ON THE WIENER PROCESS APPROXIMATION TO BAYESIAN SEQUENTIAL TESTING PROBLEMS

P. J. BICKEL[1]
UNIVERSITY OF CALIFORNIA, BERKELEY
and
J. A. YAHAV[2,3]
UNIVERSITY OF TEL AVIV

1. Introduction and summary

In 1959 Chernoff [7] initiated the study of the asymptotic theory of sequential Bayes tests as the cost of observation tends to zero. He dealt with the case of a finite parameter space. The definitive generalization of the line of attack initiated in that paper was given by Kiefer and Sacks in [13]. Their work as well as that of Chernoff, the intervening papers of Albert [1], Bessler [3], and Schwarz [19], and the subsequent work of the authors [4] used implicitly or explicitly the theory of large deviations and applied only to situations where hypothesis and alternative were separated or at least an indifference region was present.

In the meantime in 1961 Chernoff [8] began to study the problem of testing $H: \theta \leqq 0$ versus $K: \theta > 0$ on the basis of observation of a Wiener process with drift θ per unit time as an approximation to the discrete time normal observations problem. Having made the striking observation that study of the asymptotic behavior of the Bayes procedures for any normal prior was in this case equivalent to the study of the Bayes procedure with Lebesgue measure as prior and unit cost of observation, he reduced this problem for suitable loss functions to the solution of a free boundary problem for the heat equation. In subsequent work ([2], [9], [10] and [16]) the nature of this solution was investigated by Chernoff and others.

In this paper we are concerned with the problem of testing $H: \theta \leqq 0$ versus $K: \theta > 0$ by sampling sequentially from a member of one parameter exponential (Koopman-Darmois) family of distributions (see equation (3.1)) at cost c per observation. We will assume the simple zero-one loss structure in which an error in decision costs one unit while being right costs nothing.

[1]Prepared with the partial support of the Office of Naval Research, Contract NONR N00014-69-A-0200-1038.
[2]Prepared with the partial support of U.S. Public Health Grant GM-10525(07).
[3]This research was done while the author was visiting the University of California, Berkeley.

Our main result, Theorem 4.2, states that if we assume a bounded continuous prior density ψ on the parameter space and that an observation has mean zero and variance one if $\theta = 0$, then our problem is asymptotically equivalent to the analogous Wiener process problem with drift θ per unit time, the same loss and cost structure and prior "density" $\equiv \psi(0)$. Chernoff's observation applies here also and this asymptotic problem is equivalent to the problem for fixed cost. A formal result in this direction was obtained for the special case of Bernoulli trials by Moriguti and Robbins [18]. Our technique may be viewed as an extension to the sequential case of an approach of Wald [21] and LeCam [14]. It is clearly applicable to other testing, estimation, and general decision problems.

We begin by examining the Wiener process problem and the embedded discrete time normal observation problem for a general continuous and bounded prior density ψ. Our first two results, Lemmas 1.1 and 2.2, establish the asymptotic relation between the Wiener process problem with prior density ψ and the same problem with prior density $\equiv \psi(0)$. Our basic tool is the similarity transform used by Chernoff in [8] and a weak compactness theorem which is a special case of an unpublished result of LeCam. A statement and proof of the latter for our special case is given in the Appendix (Theorem A.1). The validity of this result requires the use of randomized procedures. These are employed throughout the paper, despite the fact that the Bayes procedures for all our problems are nonrandomized. Randomization also plays an important role in considering the relation between the discrete and continuous time problems where we make heavy use of sufficiency. Reference to Chapter 7 of Ferguson [12] may prove helpful.

In Section 3 we show essentially that the exponential family problem is asymptotically at least as hard as the Wiener process problem. To do this we successively, without substantial loss, reduce the problem to one in which observation is carried out in blocks, the parameter space is shrunk to a neighborhood of zero, and the time of observation is truncated. At this stage we use a Berry-Esseen type bound essentially due to Petrov [19] to show that the normal approximation is valid and then apply the results of Section 2. This approximation theorem is given as Lemma 3.3 and its proof is given in the Appendix.

Finally, in the fourth section we show that the Wiener process problem is at least as difficult asymptotically as the exponential family problem. In doing so, we exhibit implicitly a sequence of procedures, independent of ψ, for which the bound of Section 3 is achieved.

Some concluding remarks and statements of open problems are given in the last section.

2. The normal theory problem

In this section we shall describe randomized sequential procedures in continuous and discrete time and derive asymptotic results for the Wiener process problem and its discrete time approximations.

Let $\tilde{C}[0, \infty)$ be the set of all continuous functions defined on $[0, \infty)$ such that $\lim_{t\to\infty} x(t)/t^2 = 0$ endowed with the norm $\|x\| = \sup_t |x(t)|/(1 + t^2)$. The space $\tilde{C}$ is complete separable and metric. Let $\mathscr{B}$ denote the class of Borel sets on $\tilde{C}[0, \infty)$ (the product sigma field) and let $\mathscr{B}_t$ denote the Borel field generated by the maps $x \to x(s)$ for $0 \leqq s \leqq t$.

Let $\Omega = \tilde{C}[0, \infty) \times [0, 1]$, $\mathscr{A}$ be the product Borel field and Q_θ, $-\infty < \theta < \infty$, be the probability measure on $(\Omega, \mathscr{A})$ such that the stochastic process W and random variable U given by $W(x, z) = x$, $U(x, z) = z$ are independent and respectively a Wiener process with drift θ per unit time and a uniformly distributed variable on $[0, 1]$. The subscript θ will be used in this section when calculating expectations with respect to those measures or related measures of the discrete time problem. We are interested in testing $H: \theta \leqq 0$ versus $K: \theta > 0$ with zero-one loss and cost c per unit time. A sequential procedure $\pi = (\delta, \tau)$ for this problem consists of a randomized stopping time τ and a randomized rule δ. Rigorously τ is a measurable map from Ω to $[0, \infty)$ such that for every $z \in [0, 1]$ and $t \leqq \infty$ the event $[\tau(\cdot, z) < t] \in \mathscr{B}_t$. To describe δ we begin by defining the pre τ field $\bar{\mathscr{B}}_\tau$. This is simply the class of all events $A \in \mathscr{A}$ such that for every $z \in [0, 1]$ and every $t \leqq \infty$ the z section of $A \cap [\tau < t]$, that is, $\{x: (x, z) \in A \cap [\tau < t]\}$, is $\mathscr{B}_t$ measurable. Given τ, δ is any map from Ω to $[0, 1]$ which is $\bar{\mathscr{B}}_\tau$ measurable. The use of these procedures should be clear. Having observed $U = z$, we employ $\tau(\cdot, z)$ and on stopping reject with probability $\delta(\cdot, z)$ and accept otherwise.

If $l(d, \theta)$ is our zero-one loss function we write the contitional risk of π given θ for observation cost c as,

$$(2.1)\qquad \begin{aligned} R_\theta(\pi, c) &= E_\theta[l(\delta, \theta)] + cE_\theta(\tau) \\ &= \varepsilon(\theta)\, E_\theta(\delta) + [1 - \varepsilon(\theta)]E_\theta(1 - \delta) + cE_\theta(\tau), \end{aligned}$$

where $\varepsilon(\theta) = 1$ if $\theta \leqq 0$ and 0 otherwise.

Let M_c be the bimeasurable transformation of Ω onto itself given by,

$$(2.2)\qquad M_c(x, z)(t) = \left[\frac{1}{\sqrt{c}}\, x(ct), z\right].$$

This is the similarity transformation suitable for this problem. Then,

$$(2.3)\qquad P_\theta M_c^{-1} = P_{\theta\sqrt{c}}.$$

M_c induces a mapping of the space of decision procedures onto itself as follows:

$$(2.4)\qquad M_c: \pi \to \pi_c = (\tau_c, \delta_c)$$

where

$$(2.5)\qquad \tau_c(x, z) = c\tau[M_c(x, z)],$$

$$(2.6)\qquad \delta_c(x, z) = \delta[M_c(x, z)].$$

Then,

$$E_\theta(\delta_c) = E_{\theta\sqrt{c}}(\delta), \qquad E_\theta(\tau_c) = cE_{\theta\sqrt{c}}(\tau), \tag{2.7}$$

so that

$$R_\theta(\pi_c, 1) = R_{\theta\sqrt{c}}(\pi, c). \tag{2.8}$$

Let ψ be any nonnegative measurable function on R. Define

$$\underline{R}(\psi, c) = \inf_\pi \int_{-\infty}^{\infty} R_\theta(\pi, c)\psi(\theta)\, d\theta. \tag{2.9}$$

LEMMA 2.1.

$$\frac{1}{\sqrt{c}} \underline{R}(\psi, c) = \underline{R}(\psi(\cdot\sqrt{c}), 1). \tag{2.10}$$

PROOF. By (2.8),

$$\begin{aligned} \int_{-\infty}^{\infty} R_\theta(\pi, c)\psi(\theta)\, d\theta &= \sqrt{c} \int_{-\infty}^{\infty} R_{\theta\sqrt{c}}(\pi, c)\psi(\theta\sqrt{c})\, d\theta \\ &= \sqrt{c} \int_{-\infty}^{\infty} R_\theta(\pi_c, 1)\psi(\theta\sqrt{c})\, d\theta. \end{aligned} \tag{2.11}$$

Since the correspondence between π and π_c is one to one onto, the result follows by taking the infima over π on both sides.

