - \varphi(z)\right\}h(z)dz - \rho(\theta)\kappa_1(\theta)\int_{\mathbf{R}} zh(z)\varphi(z)dz\right|$$
$$\leq \int \left|\sqrt{a}\left\{\varphi\left[z - \frac{1}{\sqrt{a}}\rho(\theta)\kappa_1(\theta)\right] - \varphi(z)\right\} - \rho(\theta)\kappa_1(\theta)z\varphi(z)\right|(1 + |z|)dz,$$

which is independent of $h \in H_1$, approaches zero as $a \to \infty$ for each θ, and is bounded by a constant multiple of $1 + |\rho'|$. So, $\mathrm{II}_a \to 0$ uniformly in $h \in H_1$, by the dominated convergence theorem, to complete the proof.

Corollary 6 does not assert that $F_{a,\theta}(z) - \Phi^{(1)}_{a,\theta}(z) = o(1/\sqrt{a})$ as $a \to \infty$ for any fixed θ or z. Rather, $\Phi^{(1)}_{a,\theta}(z)$ may be regarded as an approximation to appropriate averages of $F_{a,\omega}(z)$ in small neighborhoods about θ. Alternatively, the integral in (21) defines a linear functional on (the linear span of) $\Xi_0 \cap \Xi_1$, and (21) may be regarded as a form of weak convergence of the integrands.

7. Estimation after Sequential Testing

One of the more interesting applications of Theorem 1 is to approximate the sampling distributions of maximum likelihood estimators in sequential problems. To understand the nature of the difficulty consider the simple problem in which $X_1, X_2, \ldots$ are i.i.d. $N(\theta, 1)$, where $0 < \theta < \infty$ is unknown, and

$$t = t_a = \inf\{n \geq 1 : S_n > a\}.$$

Then $\hat{\theta}_t = \overline{X}_t$ and $Z_t = \sqrt{t}(\Theta - \overline{X}_t)$ on $\{\overline{X}_t > 0\}$. It follows easily from Anscombe's (1952) Theorem that Z_t is asymptotically standard normal under P_θ as $a \to \infty$ for every $0 < \theta < \infty$; but the simulations of Woodroofe and Keener (1987) indicate that this normality may not provide a good approximation to the distribution of Z_t for values of a of practical interest. See Table 1 below.

In this problem, it is easily seen that $\rho(\theta) = \sqrt{\theta}$, $0 < \theta < \infty$, so that $\rho(\theta)\kappa_1(\theta) = -1/2\sqrt{\theta}$, $0 < \theta < \infty$, and Corollary 5 suggests the approximation

$$F_{a,\theta}(z) \cong \Phi^{(1)}_{a,\theta}(z) \cong \Phi(z) + \frac{1}{2\sqrt{a\theta}}\varphi(z) = \Phi^*_{a,\theta}(z), \tag{22}$$

say, for $-\infty < z < \infty$ and $0 < a, \theta < \infty$.

In Table 1 below Monte Carlo estimates of $P_\theta\{Z_t \leq z\}$ are reported for $a = 12$, $\theta = 0.5$ and 1 and selected values of z, along with direct normal approximation $\Phi(z)$ and the right side of (22). The most striking aspect of the simulations is the amount by which normal approximation underestimates the probabilities. The approximation (22) is closer in all cases reported and much closer in some.

Table 1. Simulations of $P_\theta\{Z_t \leq z\}$ for $a = 12$

	$\theta = 0.5$			$\theta = 1.0$	
z	M.C.	$\Phi^*_{a,\theta}(z)$	$\Phi(z)$	M.C.	$\Phi^*_{a,\theta}(z)$
−1.80	.049	.052	.036	.046	.047
−1.20	.151	.145	.115	.136	.143
−0.60	.338	.342	.274	.317	.323
0.00	.576	.582	.500	.556	.558
0.60	.789	.795	.726	.766	.774
1.20	.921	.925	.885	.911	.913
1.80	.978	.980	.964	.975	.975

Note: $\Phi^*_{a,\theta}(z)$ is the right side of (22).
Source (of the simulations): Woodroofe and Keener (1987); based on 40,000 replications.

8. Higher Order Expansions

For higher order expansions the coefficient of $1/\sqrt{a}$ in Theorem 1 must be computed with more care.

