The Annals of Statistics
1973, Vol. 1, No. 2, 231-240

ON THE ASYMPTOTIC SHAPE OF BAYESIAN SEQUENTIAL TESTS OF $\theta \leqq 0$ VERSUS $\theta > 0$ FOR EXPONENTIAL FAMILIES[1]

BY P. J. BICKEL

Princeton University and University of California, Berkeley

We show that, in a weak sense, as the cost of observation tends to zero, the shape of the continuation region of the Bayes solution for the exponential family problem given above is approximated by that of a corresponding problem for the Wiener process with drift. The approach is an extension of that used in Bickel and Yahav, *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* (1971).

1. Introduction. Consider the problem of testing $H: \theta \leqq 0$ versus $K: \theta > 0$ by sequential sampling in the following two models.

A. (i) We observe $X_1, X_2, \cdots$, which are i.i.d. (given θ) with density

$$p(x, \theta) = e^{\theta x - b(\theta)} \tag{0}$$

with respect to some σ finite measure μ.

(ii) The parameter θ ranges over Θ which includes 0 as an interior point. Moreover,

$$b'(0) = 0\,, \qquad b''(0) = 1\,.$$

(iii) The terminal loss is 0 or 1 according as the decision to accept or reject H is made correctly or not.

(iv) The cost of observation is c per unit time.

(v) We are given a generalized prior density ψ_c (with respect to Lebesgue measure) with which we measure our performance. Thus if we denote the risk of a procedure given θ by $B_\theta(\pi, c)$, our aim is to choose π so as to minimize $\int_\Theta B_\theta(\pi, c)\psi_c(\theta)\, d\theta$. We denote this minimum (Bayes) risk by $B(c, \psi_c)$.

B. (i) We observe $X(t) = \theta t + W(t)$, $t \geqq 0$ where W is a standard Wiener process and θ is unknown.

(ii) The parameter θ ranges over R.

(iii) The terminal loss is as in problem A.

(iv) The cost of observation is c per unit time.

(v) We are given ψ_c as above and our aim is the same. The risk of a procedure π is denoted by $R_\theta(\pi, c)$ (given θ) and the Bayes risk by $R(c, \psi_c)$.

In [2], to which we refer the reader for a precise definition of the notion of procedure in the two problems, the following theorem was proved (Theorem 4.2).

Received September 16, 1971.

[1] Research performed while the author was a fellow of the John Simon Guggenheim Foundation on sabbatical leave at Princeton University, and with the partial support of the O.N.R. under contract N00014-67-A0114-0004 with the University of California at Berkeley.

THEOREM. *Let ψ be a bounded probability density which is continuous at* 0. *Then,*

$$\frac{B(c, \psi)}{c^{\frac{1}{2}}} = \frac{R(c, \psi)}{c^{\frac{1}{2}}} + o(1) = \psi(0)R(1, 1) + o(1) \tag{1}$$

where the second 1 *in* $R(1, 1)$ *denotes the density of Lebesgue measure,* $\psi \equiv 1$.

Moreover, there exists a sequence of policies $\{\pi_c\}$ *independent of* ψ *and the exponential family such that*

$$c^{-\frac{1}{2}} \int_{-\infty}^{\infty} B_\theta(\pi_c, c)\psi(\theta)\, d\theta \to \psi(0)R(1, 1)\,. \tag{2}$$

This type of result, convergence of the Bayes risk, was given by Chernoff [3] and Kiefer and Sacks [6] for testing problems as above, when hypothesis and alternative are separated at least by an indifference region.

It is well known that in these two problems the optimal stopping times are describable in terms of a *continuation region* (or the complementary stopping region) Q, a subset of points of the right half plane such that one continues sampling if and only if (n, S_n) (respectively $(t, X(t))$) is in Q, where $S_m = \sum_{j=1}^m X_j$, $m \geqq 0$. Of course the continuation region depends on which problem we are considering, c, and ψ. In 1962 Schwarz [8] showed, in the analogue of problem A, when hypothesis and alternative were separated at least by an indifference region, that the continuation regions suitably rescaled in space and time converged to a limiting shape. Our main result, Theorem 2, establishes a weak version of Schwarz' result for problem A. Of course, both rescaling and limiting shape are different. The latter, as the result of [2] suggests, is the continuation region for problem B with $c = 1$ and $\psi \equiv 1$.