All limits in the sequel are taken as $c \to 0$.

LEMMA 2:2. *Let ψ be as above, bounded and continuous at zero. Then,*

$$\lim \frac{1}{\sqrt{c}} \underline{R}(\psi, c) = \psi(0)R^*(1) \tag{2.12}$$

where $R^*(1) = \underline{R}(1, 1) = \inf_\pi \int_{-\infty}^{\infty} R_\theta(\pi, 1)\, d\theta$.

PROOF. Note that $R^*(1)$ is finite (see, for example, the procedure of [5]). By Lemma 1.1, our hypothesis, and the dominated convergence theorem, we must have,

$$\begin{aligned} \lim \frac{1}{\sqrt{c}} \underline{R}(\psi, c) &= \lim \underline{R}[\psi(\cdot\sqrt{c}), 1] \\ &\leqq \psi(0)R^*(1). \end{aligned} \tag{2.13}$$

On the other hand by Theorem A.1 there exists a procedure $\pi(c)$ such that $\underline{R}(\psi(\cdot\sqrt{c}), 1) = \int_{-\infty}^{\infty} R_\theta(\pi, 1)\psi(\theta\sqrt{c})\, d\theta$. Further given any sequence $c_n \downarrow 0$ there exists a procedure π and a subsequence $\{n_k\}$ such that,

$$R_\theta(\pi, 1) \leqq \underline{\lim}_k R_\theta(\pi(c_{n_k}), 1). \tag{2.14}$$

Then by Fatou's lemma, and the continuity of ψ,

$$\liminf_k R[\psi(\cdot, \sqrt{c_{n_k}}), 1] \geqq \psi(0) \int_{-\infty}^{\infty} R_\theta(\pi, 1)\, d\theta \tag{2.15}$$
$$\geqq \psi(0) R^*(1).$$

The lemma follows.

Consider our problem with the modification that if you sample beyond time T then there is no terminal loss and no cost for additional sampling. Let $\underline{R}_\theta(\pi, c, T)$ denote the conditional risk given θ for the modified problem. Formally,

$$\underline{R}_\theta(\pi, c, T) = E_\theta(\delta I[\tau \leqq T])\varepsilon(\theta) + [1 - \varepsilon(\theta)] \tag{2.16}$$
$$E_\theta[(1 - \delta) I[\tau \leqq T]] + c E_\theta[\min(\tau, T)],$$

where $I[A]$ is the indicator of A.

Given any procedure π, let $\pi_T = (\delta_T, \tau_T)$ and a truncation of π be defined by $\tau_T = \min(\tau, T)$, $\delta_T = \delta$ if $\tau < T$ and δ_T minimizes the posterior Bayes risk given $\mathscr{B}_T$.

Let

$$\bar{R}^*(1, T) = \inf_\pi \int_{-\infty}^{\infty} R_\theta(\pi_T, 1)\, d\theta, \tag{2.17}$$

$$\underline{R}^*(1, T) = \inf_\pi \int_{-\infty}^{\infty} \underline{R}_\theta(\pi, 1, T)\, d\theta. \tag{2.18}$$

LEMMA 2.3. *Let $K_c \to \infty$ as $c \to 0$, and let $\psi(\theta)$ be as in Lemma 2.2. Then*

$$\liminf_\pi \frac{1}{\sqrt{c}} \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} \underline{R}_\theta\left(\pi, c, \frac{T}{c}\right) \psi(\theta)\, d\theta = \psi(0) \underline{R}^*(1, T), \tag{2.19}$$

$$\liminf_\pi \frac{1}{\sqrt{c}} \int_{-\infty}^{\infty} R_\theta(\pi_{T/c}, c) \psi(\theta)\, d\theta = \psi(0) \bar{R}^*(1, T). \tag{2.20}$$

PROOF. By arguing as in the proof of Lemma 2.1 we have

$$\inf_\pi \frac{1}{\sqrt{c}} \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} \underline{R}_\theta\left(\pi, c, \frac{T}{c}\right) \psi(\theta)\, d\theta = \inf_\pi \int_{-K_c}^{K_c} \underline{R}_\theta(\pi, 1, T) \psi(\theta\sqrt{c})\, d\theta. \tag{2.21}$$

By arguing as in the proof of Lemma 2.2, we get that the right side of (2.21) converges as $c \to 0$ to the right side of (2.19) which proves (2.19). Exactly the same type of arguments prove (2.20) which completes the proof.

LEMMA 2.4.

$$\lim_{T \to \infty} \underline{R}^*(1, T) = R^*(1), \tag{2.22}$$

$$\lim_{T \to \infty} \bar{R}^*(1, T) = R^*(1). \tag{2.23}$$

PROOF. Clearly we have

$$\underline{R}^*(1, T) \leqq R^*(1) \leqq \bar{R}^*(1, T). \tag{2.24}$$

By a weak compactness argument there exists for fixed T a $\tilde{\pi}(T)$ such that $\underline{R}^*(1, T) = \underline{R}(\tilde{\pi}, 1, T)$. Hence

$$\begin{aligned}\bar{R}^*(1, T) - \underline{R}^*(1, T) &= \bar{R}^*(1, T) - \underline{R}(\tilde{\pi}, 1, T) \\ &\leqq R(\tilde{\pi}_T, 1) - \underline{R}(\tilde{\pi}, 1, T) \\ &\leqq \int_{-\infty}^{0} P_\theta\{W(T) > 0\}\, d\theta + \int_0^\infty P_\theta\{W(T) < 0\}\, d\theta.\end{aligned} \tag{2.25}$$

The right side of (2.25) converges to zero as $T \to \infty$ which completes the proof of the lemma.

Before giving our final lemma we review two ways of defining sequential procedures for discrete time problems. Let $\mathscr{X} = R^\infty \times [0, 1]$ be the product of a countable number of copies of R and $[0, 1]$, and let $\mathscr{D}$ be the Borel field on this space. A randomized stopping time τ is now a measurable map from $\mathscr{X}$ to the natural numbers $\{0, 1, 2, \cdots, \infty\}$ such that the event $[\tau(\cdot, z) \leqq n]$ is, for every z and n, measurable with respect to the σ-field $\mathscr{B}_n^*$ generated by the map $(x_1, x_2, \cdots) \to (x_1, x_2, \cdots, x_n)$ on R^∞. We shall always suppose that the probability measure on $\mathscr{D}$ is such that the random sequence $(X_1, X_2, \cdots)$ and the random variable U given by $(X_1, X_2, \cdots)(x_1, x_2, \cdots, z) = (x_1, x_2, \cdots)$ and $U(x_1, x_2, \cdots, z) = z$ are independent and U is uniform on $[0, 1]$. Similarly a decision rule δ is any measurable map from $\mathscr{X}$ to $[0, 1]$ which is measurable with respect to $\tilde{\mathscr{B}}_\tau$ the ψ-field of all events $A \in \mathscr{D}$ such that the z section of $A \cap [\tau \leqq n]$ is in $\mathscr{B}_n^*$ for every n and z.