Lemma 4. *Let $t = t_a$, $a \geq 1$, be stopping times for which* (17) *holds. Suppose that $\xi \in \Xi_0 \cap \Xi_1$ has compact support Ω_ξ, say; let Ω_0 denote a compact interval for which $\Omega_\xi \subseteq \Omega_0^o \subseteq \Omega_0 \subseteq \Omega$; and let A be the event $A = \{t \geq \eta a$ and $\hat{\theta}_t \in \Omega_0\}$. If $G : \Omega \times \Omega \to \mathbf{R}$ is continuously differentiable, then*

$$\lim_{a\to\infty} \int_A \sqrt{a}\left|E_\xi^t\{G(\Theta, \Theta) - G(\hat{\theta}_t, \Theta)\}\right| dP_\xi = 0.$$

Proof. It suffices to prove the lemma with $\sqrt{a}$ replaced by $\sqrt{t}$. Now, $\sqrt{n}\{G(\Theta, \Theta) - G(\hat{\theta}_n, \Theta)\}$ may be written as the sum of

$$\mathrm{I}_n = \frac{G_{10}(\hat{\theta}_n, \hat{\theta}_n)}{\sqrt{\psi''(\hat{\theta}_n)}} Z_n$$

and

$$\mathrm{II}_n = \sqrt{n}\{G(\Theta,\Theta) - G(\hat{\theta}_n,\Theta)\} - \frac{G_{10}(\hat{\theta}_n,\hat{\theta}_n)}{\sqrt{\psi''}(\hat{\theta}_n)} Z_n$$

for $n = 1,2,\ldots$. Using Proposition 4 and the continuity of G_{10} and ψ'' it is easily seen that there is a constant C for which

$$\begin{aligned}\int_A \left|E_\xi^t(\mathrm{I}_t)\right| dP_\xi &= \int_A \left|\frac{G_{10}(\hat{\theta}_t,\hat{\theta}_t)}{\sqrt{\psi''}(\hat{\theta}_t)}\right| \left|E_\xi^t(Z_t)\right| dP_\xi \\ &\le \int_A \frac{1}{\sqrt{t}}\left|\frac{G_{10}(\hat{\theta}_t,\hat{\theta}_t)}{\sqrt{\psi''}(\hat{\theta}_t)} K_1^\xi(\hat{\theta}_t,\Theta)\right| dP_\xi \le C\int_A \frac{1}{\sqrt{t}} dP_\xi,\end{aligned}$$

which approaches zero as $a \to \infty$. Using the Mean Value Theorem, it is easily seen that $\mathrm{II}_t \to 0$ in probability as $a \to \infty$ and that $|\mathrm{II}_t| \le C|Z_t|$ w.p.1 on A for all $a \ge 1$ for some constant C. So,

$$\int_A |E_\xi^t(\mathrm{II}_t)| dP_\xi \le \int_A |\mathrm{II}_t| dP_\xi \to 0$$

as $a \to \infty$. The lemma follows.

In the next theorem let $t = t_a$ denote stopping times for which (16) and (17) hold; and

$$\begin{aligned}R_{2,a}(\xi;h) = a\{E_\xi^t[h(Z_t) - \Phi h] &- \frac{1}{\sqrt{t}}(\Phi U h) E_\xi^t[\tilde{K}_1^\xi(\Theta)] \\ &- \frac{1}{a}(\Phi U^2 h) E_\xi^t[\rho^2(\Theta)\tilde{K}_2^\xi(\Theta)]\}\end{aligned}$$

for $h \in H$ and $\xi \in \Xi_0 \cap \Xi_2$, where $U^2 = U \circ U$.

Theorem 2. *Let $t = t_a$, $a \ge 1$, be stopping times for which* (16), (17) *and* (19) *hold for some $q > 1$. If $\xi \in \Xi_0 \cap \Xi_2$, then*

$$\lim_{a\to\infty} E_\xi\left\{\operatorname*{ess\,sup}_{h\in H_2} |R_{2,a}(\xi;h)|\right\} = 0.$$

Proof.

$$\lim_{a\to\infty} \int_{\{t\le \eta a\}} \operatorname*{ess\,sup}_{h\in H_2} \left|R_{2,a}(\xi;h)\right| dP_\xi = 0$$

as $a \to \infty$, may be established as in the proof of Theorem 1, since $q > 1$.