2. The results. We recall a well-known characterization of the continuation regions in these problems which we shall from now on denote by $Q_A(c, \psi)$ and $Q_B(c, \psi)$ (with dependence on (c, ψ) usually suppressed). Denote the posterior density of θ given $S_n = x/c^{\frac{1}{2}}$, $n = [t/c]$ by

$$\psi_c(\theta \mid x, t) = \left[\exp\left\{\frac{\theta x}{c^{\frac{1}{2}}} - \left[\frac{t}{c}\right]b(\theta)\right\}\right]\psi(\theta)A(x, t, c) \tag{3}$$

with $A(x, t, c)$ defined by $\int_{-\infty}^{\infty} \psi_c(\theta \mid x, t)\, d\theta = 1$. Note that (3) is well-defined for all $t \geqq 0$ real. Let,

$$B(x, t, c) = B(c, \psi_c(\cdot \mid x, t))\,, \tag{4}$$

the expected future risk given that sampling has proceeded to time $n = [t/c]$ and $S_n = x/c^{\frac{1}{2}}$. Let,

$$Y(x, t, c) = \min\{\textstyle\int_0^\infty \psi_c(\theta \mid x, t)\, d\theta,\ 1 - \int_0^\infty \psi_c(\theta \mid x, t)\, d\theta\} \tag{5}$$

the stopping risk at time $n = [t/c]$ given that $S_n = x/c^{\frac{1}{2}}$. Then,

$$Q_A = \{(s, n) : Y(sc^{\frac{1}{2}}, nc, c) > B(sc^{\frac{1}{2}}, nc, c)\} \tag{6}$$

where n ranges over the natural numbers. *We shall however suppose throughout* *the sequel that* Q_A *is given by* (6) *where n ranges over* R^+. Similarly let

$$R(x, t, c) = R(c, \phi_c) \tag{7}$$

where ϕ_c is the density of the normal $[xc^{\frac{1}{2}}/t, c/t]$ distribution, the posterior distribution of θ given $X(\tau) = x/c^{\frac{1}{2}}$, $\tau = t/c$ when $\psi \equiv 1$, and

$$\tilde{Y}(x, t, c) = \min\{\textstyle\int_0^\infty \phi_c(\theta \mid x, t)\, d\theta, 1 - \int_0^\infty \phi_c(\theta \mid x, t)\, d\theta\}\,. \tag{8}$$

Then,

$$Q_B(c, 1) = \{(s, \tau) : \tilde{Y}(sc^{\frac{1}{2}}, \tau c, c) > R(sc^{\frac{1}{2}}, \tau c, c)\} \tag{9}$$

with equality holding on the complement of Q_B.

Our scaling of time and state space is, of course, such that,

$$\begin{aligned} \tilde{Y}(x, t, c) &\equiv \tilde{Y}(x, t, 1) = \Phi\left[\frac{-|x|}{t^{\frac{1}{2}}}\right], \\ R(x, t, c) &\equiv R(x, t, 1)\,, \end{aligned} \tag{10}$$

(cf. [5]). Define the homothety T_c of $R \times R^+$ onto itself by, $T_c(s, \tau) = (sc^{\frac{1}{2}}, \tau c)$. Then, (10) implies that

$$T_c(Q_B(c, 1)) = Q_B(1, 1)\,. \tag{11}$$

To apply this characterization we need the following simple extension of the theorem stated in the introduction.

THEOREM 1. *Let* $\psi_c(\cdot)$ *be a sequence of nonnegative measurable functions such that*,

(a) $\sup_{c,\theta} c^{\frac{1}{2}}\psi_c(\theta) < \infty$
(b) $c^{\frac{1}{2}}\psi_c(\theta c^{\frac{1}{2}}) \to \gamma(\theta) \neq 0$ *as* $c \to 0$
(c) $\int_{-\infty}^{\infty} \psi_c(\theta)\, d\theta = o(\exp(c^{-\gamma}))$ *for all* $\gamma > 0$.

Then,

$$B(c, \psi_c) = R(1, \gamma) + o(1) \qquad \text{as } c \to 0\,. \tag{12}$$

That Theorem 1 is indeed an extension of Theorem 4.2 of [2] may be seen by taking $\psi_c(\theta) = \psi(\theta)/c^{\frac{1}{2}}$. The proof of this result follows exactly the lines of the proof of Theorem 4.2 and will not be given. A ready consequence is,

COROLLARY. *Let* ψ *be bounded and continuous at* 0. *Then*,

$$B(x, t, c) = R(x, t, 1) + o(1) \qquad \text{as } c \to 0 \tag{13}$$

uniformly for $|x| \leqq M < \infty$, $\varepsilon \leqq t \leqq \varepsilon^{-1}$, $\varepsilon > 0$.