In this formulation (which we refer to as I) a procedure $\pi = (\delta, \tau)$ has the same interpretation as in the continuous time problem. On the other hand, following Ferguson [12] we can define a stopping rule τ by a sequence of functions $(\psi_0, \psi_1, \psi_2, \cdots)$ where ψ_j is a $\mathscr{B}_j^*$ measurable function from R^∞ to $[0, 1]$ and $\Sigma_{j=0}^\infty \psi_j \leqq 1$. If τ is a stopping time in the sense of (I), then the ψ_j are given by

$$\psi_j(x_1, x_2, \cdots) = \lambda[z: \tau(x_1, x_2, \cdots, z) = j], \tag{2.26}$$

where λ is Lebesgue measure. Conversely, it is a well-known result of Wald and Wolfowitz [23] that given a stopping time in this second mode as $(\psi_0, \psi_1, \cdots)$ there is a stopping time in the sense of I satisfying (2.26) (see the proof of Theorem A.1). Similarly, a terminal decision rule is specified in the second mode as a sequence $(\delta_0, \delta_1, \cdots)$ of functions from R^∞ to $[0, 1]$ such that δ_j is measurable $\mathscr{B}_j^*$. Again given δ of type I,

$$\delta_j(x_1, x_2, \cdots) = \int_0^1 \delta(x_1, x_2, \cdots, z)\, dz \tag{2.27}$$

and by [23] to any policy $((\delta_0, \delta_1, \cdots), (\psi_0, \psi_1, \cdots))$ there corresponds a policy $\pi = (\delta, \tau)$ satisfying (2.26) and (2.27). Now suppose that Q_θ (abusing notation) makes the X_i in $(X_1, X_2, \cdots)$ independent normal random variables with mean θ and variance one. If π is a policy as above we write the conditional risk given θ for the usual sequential testing problem as

$$(2.28)\quad \begin{aligned} R_\theta(\pi, c) &= \varepsilon(\theta)E_\theta(\delta) + [1 - \varepsilon(\theta)]E_\theta(1 - \delta) + cE_\theta(\tau) \\ &= \varepsilon(\theta) \sum_{j=0}^{\infty} E_\theta(\psi_j\delta_j) + [1 - \varepsilon(\theta)] \sum_{j=0}^{\infty} E_\theta[\psi_j(1 - \delta_j)] \\ &\quad + c \sum_{j=0}^{\infty} jE_\theta(\psi_j), \end{aligned}$$

if $\Sigma_{j=0}^{\infty} \psi_j = 1$ a.s. Q_θ, and $= \infty$ otherwise. We shall refer to this as the discrete time normal problem. Evidently any policy π as above for the given Q_θ may be considered as a policy in Wiener process problem with the same risk. We shall want to consider the normal block problem in which we are permitted to sample in blocks of size N only and are told only the block sums $S_N = \Sigma_{i=1}^{N} X_i$, $S_{2N} = \Sigma_{i=1}^{2N} X_i$, and so forth. Of course statistically, because of sufficiency, this last restriction has no effect on the difficulty of the problem. Let $\mathscr{B}(S_N, S_{2N}, \cdots, S_{jN})$ be the σ-field induced on R^∞ by the maps $(x_1, x_2, \cdots) \to (\Sigma_{i=1}^{N} x_i, \Sigma_{i=N+1}^{2N} x_i, \cdots, \Sigma_{i=(j-1)N+1}^{jN} x_i)$. Formally a block procedure π is any procedure in the discrete time problem such that τ only takes on the values $0, N, 2N, \cdots, jN, \cdots$ with probability one, for every z, j.

$$(2.29)\qquad [\tau(\cdot, z) = jN] \in \mathscr{B}(S_N, S_{2N}, \cdots, S_{jN}),$$

and for every c, z, j,

$$(2.30)\qquad [\delta(\cdot, z) \leqq c] \cap [\tau(\cdot, z) = jN] \in \mathscr{B}(S_N, \cdots . S_{jN}).$$

We can now state:

LEMMA 2.5. *For every procedure π in the Wiener process problem there exists a normal block procedure $\pi^{(N)}$ such that,*

$$(2.31)\qquad |R_\theta(\pi, c) - R_\theta(\pi^{(N)}, c)| \leqq Nc.$$

PROOF. In view of our remarks we can give $\pi^{(N)}$ in the second formulation. Define

$$(2.32)\qquad \psi_j^{(N)} = 0 \qquad \text{for} \quad j \neq iN,$$

$$(2.33)\qquad \psi_{iN}^{(N)} = Q_\theta[(i - 1)N < \tau \leqq iN/W(N), W(2N), \cdots, W(iN)].$$

(The / indicates a suitable version of the conditional probability.) Note that since $W(iN)$ is sufficient for $\mathscr{B}_{iN}$, the $\psi_{iN}^{(N)}$ may be chosen independent of θ. Strictly speaking $\psi_{iN}^{(N)}$ is a function on $\tilde{C}[0, \infty]$ not R^∞. However by the usual arguments we may in fact take $\psi_{iN}^{(N)}$ to be a function of the variables $W(N), W(2N), \cdots, W(iN)$ only. It is clear that if the definition (2.33) defines a stopping time at all, it will be a block time. To check that it is a stopping time we need only to show that,

$$\sum_{j=0}^{\infty} \psi_j^{(N)} \leqq 1. \tag{2.34}$$

It is enough to show that $\Sigma_{j=0}^{iN} \psi_j^{(N)} \leqq 1$ for every i. We have (for a suitable choice of the conditional probability)

$$\begin{aligned} 1 &\geqq Q_\theta[\tau \leqq iN/W(N), W(2N), \cdots, W(iN)] \\ &= \psi_0^{(N)} + \sum_{j=1}^{iN} Q_\theta[(j-1)N < \tau \leqq jN/W(N), W(2N), \cdots, W(iN)] \\ &= \psi_0^{(N)} + Q_\theta[0 < \tau \leqq N/W(N), W(2N) - W(N), \\ &\qquad\qquad W(3N) - W(N), \cdots, W(iN) - W(N)] \\ &\quad + Q_\theta[N < \tau \leqq 2N/W(N), W(2N), W(3N) - W(2N), \cdots, \\ &\qquad\qquad W(iN) - W(2N)] \\ &\quad + \cdots + Q_\theta[(i-1)N < \tau \leqq iN/W(N), W(2N), W(3N), \cdots, W(iN)] \\ &= \psi_0^{(N)} + Q_\theta[0 < \tau \leqq N/W(N)] + Q_\theta[N < \tau \leqq 2N/W(N), W(2N)] \\ &\quad + \cdots + Q_\theta[(i-1)N < \tau \leqq iN/W(N). W(2N), \cdots, W(iN)] \\ &= \sum_{j=0}^{iN} \psi_j^{(N)}. \end{aligned} \tag{2.35}$$

Now define the $\delta_i^{(N)}$ by,

$$\psi_{iN}^{(N)} \delta_{iN}^{(N)} = E_\theta\{\delta I[(i-1)N < \tau \leqq iN]/W(N), \cdots, W(iN)\} \quad \text{for} \quad i = 0, 1, \cdots, \tag{2.36}$$

and $\delta_i^{(N)} = 0$ otherwise.

It is clear that $\pi^{(N)} = \big((\psi_0^{(N)}, \cdots), (\delta_0^{(N)}, \delta_1^{(N)}, \cdots)\big)$ is a block procedure and,

$$\sum_{j=0}^{\infty} E_\theta(\psi_j^{(N)} \delta_j^{(N)}) = E_\theta(\delta) \tag{2.37}$$

while

$$\begin{aligned} \sum_{j=0}^{\infty} jE_\theta(\psi_j^{(N)}) &= \sum_{i=0}^{\infty} iN\, E_\theta(\psi_{iN}^{(N)}) \\ &= \sum_{i=0}^{\infty} iN\, Q_\theta[(i-1)N < \tau \leqq iN] \\ &\leqq E_\theta(\tau) + N. \end{aligned} \tag{2.38}$$

The lemma follows.

3. The exponential family problem: lower bound

In this section we introduce the exponential family model and derive a lower bound to the Bayes risk of the testing problem in terms of the Wiener process problem of the previous section. Without loss of generality we shall throughout this section suppose that $X_1, X_2, \cdots$ are the coordinate projections of R^∞ and are thus defined on the space of the previous section. We let P_θ be a probability measure on $\mathscr{X}$ which makes the X_i independent and identically distributed with density f_θ, with respect to some nondegenerate σ-finite measure μ on R. We take f_θ to be the function

$$f_\theta(x) = e^{\theta x - b(\theta)}, \tag{3.1}$$

where θ ranges over a set Θ such that zero is an interior point of Θ. (As in the previous section whatever be θ, U is independent of the X_i and uniformly distributed on $[0, 1]$.) Let ψ be as in Lemmas 2.3 and 2.4 a bounded probability density (with respect to Lebesgue measure) on Θ and continuous at zero. As before we wish to test $H: \theta \leqq 0$ versus $K: \theta > 0$ with zero-one loss, and at cost c per observation. Evidently the definitions of sequential procedure introduced in connection with the normal discrete time problem are appropriate for this exponential family problem also, the only difference being that risks must be calculated under P_θ rather than Q_θ. Since we shall occasionally have to talk about both problems we shall use the superscripts P, Q on expectations where this is necessary to avoid ambiguity.

Note that

$$E_\theta^{(P)}(X_1) = b'(\theta), \qquad \mathrm{Var}_\theta^{(P)}(X_1) = b''(\theta). \tag{3.2}$$

We shall suppose that $b(0) = b'(0) = 0$ and $b''(0) = 1$. The general case reduces to this special one. To see this consider $Y_i = [X_i - b'(0)]/[b''(0)]^{1/2}$. The Y_i are a sequence of observations distributed according to an exponential family with density

$$g_\theta(y) = \exp\{\theta[b''(0)]^{1/2}y - c(\theta)\}, \tag{3.3}$$

with respect to a suitable underlying measure.