For the integral over $\{t > \eta a\}$, it is convenient to write (using Proposition 4)

$$R_{2,a}(\xi;h) = \mathrm{I}_a + \mathrm{II}_a + \mathrm{III}_a + \mathrm{IV}_a,$$

where

$$\begin{aligned}
\mathrm{I}_a &= \sqrt{a}\sqrt{(\tfrac{a}{t})}(\Phi U h)E_\xi^t[K_1^\xi(\hat{\theta}_t,\Theta) - \tilde{K}_1^\xi(\Theta)],\\
\mathrm{II}_a &= \frac{a}{t}E_\xi^t\{[K_2^\xi(\hat{\theta}_t,\Theta) - E_\xi^t[\tilde{K}_2^\xi(\Theta)]]U^2h(Z_t)\},\\
\mathrm{III}_a &= \frac{a}{t}E_\xi^t[\tilde{K}_2^\xi(\Theta)]E_\xi^t[U^2h(Z_t) - \Phi U^2 h],\\
\mathrm{IV}_a &= (\Phi U^2 h)E_\xi^t\Big\{\tilde{K}_2^\xi(\Theta)\Big[\frac{a}{t} - \rho^2(\Theta)\Big]\Big\}.
\end{aligned}$$

The analyses of $\mathrm{II}_a - \mathrm{IV}_a$ are similar to and simpler than those of $\mathrm{I}_a - \mathrm{III}_a$ in Theorem 1. For I_a, let Ω_0 and A be as in Lemma 4; let $1 < \alpha < \infty$ be such that $|K_1^\xi|^\alpha \in G_\xi$; and let β denote the conjugate value. Then, by Lemma 4,

$$\sqrt{a}\int_A \sqrt{(\tfrac{a}{t})}\Big|[E_\xi^t[K_1^\xi(\hat{\theta}_t,\Theta) - \tilde{K}_1^\xi(\Theta)]\Big|dP_\xi \to 0;$$

and

$$\begin{aligned}
&\sqrt{a}\int_{A'\cap\{t>\eta a\}} \sqrt{(\tfrac{a}{t})}\Big|[E_\xi^t[K_1^\xi(\hat{\theta}_t,\Theta) - \tilde{K}_1^\xi(\Theta)]\Big|dP_\xi\\
\le &2\sqrt{\frac{a}{\eta}}E_\xi\Big\{\sup_{n\ge\eta a}\Big|K_1^\xi(\hat{\theta}_n,\Theta)\Big|^\alpha\Big\}^{\frac{1}{\alpha}} \times P_\xi\{\hat{\theta}_n \notin \Omega_0, \exists\, n \ge \eta a\}^{\frac{1}{\beta}},
\end{aligned}$$

which approaches zero as $a \to \infty$ by Propositions 1 and 3.

For the Corollaries below, suppose that ρ' is absolutely continuous; recall the expression for $\tilde{K}_2^\xi(\theta)$; and let

$$\kappa_2 = \frac{1}{\psi''}\cdot\frac{(\rho^2)''}{\rho^2} - \frac{\psi'''}{\psi''^2}\cdot\frac{(\rho^2)'}{\rho^2} + \frac{15}{36}\cdot\frac{\psi'''^2}{\psi''^3} - \frac{1}{4}\cdot\frac{\psi^{(4)}}{\psi''^2}$$

and

$$\overline{\kappa}_2(\xi) = \int_\Omega \kappa_2(\theta)\rho^2(\theta)\xi(\theta)d\theta.$$

Corollary 7. *If, for every compact $\Omega_0 \subseteq \Omega$,*

$$\lim_{a\to\infty}\sqrt{a}\int_{\Omega_0}\Big|E_\theta\Big(\sqrt{\frac{a}{t}}\Big) - \rho(\theta)\Big|d\theta = 0, \tag{23}$$

then

$$\lim_{a\to\infty}\sup_{h\in H_2} a\left|\left[E_\xi(h(Z_t)) - \Phi h\right] - \frac{1}{\sqrt{a}}\Phi U h\cdot\overline{\kappa}_1(\xi) - \frac{1}{a}\Phi U^2 h\cdot\overline{\kappa}_2(\xi)\right| = 0 \quad (24)$$

for every $\xi \in \Xi_0 \cap \Xi_2$.