PROOF. It suffice to show that if $x_c \to x$, $t_c \to t > 0$ $\psi_c(\cdot \mid x_c, t_c)$ satisfy (a), (b), and (c). Since b is convex $\theta x_c/c^{\frac{1}{2}} - [t_c/c]b(\theta)$ is maximized by the solution θ_c of

$$b'(\theta) = \frac{x_c}{c^{\frac{1}{2}}}\left[\frac{t_c}{c}\right]^{-1} \sim \frac{xc^{\frac{1}{2}}}{t}\,. \tag{14}$$

Since $b'(0) = 0$, $b''(0) > 0$ we obtain,

$$\theta_c \sim \frac{xc^{\frac{1}{2}}}{t}, \tag{15}$$

and hence $\exp\{(\theta/c^{\frac{1}{2}})x_c - [t_c/c]b(\theta)\}$ is uniformly bounded.

By Taylor expansion we also obtain,

$$\theta x_c - \left[\frac{t_c}{c}\right] b(\theta c^{\frac{1}{2}}) = \left(\theta x - \frac{t\theta^2}{2}\right) + o(1) \tag{16}$$

uniformly on compact sets of θ.

Moreover, there exists a $\delta > 0$ such that

$$\theta x_c - \left[\frac{t_c}{c}\right] b(\theta c^{\frac{1}{2}}) \leqq \theta x_c - \frac{\mu t \theta^2}{2} \tag{17}$$

for some $\mu > 0$, and all $|\theta| \leqq \delta c^{-\frac{1}{2}}$. Thus by dominated convergence,

$$\lim_c \int_{-\delta c^{-\frac{1}{2}}}^{\delta c^{-\frac{1}{2}}} \left[\exp\left\{\theta x_c - \left[\frac{t_c}{c}\right] b(\theta c^{\frac{1}{2}})\right\}\right] \psi(\theta c^{\frac{1}{2}})\, d\theta = \left(\frac{2\pi}{t}\right)^{\frac{1}{2}} \left[\exp \frac{x^2}{2t}\right] \psi(0) \tag{18}$$

for some $\delta > 0$. On the other hand by the convexity of b and (15) we can bound for c sufficiently small,

$$\begin{aligned} &\int_{\delta c^{-\frac{1}{2}}}^{\infty} \left[\exp\left\{\theta x_c - \left[\frac{t_c}{c}\right] b(\theta c^{\frac{1}{2}})\right\}\right] \psi(\theta c^{\frac{1}{2}})\, d\theta \\ &\qquad \leqq c^{-\frac{1}{2}} \exp\left\{\frac{\delta x_c}{c^{\frac{1}{2}}} - \left[\frac{t_c}{c}\right] b(\delta)\right\} \\ &\qquad \leqq c^{-\frac{1}{2}} \exp\left\{\frac{\delta x}{c^{\frac{1}{2}}} - \frac{t\mu\delta^2}{c}\right\} = o(1) \end{aligned} \tag{19}$$

for some $\mu > 0$. A similar bound applies to $\int_{-\infty}^{-\delta c^{-\frac{1}{2}}}$ for the same expression. In view of (18) and (19),

$$c^{\frac{1}{2}} A(x_c, t_c, c) \to \left(\frac{2\pi}{t}\right)^{\frac{1}{2}} \left(\exp \frac{x^2}{2t}\right) \psi(0). \tag{20}$$

Therefore (a) holds as does (b) with $\gamma(\theta) = \phi_1(\theta \mid x, t)$ and the corollary follows.

REMARK. The same argument and obvious estimates (cf. [1] Section 3) show that if $K(c) \uparrow \infty$, $K(c) = o(c^{-\frac{1}{2}})$ then,

$$\int_{-K(c)}^{K(c)} |\theta|^r |c^{\frac{1}{2}} \psi(c^{\frac{1}{2}}\theta) - \phi_1(\theta \mid x, t)|\, d\theta \to 0 \tag{21}$$

uniformly for $|x| \leqq M < \infty$, $\varepsilon \leqq t \leqq \varepsilon^{-1}$.

A direct argument or a proof parallel to that of Theorem 1 shows that,

$$\tilde{Y}(x, t, 1) = \tilde{Y}(x, t) + o(1) \tag{22}$$

uniformly in $|x| \leqq M < \infty$, $\varepsilon \leqq t \leqq \varepsilon^{-1}$, $\varepsilon > 0$.

The result we are aiming at is, of course, $T_c(Q_A) \sim Q_B(1, 1)$. To make this precise we need another notion.