If we change parameters to $\eta = \theta[b''(0)]^{1/2}$ we are back in the previous case although this does, of course, give the prior density $[b''(0)]^{1/2}\psi\{\cdot[b''(0)]^{-1/2}\}$ for η. Also note that there is no loss of generality in assuming that the X_i are real valued. If X takes vector values (or even abstract values) and follows a one parameter exponential family with density of the form,

$$f_\theta(x) = e^{\theta t(x) - b(\theta)} \tag{3.4}$$

then $t(X_1)$, $t(X_1) + t(X_2)$, $\cdots$ is a sequence of transitive sufficient statistics (see [12], Chapter 7) for the problem and of course $t(X_i)$ is a random variable following an exponential family probability law of the original form.

For any procedure π (in form I and II) define $B_\theta(\pi, c)$ to be the conditional risk of π given θ. Define the average risk of π, as usual, by

$$B(\pi, c, \psi) = \int_{-\infty}^{\infty} B_\theta(\pi, c)\psi(\theta)\, d\theta, \tag{3.5}$$

and let $\pi^*(c, \psi)$ denote the Bayes procedure for this problem which minimizes $B(\pi, c, \psi)$ over all π. For convenience we refer to these procedures as π^* in the sequel.

We shall prove

THEOREM 3.1. *Under the conditions of this section,*

$$\liminf_{c\to 0} \frac{1}{\sqrt{c}} B(\pi^*, c, \psi) \geqq \psi(0)R^*(1). \tag{3.6}$$

The proof proceeds by a series of lemmas. Let block procedures be defined as in the previous section.

LEMMA 3.1. *For every π, there exists a block procedure $\pi^{(N)}$ such that,*

$$|B_\theta(\pi^{(N)}, c) - B_\theta(\pi, c)| \leqq Nc \tag{3.7}$$

for every θ.

PROOF. The method of proof is the same as that of Lemma 2.5. Define $\pi = ((\psi_0, \psi_1, \cdots), (\delta_0, \delta_1, \cdots))$,

$$\psi_j^{(N)} = 0, \qquad j \neq iN, \qquad \psi_{iN}^{(N)} = \sum_{j=(i-1)N+1}^{iN} E_\theta^{(P)}(\psi_j | S_N, \cdots, S_{iN}), \tag{3.8}$$

and

$$\delta_{iN}^{(N)}\, \psi_{iN}^{(N)} = \sum_{j=(i-1)N+1}^{iN} E_\theta^{(P)}[\delta_j \psi_j | S_N, \cdots, S_{iN}]. \tag{3.9}$$

Crucial use is made as before of the sufficiency of S_{iN} for P_θ on $\mathscr{B}_{iN}^*$ and the independence of the increments of the S_n process.

For any π let $\underline{B}_\theta(\pi, T)$ denote the conditional risk of π given θ for a modified version of the exponential family problem in which there is neither terminal loss nor additional cost of observation incurred after time T. Thus,

$$\begin{aligned}\underline{B}_\theta(\pi, c, T) = \varepsilon(\theta)E_\theta^{(P)}(\delta I[\tau \leqq T]) + cE_\theta^{(P)}[\min(\tau, T)] \\ + [1 - \varepsilon(\theta)]E_\theta^{(P)}[(1 - \delta)I[\tau \leqq T]].\end{aligned} \tag{3.10}$$

Let $\underline{R}_\theta(\pi, c, T)$ denote the same conditional expectation when the observations come from the normal distribution with mean θ and variance one, that is, when the expectation is taken with respect to Q_θ rather than P_θ. We shall also consider truncated procedures π_T defined in the natural way.

LEMMA 3.2. *For every π, there exists a block procedure $\pi^{(N)}$ such that,*

$$|\underline{B}_\theta(\pi^{(N)}, c, T) - \underline{B}_\theta(\pi, c, T)| \leqq Nc. \tag{3.11}$$

PROOF. As in Lemma 3.1.

Note that both lemmas apply to R_θ, $\underline{R}_\theta$ as a special case.

Let $P_{\theta,n}$ be the measure corresponding to the distribution of $S_n = \sum_{i=1}^n X_i$ where the X_i are independent and identically distributed according to f_θ. Let

$Q(\xi, \sigma^2)$ be the measure corresponding to the normal distribution with mean ξ and variance σ^2. Given a signed measure R defined on a σ-field $\mathscr{A}$ let $\|R\| = \sup_{A\in\mathscr{A}} |R(A)|$. Recall that if P, Q are probability measures dominated by a σ-finite measure μ then

$$\|P - Q\| = \frac{1}{2}\int \left| \frac{dP}{d\mu} - \frac{dQ}{d\mu} \right| d\mu. \tag{3.12}$$

We need the following lemma which may be derived in the same fashion as a known result of Petrov [19].

LEMMA 3.3. *Let $\mathscr{F}$ be a family of densities (with respect to Lebesgue measure) on R. Suppose that $Z_1, \cdots, Z_n$ are independent and identically distributed according to $f \in \mathscr{F}$. Let $U_{f,n}$ be the probability induced by $n^{-1/2}\sum_{i=1}^n X_i$ and let Φ be the standard normal measure on R. Suppose that $\mathscr{F}$ satisfies the following conditions:*

(i) *$\mathscr{F}$ is precompact when considered as a subset of L_1 with the usual topology;*

(ii) $c_1(\mathscr{F}) = \sup\{f(x): x \in R, f \in \mathscr{F}\} < \infty$;

(iii) $\int_{-\infty}^{\infty} xf(x)\,dx = 0$. $\quad \int_{-\infty}^{\infty} x^2 f(x)\,dx = 1$ *for every* $f \in \mathscr{F}$:

(iv) $c_2(\mathscr{F}) = \sup\left\{\int_{-\infty}^{\infty} |x|^3 f(x)\,dx : f \in \mathscr{F}\right\} < \infty$.

Then,

$$\sup\{\|U_{f,n} - \Phi\| : f \in \mathscr{F}\} \leqq \frac{c(\mathscr{F})}{\sqrt{n}}. \tag{3.13}$$

PROOF. See Appendix.

LEMMA 3.4. *There exist $K_1, K_2(M)$ such that*

$$\|Q_{(\xi, c^2)} - Q_{(0,1)}\| \leqq K_1|\xi| + K_2|\sigma^2 - 1| \tag{3.14}$$

for all ζ and σ^2 such that $|\sigma^2 - 1| \leqq M$.

PROOF.

$$\begin{aligned} \|Q_{(\xi,\sigma^2)} - Q_{(0,1)}\| &\leqq \|Q_{(\xi,\sigma^2)} - Q_{(\xi,1)}\| + \|Q_{(\xi,1)} - Q_{(0,1)}\| \\ &= \|Q_{(0,\sigma^2)} - Q_{(0,1)}\| + \|Q_{(\xi,1)} - Q_{(0,1)}\|. \end{aligned} \tag{3.15}$$

By (2.12) for $\xi > 0$,

$$\begin{aligned} &\|Q_{(\xi,1)} - Q_{(0,1)}\| \\ &= \frac{1}{2}\left\{\int_{-\infty}^{\xi/2} [\phi(t) - \phi(t-\xi)]\,dt + \int_{\xi/2}^{\infty} [\phi(t-\xi) - \phi(t)]\,dt\right\} \\ &= \Phi(\xi/2) - \Phi(-\xi/2). \end{aligned} \tag{3.16}$$

So, in general we get

$$\|Q_{(\xi,1)} - Q_{(0,1)}\| \leqq K_1|\xi|. \tag{3.17}$$

Similarly for $|\sigma^2 - 1| \leqq M_2$,

$$\|Q_{(0,\sigma^2)} - Q_{(0,1)}\| \leqq K_2|\sigma^2 - 1|. \tag{3.18}$$

From Lemmas 3.3, 3.4 we obtain

LEMMA 3.5. *Suppose that* $\{Z_i\}$ *are independent and identically distributed with density* g_θ *(with respect to Lebesgue measure) where, the set* $\{g_\theta: |\theta| \leqq \varepsilon\}$ *satisfies conditions* (i), (ii) *and* (iv) *of Lemma* 3.3 *for some* $\varepsilon > 0$, *and further*

$$\begin{aligned} e(\theta) &= E_\theta(Z_1) = \theta + O(\theta^2) \\ v(\theta) &= V_\theta(Z_1) = 1 + O(|\theta|). \end{aligned} \tag{3.19}$$

Let U_{θ_n} *denote the distribution of* $(1/\sqrt{n})\Sigma_{i=1}^n Z_i$.