Proof. By the definition of $R_{2,a}$, the term on left side of (24) may be written as the absolute value of the sum of

$$\mathrm{I}_a = \sqrt{a}\Phi U h\left\{E_\xi\left\{\sqrt{\frac{a}{t}}E_\xi^t[\tilde{K}_1^\xi(\Theta)]\right\} - \overline{\kappa}_1(\xi)\right\},$$

$$\mathrm{II} = \Phi U^2 h\left\{E_\xi\{\rho^2(\Theta)\tilde{K}_2^\xi(\Theta)\} - \overline{\kappa}_2(\xi)\right\},$$

and

$$\mathrm{III}_a = E_\xi\{R_{2,a}(\xi; h)\}$$

for all $h \in H_2$, $a \geq 1$, and $\xi \in \Xi_0 \cap \Xi_2$. Here $\mathrm{III}_a \to 0$ uniformly in $h \in H_2$, as $a \to \infty$, by Theorem 2; and $\mathrm{II} = 0$ for all $h \in H_2$, by an integration by parts, as in the proof of Corollary 5. For the first term, there is a constant C for which

$$\begin{aligned}
|\mathrm{I}_a| &= \sqrt{a}|\Phi U h|\left|E_\xi\left[\sqrt{\frac{a}{t}}\tilde{K}_1^\xi(\Theta)\right] - E_\xi\left[\rho(\Theta)\tilde{K}_1^\xi(\Theta)\right]\right| \\
&\leq 2C\sqrt{a}\left|E_\xi\left\{\left[\sqrt{\frac{a}{t}} - \rho(\Theta)\right]\tilde{K}_1^\xi(\Theta)\right\}\right| \\
&\leq 2C\sqrt{a}\int_\Omega\left|E_\theta\left(\sqrt{\frac{a}{t}}\right) - \rho(\theta)\right||\tilde{K}_1^\xi(\theta)|\xi(\theta)d\theta,
\end{aligned}$$

which is independent of $h \in H_2$ and approaches zero as $a \to \infty$, by (23).

Corollary 7 may be restated in a form analogous to Corollary 6. Recall the definition of $F_{a,\theta}$ and define signed measures $\Phi_{a,\theta}^{(2)}$ by

$$\Phi_{a,\theta}^{(2)}h = \Phi h + \frac{1}{\sqrt{a}}\kappa_1(\theta)\Phi U h + \frac{1}{a}\kappa_2(\theta)\Phi U^2 h$$

for $h \in H$ and $\theta \in \Omega$. Then $\Phi_{a,\theta}^{(2)}$ provides a higher order, very weak approximation to $F_{a,\theta}$ in following sense.

Corollary 8. *If* (23) *holds for every compact* $\Omega_0 \subseteq \Omega$, *then*

$$\lim_{a\to\infty}\sup_{h\in H_2} a\left|\int_\Omega\left[F_{a,\theta}h - \Phi_{a,\theta}^{(2)}h\right]\xi(\theta)d\theta\right| = 0$$

for all $\xi \in \Xi_0 \cap \Xi_2$.

Acknowledgements

It is a pleasure to acknowledge conversations with Bob Keener and a careful report by an anonymous referee. The research for this paper was supported by the U.S. Army Research Office and the National Science Foundation.

References

Anscombe, F. J. (1952). Large sample theory of sequential estimation. *Proc. Cambridge. Philos. Soc.* **48**, 600-607.

Berger, J. O. and Walpole, R. L. (1984). *The Likelihood Principle.* IMS, Hayward, CA.

Bickel, P., Goetze, F. and Van Zwet, W. (1985). A simple analysis of third order efficiency of estimates. *Proc. Neyman-Kiefer Conference* **2**, 749-768.

Brown, L. D. (1986). *Fundamentals of Statistical Exponential Families.* IMS, Hayward, CA.

Chow, Y. S., Robbins, H. and Siegmund, D. (1971). *Great Expectations.* Houghton Mifflin.

Ghosh, J. K., Sinha, B. and Joshi, S. (1985). Expansions for posterior distributions and integrated Bayes risk. In *Proc. 3rd Purdue Symp. Statist. Dec. Th. Rel. Topics* **1**, 403-456.

Johnson, R. A. (1970). Asymptotic expansions associated with posterior distributions. *Ann. Math. Statist.* **41**, 851-864.

Stein, C. (1987). *Approximate Computation of Expectations.* IMS, Hayward, CA.

Woodroofe, M. (1986). Very weak expansions for sequential confidence intervals. *Ann. Statist.* **14**, 1049-1067.

Woodroofe, M. (1989). Very weak expansions for sequentially designed experiments: linear models. *Ann. Statist.* **17**, 1087-1102.

Woodroofe, M. and Keener, R. (1987). Asymptotic expansions in boundary crossing problems. *Ann. Probab.* **15**, 102-114.

Department of Statistics, University of Michigan, Ann Arbor, MI 48109-1027, U.S.A.

(Received February 1990; accepted July 1991)