DEFINITION. If S is a subset of a metric space and δ is > 0 let $S^\delta = \{x \in S: B(x, \delta) \subset S\}$ where $B(x, \delta)$ is the open δ-ball about x. It is easy to see that S^δ is closed and $\bigcup_{\delta>0} S^\delta = S^0$—the interior of S. We can now state,

THEOREM 2. *For every* $\varepsilon > 0$, $\delta > 0$, $\eta > 0$ *there exists* $c_0 > 0$ *sufficiently small such that if* $c \leqq c_0$,

$$[\{[Q_B(1, 1)]'\}^\eta]' \supset [T_c(Q_A)]^\delta \cap \Lambda_\varepsilon \tag{23}$$

$$T_c(Q_A) \supset [Q_B(1, 1)]^\delta \cap \Lambda_\varepsilon \tag{24}$$

where $'$ *denotes complementation and* $\Lambda_\varepsilon = \{(x, t): \varepsilon \leqq t \leqq \varepsilon^{-1}\}$.

Thus a region slightly larger than the continuation region for B with $\psi = c = 1$ contains a region slightly smaller than the continuation region for A and conversely. If we knew that the $T_c(Q_A)$ had no enduring spikes (23) and (24) would imply Hausdorff convergence of $T_c(Q_A) \cap \Lambda_\varepsilon$ to $Q_B(1, 1) \cap \Lambda_\varepsilon$ and hence convergence of the boundary of $T_c(Q_A)$ to that of $Q_B(1, 1)$. We have, however, not been able to show this.

PROOF OF THEOREM 2. By a weak compactness argument it is easy to show that $R(x, t, 1)$ is continuous. Thus

$$\inf\{\tilde{Y}(x, t, 1) - R(x, t, 1): (x, t) \in [Q_B(1, 1)]^\delta \cap \Lambda_\varepsilon\} > 0 .$$

Furthermore it is well known that for each $t > 0$ $\{x: \tilde{Y}(x, t, 1) > R(x, t, 1)\}$ is a finite symmetric interval. Assertion (24) is then an immediate consequence of (13) and (22). Now suppose $(x, t) \in [T_c(Q_A)]^\delta$. By taking the suitable metric we can suppose $(y, t + \delta) \in T_c(Q_A)$ for $x - \delta < y < x + \delta$. Let $\tau = c$[first $m \geqq t/c$ such that $[(c^{\frac{1}{2}}S_m - x, cm) \in [T_c(Q_A)]'] - t$. Then

$$B(x, t, c) = E\left(\left[Y(c^{\frac{1}{2}}S_{(\tau+t)/c}, \tau + t, c) + c\left[\frac{\tau}{c}\right]\right] \middle| n = \left[\frac{t}{c}\right], S_n = \frac{x}{c^{\frac{1}{2}}}\right) \tag{25}$$

Let $\varepsilon(c)$ be any sequence tending $\downarrow 0$ as $c \downarrow 0$ such that $c^{\frac{1}{2}} = o(\varepsilon(c))$.
For convenience, define

$$W(q, c) = c^{\frac{1}{2}}S_{[q/c]} \qquad q \geqq 0 . \tag{26}$$

Let P^*, (E^*) denote conditional probability given $n = [t/c]$, $S_n = x/c^{\frac{1}{2}}$. Then,

$$\begin{aligned} Y(x, t, c) > B(x, t, c) \\ \geqq \varepsilon(c)P^*[\tau \geqq \varepsilon(c)] + E^*(B(W(\tau + t, c), \tau + t, c)I_{[\tau \geqq \varepsilon(c)]}) . \end{aligned} \tag{27}$$

Then,

$$\begin{aligned} E^*(B(W(\tau + t, c), \tau + t, c)I_{[\tau \geqq \varepsilon(c)]}) \\ = E^*(B(W(t + \varepsilon(c), c), t + \varepsilon(c), c)I_{[\tau \geqq \varepsilon(c)]}) \end{aligned} \tag{28}$$

(cf. [9], for example). Now in view of the definition of l (27) implies that,

$$\begin{aligned} Y(x, t, c) > \varepsilon(c)P^*[\tau \geqq \varepsilon(c)] + E^*(B(W(t + \varepsilon(c), c), t + \varepsilon(c), c)) \\ - P^*[\tau < \varepsilon(c)] . \end{aligned} \tag{29}$$