Then there exists a $\delta > 0$ *and constants* d_1, d_2, d_3 *such that*

$$\sup\{\|U_{\theta_n} - Q_{\theta_n}\|: |\theta| \leqq \delta\} \leqq \frac{d_1}{\sqrt{n}} + d_2|\theta| + d_3\theta^2\sqrt{n}. \tag{3.20}$$

PROOF.

$$\|U_{\theta_n} - Q_{\theta_n}\| \leqq \|U_{\theta_n} - Q_{(ne(\theta), nv(\theta))}\| + \|Q_{(ne(\theta), nv(\theta))} - Q_{(n\theta, n)}\|. \tag{3.21}$$

By Lemma 3.3 and our assumptions on $\{g_\theta: |\theta| \leqq \delta\}$,

$$\sup\{\|U_{\theta_n} - Q_{(ne(\theta), nv(\theta))}\|: |\theta| \leqq \delta\} \leqq \frac{c(\delta)}{\sqrt{n}} \tag{3.22}$$

for $\delta \leqq \varepsilon$.

On the other hand, by Lemma 3.4,

$$\begin{aligned} \|Q_{(ne(\theta), nv(\theta))} - Q_{(n\theta, n)}\| &= \|Q_{(\sqrt{n}(e(\theta)-\theta), v(\theta))} - Q_{(0,1)}\| \\ &\leqq K(\delta)[\sqrt{n}|e(\theta) - \theta| + |v(\theta) - 1|] \end{aligned} \tag{3.23}$$

for δ sufficiently small. The result follows by (3.19).

REMARK. If μ is dominated by Lebesgue measure and

$$\sup\left\{e^{\theta x}\frac{d\mu}{dx}: |\theta| \leqq M, x \in R\right\} < \infty \tag{3.24}$$

for some $M > 0$, then we may apply Lemma 3.5 to the exponential family and deduce that

$$\|P_{\theta,n} - Q_{\theta,n}\| \leqq \frac{d_1}{\sqrt{n}} + d_2|\theta| + d_3\theta^2\sqrt{n} \tag{3.25}$$

if $|\theta| \leqq M$ for suitable d_1, d_2, d_3.

The following result is well known and is stated without proof.

LEMMA 3.6. *Let $P_1, P_2, \cdots, P_n$; $Q_1, Q_2, \cdots, Q_n$ be probability measures defined on the real line and let $P^{(n)}$, $Q^{(n)}$ be the corresponding n dimensional product measures. Then,*

$$\|P^{(n)} - Q^{(n)}\| \leqq \sum_{i=1}^{n} \|P_i - Q_i\|. \tag{3.26}$$

LEMMA 3.7. *Suppose that μ is dominated by Lebesgue measure and* $\sup\{e^{\theta x}\, d\mu/dx : x \in R, |\theta| \leqq M\} < \infty$. *If $|\theta| \leqq M$, then for any $N_c \geqq 1$*

$$\max\left\{\left|\underline{B}_\theta\left(\pi, c, \frac{T}{c}\right) - \underline{R}_\theta\left(\pi, c, \frac{T}{c}\right)\right|; |B_\theta(\pi_{T/c}, c) - R_\theta(\pi_{T/c}, c)|\right\} \tag{3.27}$$
$$\leqq 2cN_c + \frac{T}{cN_c}(2 + T)\left\{\frac{d_1}{\sqrt{N_c}} + d_2|\theta| + d_3\theta^2\sqrt{N_c}\right\}.$$

PROOF. We give the argument for $\underline{B}_\theta$, that for B_θ is identical.

$$\left|\underline{B}_\theta\left(\pi, c, \frac{T}{c}\right) - \underline{R}_\theta\left(\pi, c, \frac{T}{c}\right)\right| \leqq \left|\underline{B}_\theta\left(\pi^{(N_c)}, c, \frac{T}{c}\right) - \underline{B}_\theta\left(\pi, c, \frac{T}{c}\right)\right| \tag{3.28}$$
$$+ \left|\underline{R}_\theta\left(\pi^{(N_c)}, c, \frac{T}{c}\right) - \underline{R}_\theta\left(\pi, c, \frac{T}{c}\right)\right| + \left|\underline{B}_\theta\left(\pi^{(N_c)}, c, \frac{T}{c}\right) - \underline{R}_\theta\left(\pi^{(N_c)}, c, \frac{T}{c}\right)\right|.$$

By Lemma 3.2 the first two terms on the right side of (3.28) are each bounded by cN_c. Now,

$$\left|\underline{B}_\theta\left(\pi^{(N_c)}, c, \frac{T}{c}\right) - \underline{R}_\theta\left(\pi^{(N_c)}, c, \frac{T}{c}\right)\right| \tag{3.29}$$
$$\leqq 2|E_\theta^{(P)}(\delta^{(N_c)} I[\tau^{(N_c)} \leqq T/c]) - E_\theta^{(Q)}(\delta^{(N_c)} I[\tau^{(N_c)} \leqq T/c])|$$
$$+ c\left|E_\theta^{(P)}\left(\min\left(\tau^{(N_c)}, \frac{T}{c}\right)\right) - E_\theta^{(Q)}\left(\min\left(\tau^{(N_c)}, \frac{T}{c}\right)\right)\right|$$
$$\leqq 2\|P_\theta \underset{\sim}{S}^{-1} - Q_\theta \underset{\sim}{S}^{-1}\| + T\|P_\theta \underset{\sim}{S}^{-1} - Q_\theta \underset{\sim}{S}^{-1}\|,$$

where $\underset{\sim}{S}$ maps $(x_1, x_2, \cdots, z)$ into

$$\left(\sum_{i=1}^{N_c} x_i, \sum_{i=N_c+1}^{2N_c} x_i, \cdots, \sum_{i=(I_c-1)N_c+1}^{I_cN_c} x_i\right) \tag{3.30}$$

and $I[c] = [T/cN_c] + 1$.

Applying Lemma 3.5 (and the following remark) and Lemma 3.6 to (3.29), the result follows.

We are now able to prove Theorem 3.1. We begin by proving the theorem in the case μ is dominated by Lebesgue measure and $e^{\theta x}\, d\mu/dx$ is bounded in x for θ in some neighbourhood of zero.

Let K_c, N_c be positive numbers to be determined below. We have by the previous lemmas the following relations

$$(3.31)\qquad \begin{aligned} B(\pi^*, c, \psi) &= \int_{-\infty}^{\infty} B_\theta(\pi^*, c)\psi(\theta)\, d\theta \\ &\geqq \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} B_\theta(\pi^*, c)\psi(\theta)\, d\theta \\ &\geqq \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} \underline{B}_\theta\left(\pi^*, c, \frac{T}{c}\right)\psi(\theta)\, d\theta \\ &\geqq \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} \underline{R}_\theta\left(\pi^*, c, \frac{T}{c}\right)\psi(\theta)\, d\theta \\ &\quad - \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} \left\{2cN_c + \frac{T}{cN_c}(2 + T)\left(\frac{d_1}{\sqrt{N_c}} + d_2|\theta| \right.\right. \\ &\quad \left.\left. + d_3\theta^2\sqrt{N_c}\right)\right\}\psi(\theta)\, d\theta. \end{aligned}$$

Now since $\psi(\theta) \leqq F$ ($\psi(\theta)$ is assumed to be bounded),

$$(3.32)\qquad \begin{aligned} &\frac{1}{\sqrt{c}}\int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} \left\{2cN_c + \frac{T}{cN_c}(2 + T)\left(\frac{d_1}{\sqrt{N_c}} + d_2|\partial| + d_3\theta^2\sqrt{N^c}\right)\right\}\psi(\theta)\, d\theta \\ &\leqq 2FK_c\left\{2cN_c + \frac{d_1T(2 + T)}{cN_c^{3/2}} + \frac{T(2 + T)\, d_2K_c\sqrt{c}}{cN_c} + \frac{T(2 + T)\, d_32K_c^2}{4N_c^{1/2}3}\right\}. \end{aligned}$$

The right side converges to zero for $K_c = c^{-1/8-3\varepsilon}$, $N_c = c^{-7/8+4\varepsilon}$, and $0 < 176\varepsilon < 1$.

On the other hand considering π^* as a procedure for the Wiener process we have

$$(3.33)\qquad \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} \underline{R}\left(\pi^*, c, \frac{T}{c}\right)\psi(\theta)\, d\theta \geqq \inf_\pi \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} \underline{R}\left(\pi, c, \frac{T}{c}\right)\psi(\theta)\, d\theta$$

and the result follows by Lemmas 2.3 and 2.4.

To prove the general case, that is, where μ is not dominated by Lebesgue measure consider the following problem.

We observe $Y_1, Y_2, \cdots$ where

$$(3.34)\qquad Y_i = (X_i, Q_i)$$

with X_i as before and $\{Q_i\}$ a sequence of independent identically distributed normal random variables independent of the $\{X_i\}$ with mean $\varepsilon\theta$ and variance ε. Let $W_i = X_i + Q_i$. The sequence $\{\Sigma_{i=1}^n W_i\}$ is sufficient and transitive for this new problem. The W_i are independent identically distributed according to a one parameter exponential family of the form (3.1) with $b'(0) = 0$, $b''(0) = 1 + \varepsilon$. Furthermore the underlying measure μ of this new family satisfies the condition of the remark following Lemma 3.5.