For convenience, in what follows, write ψ_c for $\psi_c(\cdot \mid x, t)$. Since $(x, t) \in [T_c(Q_A)]^\delta$,

$$
\begin{aligned}
P^*[\tau < \varepsilon(c)] &\leqq P^*\left[\max\left\{|S_m - S_n| : n \leqq m \leqq n + \frac{\varepsilon(c)}{c}\right\} > \frac{\delta}{c^{\frac{1}{2}}}\right] \\
&= \int_{-\infty}^{\infty} P_\theta\left[\max\left\{|S_j| : 1 \leqq j \leqq \frac{\varepsilon(c)}{c}\right\} > \frac{\delta}{c^{\frac{1}{2}}}\right]\psi_c(\theta)\, d\theta \\
&\leqq \left[\int_{-M_c}^{M_c} c^{\frac{1}{2}}\psi_c(\theta c^{\frac{1}{2}})\, d\theta\right] \max\left\{P_\eta\left[\max\left\{|S_j| : 1 \leqq j \leqq \frac{\varepsilon(c)}{c}\right\}\right.\right. \\
&\qquad \left.\left. > \frac{\delta}{c^{\frac{1}{2}}}\right] : |\eta| \leqq M_c c^{\frac{1}{2}}\right\} + \int_{|\theta| > M_c} c^{\frac{1}{2}}\psi_c(\theta c^{\frac{1}{2}})\, d\theta\,.
\end{aligned}
\tag{30}
$$

Now, (if we denote the positive part of a number a by a^+),

$$
\begin{aligned}
&P_\eta\left[\max\left\{|S_j| : 1 \leqq j \leqq \frac{\varepsilon(c)}{c}\right\} > \frac{\delta}{c^{\frac{1}{2}}}\right] \\
&\leqq P_\eta\left[\max\left\{|S_j - jb'(\eta)| : 1 \leqq j \leqq \frac{\varepsilon(c)}{c}\right\} > c^{-\frac{1}{2}}\left(\delta - \frac{\varepsilon(c)}{c^{\frac{1}{2}}} b'(\eta)\right)\right] \\
&\leqq 3\varepsilon^2(c) E_\eta(X_1 - b'(\eta))^4\left[\left(\delta - \frac{\varepsilon(c)}{c^{\frac{1}{2}}} b'(\eta)\right)^+\right]^{-4},
\end{aligned}
\tag{31}
$$

by Kolmogorov's inequality. Since $b'(\eta) \sim \eta$ in a neighborhood of 0, if $M_c = o(c^{-\frac{1}{2}})$, the right-hand side of (31) is $o(\varepsilon^2(c))$ uniformly for $|\eta| \leqq M_c c^{\frac{1}{2}}$. Arguing as for (19) it also follows that under our hypothesis on M_c.

$$\int_{|\theta| > M_c} c^{\frac{1}{2}}\psi_c(\theta c^{\frac{1}{2}})\, d\theta = o(M_c^{-1}) \tag{32}$$

as $M_c \to \infty$ uniformly for $|x|$ bounded, t bounded away from 0 and ∞. We conclude from (29), (31), and (32) that $P^*[\tau < \varepsilon(c)] = o(\varepsilon(c))$. Hence

$$\liminf_c \frac{1}{\varepsilon(c)}\,[Y(x, t, c) - E^*(B(W(t + \varepsilon(c), c), t + \varepsilon(c), c))] \geqq 1 \tag{33}$$

uniformly on $[T_c(Q_A)]^\delta \cap \Lambda_\varepsilon \cap \Gamma_M$ where

$$\Gamma_M = \{(x, t) : |x| \leqq M\}\,. \tag{34}$$

To complete the argument we need a result complementary to (33) which we isolate as a lemma

LEMMA 1. *Uniformly on* $[Q_B{}']^\gamma \cap \Lambda_\varepsilon \cap \Gamma_M$,

$$[\tilde{Y}(x, t, 1) - E^*(R(W(t + \varepsilon(c), c), t + \varepsilon(c), c))] = o(\varepsilon(c))\,, \tag{35}$$

for any sequences $\varepsilon(c)$ *such that* $\varepsilon(c) = (\lambda(c))^{-1}$ *satisfying* (50), *and* $c^{\frac{1}{2}} = o(\varepsilon(c))$. *It is clear from the definition of* $\lambda(c)$ *that such sequences exist.*