If we let $B^{\varepsilon}(c, \psi)$ be the Bayes risk of the best procedure for the new problem when the cost of observation (per vector) is c and ψ is the prior density on θ, then our initial discussion leads to

$$\liminf_{c\to 0} \frac{B^{\varepsilon}(c, \psi)}{\sqrt{c}} \geqq \frac{\psi(0)}{\sqrt{(1+\varepsilon)}} R^*(1). \tag{3.35}$$

Of course, $B(\pi^*, c, \psi) \geqq B^{\varepsilon}(c, \psi)$ for every $\varepsilon > 0$. The theorem follows.

4. The exponential family problem: upper bound

The basic result of this section is:

THEOREM 4.1. *Under the conditions of Section 2*

$$\limsup \frac{1}{\sqrt{c}} B(\pi^*, c, \psi) \leqq \psi(0) R^*(1). \tag{4.1}$$

In fact, there exists a sequence of procedures $\{\pi^{**}\}$ which is independent of ψ such that

$$\lim \frac{1}{\sqrt{c}} B(\pi^{**}, c, \psi) = \psi(0) R^*(1). \tag{4.2}$$

(*Dependence on c in π^{**} is suppressed for brevity.*)

From Theorems 4.1 and 3.1, we derive immediately our main result:

THEOREM 4.2. *Under the conditions of Section 2,*

$$\lim \frac{1}{\sqrt{c}} B(\pi^*, c, \psi) = \psi(0) R^*(1). \tag{4.3}$$

We shall give the proof of (4.1) in detail for the case where μ satisfies the conditions of the remark following Lemma 3.5 and sketch the additional remarks needed for the general case and the construction of π^{**} at the end.

PROOF. Let π be any procedure for the Wiener process problem and $\pi_{T/c}$ be its truncation at T/c as in Section 2. By Lemma 2.5 there exists a block procedure $\pi_{T/c}^{(N_c)}$ which by construction is truncated at $[T/c] + N_c$ such that

$$|R_{\theta}(\pi_{T/c}^{(N_c)}, c) - R_{\theta}(\pi_{T/c}, c)| \leqq c N_c. \tag{4.4}$$

Consider the following discrete time rule which we shall denote by $\pi^{(e)}$. Take $n_c^{(1)}$ observations. Stop and reject H if $\Sigma_i\{X_i; 1 \leqq i \leqq n_c^{(1)}\} > A_c^{(1)}$, stop and accept H if $\Sigma_i\{X_i; 1 \leqq i \leqq n_c^{(1)}\} < -A_c^{(1)}$. If $|\Sigma_i\{X_i; 1 \leqq i \leqq n_c^{(1)}\}| \leqq A_c^{(1)}$, take $n_c^{(2)}$ further observations and stop and reject H if

$$\sum_i \{X_i; n_c^{(1)} + 1 \leqq i \leqq n_c^{(1)} + n_c^{(2)}\} > A_c^{(2)} \tag{4.5}$$

stop and reject H if

$$|\sum_i \{X_i; n_c^{(1)} + 1 \leqq i \leqq n_c^{(1)} + n_c^{(2)}\}| < -A_c^{(2)}. \tag{4.6}$$

If

$$|\sum_i \{X_i;\, n_c^{(1)} + 1 \leqq i \leqq n_c^{(1)} + n_c^{(2)}\}| \leqq A_c^{(2)}, \tag{4.7}$$

then disregard the first $n_c^{(1)} + n_c^{(1)}$ observations and follow the procedure $\pi_{T/c}^{(N_c)}$. Let $n_c^{(1)} = c^{-1/2+\varepsilon}$, $A_c^{(1)} = c^{-1/4}$, $n_c^{(2)} = c^{-3/4+3\varepsilon}$, $A_c^{(2)} = c^{-3/8+\varepsilon}$, and $N_c = c^{-7/8+3\varepsilon}$, where $176\varepsilon < 1$. In that case for absolutely continuous μ as above, we shall show that

$$\limsup \frac{1}{\sqrt{c}}\left[B(\pi^{(e)}, c, \psi) - R(\pi_{T/c}, c, \psi)\right] \leqq 0. \tag{4.8}$$

Given (4.8) it follows that

$$\begin{aligned} \limsup \frac{1}{\sqrt{c}} B(\pi^*, c, \psi) &\leqq \limsup \frac{1}{\sqrt{c}} \inf_\pi \int_{-\infty}^{\infty} R_\theta(\pi_{T/c}, c)\psi(\theta)\, d\theta \\ &= \psi(0)\bar{R}^*(1, T) \end{aligned} \tag{4.9}$$

by Lemma 2.3. An application of Lemma 2.4 will then complete the proof of Theorem 4.1. To begin the proof of (4.8) note that, for arbitrary K_c,

$$\begin{aligned} &B(\pi^{(e)}, c, \psi) \\ &= \int_{-\infty}^{\infty} B_\theta(\pi^{(e)}, c)\psi(\theta)\, d\theta \\ &\leqq cn^{(1)} + \int_{-\infty}^{0} P_\theta[S_{n_c^{(1)}} > A_c^{(1)}]\psi(\theta)\, d\theta \\ &\quad + \int_{0}^{\infty} P_\theta[S_{n_c^{(1)}} < -A_c^{(1)}]\psi(\theta)\, d\theta \\ &\quad + cn_c^{(2)} \int_{-\infty}^{\infty} P_\theta[|S_{n_c^{(1)}}| \leqq A_c^{(1)}]\psi(\theta)\, d\theta \\ &\quad + \int_{-\infty}^{0} P_\theta[S_{n_c^{(2)}} > A_c^{(2)}]\psi(\theta)\, d\theta + \int_{0}^{\infty} P_\theta[S_{n_c^{(2)}} < -A_c^{(2)}]\psi(\theta)\, d\theta \\ &\quad + \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} B_\theta(\pi_{T/c}^{(N_c)}, c)\psi(\theta)\, d\theta \\ &\quad + \int_{|\theta| > K_c\sqrt{c}} P_\theta[|S_{n_c^{(2)}}| \leqq A_c^{(2)}](1 + T + cN_c)\psi(\theta)\, d\theta. \end{aligned} \tag{4.10}$$

Since $P_\theta[S_n > A]$ is increasing in θ we may for arbitrary $H_c > 0$ bound the right side of (4.10) by

$$\begin{aligned} &cn_c^{(1)} + P_0[|S_{n_c^{(1)}}| > A_c^{(1)}] \\ &\quad + cn_c^{(2)}\{P_{-H_c\sqrt{c}}[S_{n_c^{(1)}} \geqq -A_c^{(1)}] + P_{H_c\sqrt{c}}[S_{n_c^{(1)}} \leqq A_c^{(1)}] + 2FH_c\sqrt{c}\} \end{aligned} \tag{4.11}$$

$$+ P_0[|S_{n_c^{(2)}}| > A_c^{(2)}] + \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} B_\theta(\pi_{T/c}^{(N_c)}, c)\psi(\theta)\, d\theta$$

$$+ (1 + T + cN_c)\{P_{K_c\sqrt{c}}[S_{n_c^{(2)}} \leqq A_c^{(2)}] + P_{-K_c\sqrt{c}}[S_{n_c^{(2)}} \geqq -A_c^{(2)}]\},$$

where F is our bound on ψ. The idea now is to show that for suitable choices of K_c, H_c all of the above are negligible save $\int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} B_\theta(\pi_{T/c}^{(N_c)}, c)\psi(\theta)\, d\theta$ and that this expression can be well approximated by $\int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} R_\theta(\pi_{T/c}^{(N_c)}, c)\psi(\theta)\, d\theta$. We collect the estimates we need in three propositions. All of these employ the well-known inequality (see, for example, Chernoff [6]),

$$P_\theta[S_n \geqq A] \leqq \min_{t \geqq 0} E_\theta^{(P)}(e^{t(S_n - A)}) = \min_{t \geqq 0} e^{n[b(t+\theta) - b(\theta)] - tA}. \tag{4.12}$$

PROPOSITION 4.1.

$$\lim \frac{1}{\sqrt{c}} P_0[|S_{n_c^{(1)}}| \geqq A_c^{(1)}] = 0, \tag{4.13}$$

$$\lim \frac{1}{\sqrt{c}} P_0[|S_{n_c^{(2)}}| \geqq A_c^{(2)}] = 0. \tag{4.14}$$

PROOF. We prove (4.13), and (4.14) is argued similarly. By (4.12),

$$\log P_0[S_{n_c^{(1)}} \geqq A_c^{(1)}] = \min_{t > 0} \{n_c^{(1)} b(t) - tA_c^{(1)}\}. \tag{4.15}$$

Since $b(0) = b'(0) = 0$ and $b''(0) = 1$ for t sufficiently small $b(t) \leqq \frac{2}{3}t^2$. Take $t = A_c^{(1)}/n_c^{(1)}$ to get

$$\log P_0[S_{n_c^{(1)}} \geqq A_c^{(1)}] \leqq -\tfrac{1}{3}c^{-\varepsilon} \to -\infty. \tag{4.16}$$

Applying a similar argument to $\log P_0[S_{n_c^{(1)}} \leqq A_c^{(1)}]$, the result follows.