Let us first note how (33) and (35) establish (23). For fixed M, η choose $\varepsilon(c)$ (by Corollary 1 and (22)) so that

$$B(x, t, c) = R(x, t, 1) + o(\varepsilon(c)) \tag{36}$$

$$Y(x, t, c) = \tilde{Y}(x, t, 1) + o(\varepsilon(c)) \tag{37}$$

uniformly on $\Lambda_\varepsilon \cap \Gamma_M$ and such that (35) is satisfied. It follows from (36) that,

$$(38)\qquad E^*[B(W(t+\varepsilon(c), c), t+\varepsilon(c), c) - R(W(t+\varepsilon(c), c), t+\varepsilon(c), 1)] \\ = o(\varepsilon(c))$$

on $\Lambda_\varepsilon \cap \Gamma_M$. This is clear since,

$$(39)\qquad P^*[|W(t+\varepsilon(c), c) - x| \geqq \gamma] = \int_{-\infty}^{\infty} P_\theta\left[|S_{(\varepsilon(c))/c}| \geqq \frac{\gamma}{c^{\frac{1}{2}}}\right]\psi_c(\theta)\,d\theta \\ = o(\varepsilon(c))$$

for any $\gamma > 0$, by arguing as in (30) and (32). Now, (33), (35), (37), and (38) imply that,

$$(40)\qquad [T_c(Q_A)]^\delta \cap \{[Q_B(1,1)]'\}^\gamma \cap \Lambda_\varepsilon \cap \Gamma_M = \phi$$

for c sufficiently small. Suppose that $M > \sup\{|x| : (x, t) \in [\{[Q_B(1,1)]'\}^\gamma]'$. It is then clear that the restriction $|x| \leqq M$ may be dropped in (40) since the t sections of $[T_c(Q_A)]^\delta \cap \Lambda_\varepsilon$ are intervals (cf. [10]). Thus, (22) and the theorem will follow once we have established Lemma 1.

We prove (35) in two steps

$$(41)\qquad E^*(R(W(t+\varepsilon(c), c), t+\varepsilon(c), 1) - \tilde{Y}(W(t+\varepsilon(c), c), t+\varepsilon(c), 1)) \\ = o(\varepsilon(c))$$

and

$$(42)\qquad \tilde{Y}(x, t, 1) - E^*(\tilde{Y}(W(t+\varepsilon(c), c), t+\varepsilon(c), 1) = o(\varepsilon(c))\,.$$

Since $\tilde{Y}(\cdot, \cdot, 1) = R(\cdot, \cdot, 1)$ on $[Q_B']^\gamma$ (41) follows from (38) and (39). To prove (42) express the left-hand side of (42) as,

$$(43)\qquad \int_{-\infty}^{\infty}\left[\Phi\left(\frac{-|x|}{t^{\frac{1}{2}}}\right) - E_\theta\left(\Phi\left(\frac{-|x+W(\varepsilon(c), c)|}{(t+\varepsilon(c))^{\frac{1}{2}}}\right)\right)\right]\psi_c(\theta \mid x, t)\,d\theta\,.$$

On $[Q_B']^\gamma \cap \Lambda_\varepsilon$, $|x|$ is bounded away from 0. Suppose without loss of generality that x is > 0. Then,

$$(44)\qquad P_\theta[W(\varepsilon(c), c) < -x] \leqq P_\theta\left[\left|S_{\varepsilon(c)/c} - \frac{\varepsilon(c)}{c}b'(\theta)\right| \geqq \frac{|x|}{c^{\frac{1}{2}}} - \frac{\varepsilon(c)}{c}b'(\theta)\right] \\ = o(\varepsilon(c))$$

uniformly for $|\theta| \leqq c^{\frac{1}{2}}o(\varepsilon(c))^{-1}$. Moreover,

$$(45)\qquad \int_{\{|\theta| > c^{\frac{1}{2}}o(\varepsilon(c))^{-1}\}} \psi_c(\theta \mid x, t)\,d\theta = \int_{\{|\eta| > o(\varepsilon(c))^{-1}\}} c^{\frac{1}{2}}\psi_c(c^{\frac{1}{2}}\eta \mid x, t)\,d\eta = o(\varepsilon(c))$$

by estimates apparent from the proof of Corollary 1. Thus if x is positive,

$$(46)\qquad \tilde{Y}(x, t, 1) - E^*(\tilde{Y}(W(t+\varepsilon(c), c), t+\varepsilon(c), 1)) \\ = \int_{-\infty}^{\infty} E_\theta\left[\Phi\left(\frac{-x}{t^{\frac{1}{2}}}\right) - \Phi\left(\frac{-(x+W(\varepsilon(c), c))}{(t+\varepsilon(c))^{\frac{1}{2}}}\right)\right]\psi_c(\theta \mid x, t)\,d\theta \\ + o(\varepsilon(c))\,.$$

Now,

$$E_\theta\left[\Phi\left(\frac{-x}{t^{\frac{1}{2}}}\right) - \Phi\left(\frac{-(x + W(\varepsilon(c), c))}{(t + \varepsilon(c))^{\frac{1}{2}}}\right)\right]$$

$$(47)\qquad = \frac{x}{t^{\frac{1}{2}}}\phi\left(\frac{x}{t^{\frac{1}{2}}}\right)\left(b'(\theta)\frac{\varepsilon(c)}{c^{\frac{1}{2}}} + x((t + \varepsilon(c))^{-\frac{1}{2}} - t^{\frac{1}{2}})\right) - \phi\left(\frac{x}{t^{\frac{1}{2}}}\right)\left(\frac{t + x^2}{2t}\right)$$

$$\times\left[\varepsilon(c)b''(\theta) + \left(b'(\theta)\frac{\varepsilon(c)}{c^{\frac{1}{2}}} + x((t + \varepsilon(c))^{-\frac{1}{2}} - t^{-\frac{1}{2}})\right)^2\right]$$

$$+ E_\theta\left(M(\theta, x, t)\left(\frac{W(\varepsilon(c), c) + x}{(t + \varepsilon(c))^{\frac{1}{2}}} - \frac{x}{t^{\frac{1}{2}}}\right)^3\right)$$

where M though random is uniformly bounded in x, t on Λ_ε, and θ in a neighborhood of 0. Now on a neighborhood of 0 by a standard inequality (see Chung [5] for example).

$$(48)\qquad E_\theta\left|W(\varepsilon(c), c) - \frac{\varepsilon(c)b'(\theta)}{c^{\frac{1}{2}}}\right|^3 \leqq K\varepsilon^{\frac{3}{2}}(c)$$

where K is a constant (used generically). Let $\lambda(c) \uparrow \infty$, $\lambda(c) = o(c^{-\frac{1}{2}})$. Then, by (47) and (48) and the usual Taylor expansion,

$$\int_{-c^{\frac{1}{2}}\lambda(c)}^{c^{\frac{1}{2}}\lambda(c)} E_\theta\left[\Phi\left(\frac{-x}{t^{\frac{1}{2}}}\right) - \Phi\left(\frac{-(x + W(\varepsilon(c), c))}{(t + \varepsilon(c))^{\frac{1}{2}}}\right)\right]\psi_c(\theta \mid x, t)\, d\theta$$

$$(49)\qquad = \phi\left(\frac{x}{t^{\frac{1}{2}}}\right)\left[\int_{-c^{\frac{1}{2}}\lambda(c)}^{c^{\frac{1}{2}}\lambda(c)}\left(\frac{x}{t^{\frac{1}{2}}}\right)\left(\theta\frac{\varepsilon(c)}{c^{\frac{1}{2}}} + (t/t + \varepsilon(c))^{\frac{1}{2}} - 1\right) - \frac{1}{2}\left(\frac{x^2}{t} + 1\right)\right]\varepsilon(c)$$

$$+ K\left[|\theta|\left[\varepsilon(c) + \frac{\varepsilon^2(c)}{c^{\frac{1}{2}}}\right] + \theta^2\left[\frac{\varepsilon(c)}{c^{\frac{1}{2}}} + \frac{\varepsilon^3(c)}{c}\right] + |\theta|^3\varepsilon^3(c)c^{-\frac{3}{2}}\right]$$

$$\times\ \psi_c(\theta \mid x, t)\, d\theta + O(\varepsilon^{\frac{3}{2}}(c))\,.$$

Now by (21) for suitable $\lambda(c) \uparrow \infty$ we have,

$$(50)\qquad \int_{-\lambda(c)c^{\frac{1}{2}}}^{\lambda(c)c^{\frac{1}{2}}} (1 + |\theta|^3)|c^{\frac{1}{2}}\psi_c(\theta c^{\frac{1}{2}} \mid x, t) - \phi_1(\theta \mid x, t)|\, d\theta = o\left(\frac{1}{\lambda(c)}\right)$$

uniformly on $\Lambda_\varepsilon \cap \Gamma_M$. If $\varepsilon(c) = (\lambda(c))^{-1}$ it follows from (49) and (50) that uniformly as above,

$$(51)\qquad \int_{-c^{\frac{1}{2}}\lambda(c)}^{\lambda(c)c^{\frac{1}{2}}}\left[E_\theta\left(\Phi\left(\frac{-x}{t^{\frac{1}{2}}}\right) - \Phi\left(\frac{-x + W(\varepsilon(c), c)}{(t + \varepsilon(c))^{\frac{1}{2}}}\right)\right)\right]\psi_c(\theta \mid x, t)\, d\theta = o(\varepsilon(c))\,,$$

and hence by (45)

$$(52)\qquad \tilde{Y}(x, t, 1) - E^*(\tilde{Y}(W(t + \varepsilon(c), c), t + \varepsilon(c), 1)) = o(\varepsilon(c))\,.$$

The lemma is proved.