PROPOSITION 4.2. *If* $H_c = c^{-1/4 - 2\varepsilon}$,

$$\lim \frac{1}{\sqrt{c}} cn_c^{(2)} P_{-H_c\sqrt{c}}[S_{n_c^{(1)}} \geqq -A_c^{(1)}] = 0, \tag{4.17}$$

$$\lim \frac{1}{\sqrt{c}} cn_c^{(2)} P_{H_c\sqrt{c}}[S_{n_c^{(1)}} \leqq A_c^{(1)}] = 0. \tag{4.18}$$

PROOF. By (4.12),

$$\log P_{-H_c\sqrt{c}}[S_{n_c^{(1)}} \geqq A_c^{(1)}] = \min_{t > 0} n_c^{(1)}\{[b(t - H_c\sqrt{c}) - b(-H_c\sqrt{c})] + tA_c^{(1)}\}. \tag{4.19}$$

For c sufficiently small, expanding b about $-H_c\sqrt{c}$ and b' about zero, we get

$$\log P_{-H_c\sqrt{c}}[S_{n_c^{(1)}} \geqq -A_c^{(1)}] \leqq -\tfrac{1}{3} n_c^{(1)} \left(\frac{H c\sqrt{c} - A_c^{(1)}}{n_c^{(1)}} \right)^2 = -\tfrac{1}{3} c^{-\varepsilon}(c^{-\varepsilon} - 1)^2 \to -\infty. \tag{4.20}$$

The result follows and a similar argument establishes (4.18).

In an entirely analogous fashion, we have

PROPOSITION 4.3. *If* $K_c = c^{-1/8 - 3\varepsilon}$,

$$\lim \frac{1}{\sqrt{c}} P_{-K_c\sqrt{c}}[S_{n_c^{(2)}} \geqq -A_c^{(2)}] = 0. \tag{4.21}$$

$$\lim \frac{1}{\sqrt{c}} P_{K_c\sqrt{c}}[S_{n_c^{(2)}} \leqq A_c^{(2)}] = 0. \tag{4.22}$$

As a consequence of Propositions 4.1 through 4.3, to prove (4.8) we need only show that

$$\limsup \frac{1}{\sqrt{c}} \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} [B_\theta(\pi_{T/c}^{(N_c)}, c) - R_\theta(\pi_{T/c}, c)]\psi(\theta)\, d\theta \leqq 0. \tag{4.23}$$

Now in view of (4.4)

$$\frac{1}{\sqrt{c}} \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} |R_\theta(\pi_{T/c}, c) - R_\theta(\pi_{T/c}^{(N_c)}, c)| \psi(\theta)\, d\theta \leqq 2FK_c c N_c \to 0. \tag{4.24}$$

Finally,

$$\frac{1}{\sqrt{c}} \left| \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} B_\theta(\pi_{T/c}^{(N_c)}, c) - R_\theta(\pi_{T/c}^{(N_c)}, c) \right| \psi(\theta)\, d\theta \leqq 2FK_c \sup \left\{ |B_\theta(\pi_{T/c}^{(N_c)}, c) - R_\theta(\pi_{T/c}^{(N_c)}, c)| : |\theta| \leqq K_c\sqrt{c} \right\} \to 0 \tag{4.25}$$

by using Lemma 3.7 and the estimates (3.32). Combining (4.24), (4.25) and (4.23), (4.8) follows.

In the general case proceed as follows. Let $Q_1, Q_2, \cdots$ be a sequence of random variables (measurable functions) defined on the unit interval such that if we put the uniform distribution on [0, 1] the Q_i are independent and normally distributed with mean zero and variance one. We may, of course, think of the Q_i as being defined on $\mathscr{X}$, depending on $(x_1, x_2, \cdots, z)$ through z only. Define $\pi_\varepsilon^{(e)}$ as follows; $\pi_\varepsilon^{(e)}$ agrees with $\pi^{(e)}$ for the first two stages of $\pi^{(e)}$. If

$$\begin{aligned} |\textstyle\sum_i \{X_i;\ 1 \leqq i \leqq n_c^{(1)}\}| &\leqq A_c^{(1)}, \\ |\textstyle\sum_i \{X_i;\ n_c^{(1)} + 1 \leqq i \leqq n_c^{(1)} + n_c^{(2)}\}| &\leqq A_c^{(2)}, \end{aligned} \tag{4.26}$$

then apply $\pi^{(e)}$ to the sequence

$$X_{n_c^{(1)} + n_c^{(2)} + 1} + Q_1, \qquad X_{n_c^{(1)} + n_c^{(2)} + 2} + Q_2, \cdots. \tag{4.27}$$

Formally,

$$(4.28)\qquad \pi_\varepsilon^{(e)}(x_1, x_2, \cdots, z) = \pi^e(x_1, \cdots, x_{n_c^{(1)}+n_c^{(2)}}, \quad x_{n_c^{(1)}+n_c^{(2)}+1} + Q_1(z), \cdots).$$

Arguing as before but now applying Lemma 3.7 to the variables

$$(4.29)\qquad Z_i = (X_{n_c^{(1)}+n_c^{(2)}+i} + Q_i)$$

which are readily seen to satisfy the condition of that lemma, we find that

$$(4.30)\qquad \limsup \frac{1}{\sqrt{c}} B(\pi^*, c, \psi) \leqq \limsup \frac{1}{\sqrt{c}} B(\pi_\varepsilon^{(e)}, c, \psi) \leqq \sqrt{1+\varepsilon}\,\psi(0)\bar{R}^*(1, T).$$

Letting $T \to \infty$ and $\varepsilon \to 0$ the result follows.

To construct a sequence of procedures which achieves the bound a slightly more involved argument is needed. First of all, arguing as before in Section 2, we show that in the Wiener process problem if $cN_cK_c \to 0$ and $T_c \to \infty$ then

$$(4.31)\qquad \lim \frac{1}{\sqrt{c}} \int_{-K_c\sqrt{c}}^{K_c\sqrt{c}} R_\theta(\tilde{\pi}_{T/c}^{(N_c)}, c)\psi(\theta)\, d\theta = \psi(0)R^*(1)$$

where $\tilde{\pi}$ is such that $\int_{-\infty}^{\infty} R_\theta(\tilde{\pi}, 1)\, d\theta = R^*(1)$. Choose $T_c \uparrow \infty$ so that $T_c^2\, c^{1/16-11\varepsilon} \to 0$, and consider the procedures

$$(4.32)\qquad (\tilde{\pi}_{T_c/c}^{(N_c)})^{(e)}$$

corresponding to $\tilde{\pi}_{T/c}^{(N_c)}$ defined in the proof of Theorem 4.1 for T_c varying as above. It is easy to check that if μ satisfies the conditions of the remark following Lemma 3.5, then

$$(4.33)\qquad \limsup \frac{1}{\sqrt{c}} \int_{-\infty}^{\infty} \{B_\theta[(\tilde{\pi}\,{}_{T_c/c}^{(N_c)})^{(e)}, c] - R_\theta\{\tilde{\pi}_{T_c/c}, c\}\,\psi(\theta)\, d\theta \leqq 0.$$

If μ does not satisfy the conditions following Lemma 3.5 the construction is even less explicit. We construct procedures $\pi_c^{(e)}$ corresponding to

$$(4.34)\qquad \tilde{\pi}_{(T_c/c,\,\varepsilon_c)}^{(N_c)}$$

to be defined below with variables $Q_i^{(c)}$ which are independent normal with mean zero and variance $\varepsilon_c \to 0$. It is necessary to examine the proof of Lemma 3.5 carefully since now d_1 will depend on c and n. It is easy to show that there exists a constant d_1^0 independent of n such that if $Z_i = X_i + Q_i^{(c)}$ then

$$(4.35)\qquad d_1(c) \leqq d_1^0 \frac{n}{\sqrt{\varepsilon_c}} \exp\{-\gamma^2\varepsilon_c n/2\},$$

and $d_1(c)$ will remain bounded above for $n = N_c$ provided that $\varepsilon_c \geqq 3\log N_c/\gamma^2 N_c$, say. For T_c, ε_c as above we have for any sequence of procedures $\{\pi\}$

$$\limsup \frac{1}{\sqrt{c}} \int_{-\infty}^{\infty} [B_\theta(\pi_{\varepsilon_c}^{(e)}, c) - R_{\theta, \varepsilon_c}(\pi, c)]\psi(\theta)\, d\theta \leqq 0, \tag{4.36}$$

where $R_{\theta, \varepsilon}$ is the risk of π for the problem in which we observe the Wiener process with drift θ per unit time and variance $1 + \varepsilon$ per unit time. Finally, it follows from the results of Section 2 that

$$\lim \frac{1}{\sqrt{c}} \inf_\pi \int_{-\infty}^{\infty} R_{\theta, \varepsilon_c}(\pi, c)\psi(\theta)\, d\theta = \psi(0)R^*(1). \tag{4.37}$$

Therefore if we take

$$\tilde{\pi}_{(T_c/c, \varepsilon_c)}^{(N_c)} \tag{4.38}$$

to be the truncated block policy corresponding in the sense of Lemma 2.5 to the procedure $\tilde{\pi}_c$ which achieves $\min_\pi \int_{-\infty}^{\infty} R_{\theta, \varepsilon_c}(\pi, c)\, d\theta$ then

$$\left(\tilde{\pi}_{(T/c, c)}^{(N_c)}\right)^{(e)} \tag{4.39}$$

achieve the bound. The theorem is proved.