3. Extensions and comments. Consider problems A and B as before with the same cost structure and prior but a more general loss function which is 0 if the decision is correct but otherwise is a bounded function l of θ such that

$$(53)\qquad l(\theta) = |\theta|^\alpha + o(|\theta|^\alpha)$$

in a neighborhood of 0 for $\alpha \geqq 0$. Using the methods of [2] it is straightforward to prove the following generalization of Theorem 1. (Retain the notation of the preceding section noting that R and B depend on α.)

THEOREM 1′. *Let ψ_c be a sequence of nonnegative measurable functions such that,*

(a) $\sup_{c,\theta} c^{(1/(\alpha+2))}\psi_c(\theta) < \infty$
(b) $c^{(1/(\alpha+2))}\psi_c(\theta c^{(1/(\alpha+2))}) \to \gamma(\theta) \neq 0$ *as* $c \to 0$
(c) $\int_{-\infty}^{\infty} \psi_c(\theta) = o(\exp(c^{-\gamma}))$ *for all* $\gamma > 0$.

Then,

$$B(c, \psi_c) = \psi(0)c^{(\alpha/(\alpha+2))}R(1, \gamma) + o(c^{(\alpha/(\alpha+2))}) . \tag{54}$$

A sequence of procedures independent of ψ and the particular exponential family which is asymptotically Bayes may also be found.

If we now redefine T_c more generally as the mapping of $R \times R^+$ onto itself given by,

$$T_c(s, \tau) = (sc^{(1/(\alpha+2))}, \tau c^{(2/(\alpha+2))}) . \tag{55}$$

Theorem 2 goes over verbatim.

The mapping (55), of course, indicates what changes are needed to make the proof of Theorem 1′ and the generalization of Theorem 2 go through.

We must consider 0 neighborhoods of orders slightly larger than $c^{(1/(\alpha+2))}$, truncate at $Tc^{(-2/(\alpha+2))}$ and then let $T \to \infty$. In the proof of Theorem 4.1 of [2] we must in general take more than 2 stages (e.g., 3 for $\alpha > 4$). Assertions (33) and (35) may also be readily checked.

The principal weakness of our Schwarz-type results in our opinion is that although they indicate that the Wiener process approximation is valid for that portion of the Bayes procedure which applies to the sample sequences contributing substantially to the risk it gives no idea of what happens for that portion of the sample space which is really likely to occur. For example, if $\alpha = 0$ clearly P[Stopping on or after time ε/c] $= O(c^{\frac{1}{2}})$ whatever be $\varepsilon > 0$. More refined studies of the shape for t near 0 would certainly seem desirable. Some numerical studies due to Lindley and Barnett [7] of the approximation of the Wiener process solution given by Chernoff for $\alpha = 1$ to the binomial-beta problem lead one to believe that the fit should be reasonably good.

REFERENCES

[1] BICKEL, P. J. and YAHAV, J. A. (1967). A.P.O. procedures in sequential analysis. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 401–413.
[2] BICKEL, P. J. and YAHAV, J. A. (1972). On the Wiener process approximation to Bayesian sequential testing problems. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 57–84.
[3] CHERNOFF, H. (1959). Sequential design of experiments. *Ann. Math. Statist.* **30** 755–770.
[4] CHERNOFF, H. (1961). Sequential tests for the mean of a normal distribution. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 79–91.
[5] CHUNG, K. L. (1951). The strong law of large numbers. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 341–353.

[6] Kiefer, J. and Sacks, J. (1963). Asymptotically optimum sequential inference and design. *Ann. Math. Statist.* **34** 705–750.

[7] Lindley, D. V. and Barnett, B. N. (1965). Sequential sampling: two decision problems with linear losses for binomial and normal random variables. *Biometrika* **52** 507–532.

[8] Schwarz, G. (1962). Asymptotic shapes of Bayesian sequential testing regions. *Ann. Math. Statist.* **33** 224–236.

[9] Siegmund, D. (1967). Some problems in the theory of optimal stopping. *Ann. Math. Statist.* 1627–1640.

[10] Sobel, M. (1953). An essentially complete class of decision functions for certain standard sequential problems. *Ann. Math. Statist.* **24** 319–337.

Department of Statistics
University of California
Berkeley, California 94720