5. Concluding remarks and open problems

The techniques of this paper are evidently not limited to the zero-one loss function considered. For different bounded loss functions we must use a different similarity transform, make different choices of K_c, H_c, N_c, and so on, obtain a different rate of convergence, but arrive at similar results. For example, if $\ell(\theta, d) = 0$ when d is the right decision and if $\ell(\theta, d) = \min\{|\theta|, 1\}$ when d is the wrong decision, then the Bayes risk of our problem is of the order of $c^{2/3}$ and the limiting coefficients of $c^{2/3}$ is $\psi(0)$ times the Bayes risk of Chernoff's problem [8] with unit cost and Lebesgue prior. We can also treat the problem of testing with shrinking indifference regions, say, of the form $[-A\sqrt{c}, B\sqrt{c}]$ for zero-one loss. The Bayes risk is of order $\sqrt{c}$ again and the coefficient is $\psi(0)$ times the risk of the Wiener process problem with unit cost, Lebesgue prior and indifference region $[-A, B]$. On the other hand if one permits ψ to vary with c, say, $\psi_c(t) = (1/\sqrt{c})\psi(t/\sqrt{c})$ for a fixed prior density, one can under suitable regularity conditions for zero-one loss obtain an asymptotic risk of order $\sqrt{c}$ with coefficient the risk of the Wiener problem with unit cost and prior density ψ. Of course such densities presupposing more and more surety that the parameter is near zero with decreasing cost are not usually reasonable.

It seems that these techniques should also apply to other decision problems for the exponential family at least locally and should prove useful in non-Bayesian problems as well.

The result may also be generalized to nonexponential families by considering, under suitable regularity conditions the variables

$$T_i = \left.\frac{\partial \log f_\theta(X_i)}{\partial \theta}\right|_{\theta = 0}. \tag{5.1}$$

To what extent an ambitious program such as that of LeCam [14] is possible in the sequential case is, however, unclear to us at present.

A great difficulty of the asymptotic theory of this paper is that in general it leads to problems for the Wiener process which, as the works of Chernoff indicate, can be solved at best approximately. In fact, from a (machine) computational point of view it might be easier, for example, to try to calculate the boundary for the Bernoulli process as an approximation to the Wiener boundary. The results of Moriguti and Robbins [17] as well as our paper indicate that such "boundary convergence" as in Schwarz [20] should hold. However, no proof is known to us.

◇ ◇ ◇ ◇ ◇

APPENDIX

We retain the notation of Section 1. Our first aim is to prove the following weak compactness theorem.

THEOREM A.1. *Let* $\pi_n = (\delta_n, \tau_n)$ *be a sequence of procedures in the Wiener process problem. Then, there exists a subsequence* $\{n_k\}$ *and a procedure* $\pi = (\delta, \tau)$ *such that,*

$$\lim_k E_\theta(\delta_{n_k}) = E_\theta(\delta) \tag{A.1}$$

whenever $\lim\sup_k E_\theta(\tau_{n_k}) < \infty$ *and*

$$\liminf_k E_\theta(\tau_{n_k}) \geqq E_\theta(\tau) \tag{A.2}$$

for every θ. (E_θ *are taken with respect to* Q_θ *throughout.*)

The proof proceeds by a series of lemmas.

The following lemma is essentially a special case of Wald's theorem [22].

LEMMA A.1. *Suppose that all of the* τ_n *have common finite range* $\{t_1 < \cdots < t_s\}$. *Then the result of Theorem* A.1 *holds for suitable* $\{n_k\}$ *and for* $\pi = (\delta, \tau)$ *such that* τ *has the same range with* Q_θ *probability one. Furthermore, if* $\pi_n' = (\delta_n', \tau_n')$ *is another sequence of procedures with* τ_n' *having the same range and* $\tau_n' \leqq \tau_n$ *for all n, then we may choose* $\{n_k\}$ *to be the same for both sequences and choose the "limiting"* $\pi' = (\delta', \tau')$ *such that* $\tau' \leqq \tau$.

PROOF. We write the (δ_n, τ_n) in the second form of Section 3, $\tau_n = (\psi_{0n}, \psi_{1n}, \cdots, \psi_{sn})$, $\delta_n = (\delta_{0n}, \delta_{1n}, \cdots, \delta_{sn})$ with

$$\psi_{in}(x) = \lambda[z: \tau_n(x, z) = t_i] \tag{A.3}$$

$$\delta_{in}(x) = \int_0^1 \delta_{in}(x, z)\, dz. \tag{A.4}$$

Apply the weak compactness theorem (for tests) to $L_1(\Omega, \mathscr{B}_{t_j}, Q_0)$ (see Lehmann [15], p. 354) and the diagonal process to obtain a sequence $\{n_k'\}$ and $\mathscr{B}_{t_j}$ measurable functions ψ_j measurable functions ψ_j. $j = 1, \cdots, s$ such that

(A.5) $$\iint \psi_{jn_k}(x)g(x)Q_0(dx, dz) \to \iint \psi_j(x)g(x)Q_0(dx, dz)$$

for every g which is $\mathscr{B}_{t_j}$ measurable and such that $\iint |g(x)| Q_0(dx, dz) < \infty$. (The theorem is applicable since Ω is a complete separable metric space.) The ψ_j are evidently nonnegative. Further, if g is measurable and $Q_0 W^{-1}$ integrable,

(A.6) $$\begin{aligned}\iint \psi_{jn_k}(x)g(x)Q_0(dx, dz) &= E_0[\psi_{jn_k}(W)g(W)] \\ &= E_0\{\psi_{jn_k}(W)E_0[g(W)|\beta_{t_j}]\} \to E_0\{\psi_j(W)E_0[g(W)|\mathscr{B}_{t_j}]\} \\ &= E_0[\psi_j(W)g(W)]\end{aligned}$$

by (A.5).

Therefore,

(A.7) $$\begin{aligned}Q_0\left[\sum_{j=1}^{s} \psi_j(W) > 1\right] &= E_0\left\{\left[\sum_{j=1}^{s} \psi_{jn_k}(W)\right] I\left[\sum_{j=1}^{s} \psi_j(W) > 1\right]\right\} \\ &\to E_0\left\{\sum_{j=1}^{s} \psi_j(W) I\left[\sum_{j=1}^{s} \psi_j(W) > 1\right]\right\}.\end{aligned}$$

By the same argument $E_0[\Sigma_{j=1}^{s} \psi_j(W)] = 1$.

Hence, since on $\mathscr{B}_{t_s}$ the $Q_\theta W^{-1}$ are equivalent

(A.8) $$Q_\theta\left[\sum_{j=1}^{s} \psi_j(W) = 1\right] = 1.$$

Evidently we may choose versions of the ψ_j such that $\psi_j \geqq 0$ and $\Sigma_{j=1}^{s} \psi_j = 1$ for all θ. Finally we conclude that $\psi = (\psi_1, \cdots, \psi_s)$ is a stopping time and

(A.9) $$\begin{aligned}E_\theta(\tau) &= \sum_{j=1}^{s} t_j E_\theta(\psi_j) \\ &= \sum_{j=1}^{s} t_j E_0\{\psi_j(W) \exp [\theta W(t_j) - \tfrac{1}{2}\theta^2 t_j]\} \\ &= \lim_k \sum_{j=1}^{s} t_j E_0\{\psi_{jn_k}(W) \exp [\theta W(t_j) - \tfrac{1}{2}\theta^2 t_j]\} \\ &= \lim_k E_\theta(\tau_{n_k}).\end{aligned}$$

Now we can by diagonalization and a similar argument obtain a further subsequence $\{n_k\}$ and $\mathscr{B}_{t_j}$ measurable functions γ_j such that,

(A.10) $$E_0[\delta_{jn_k}(W)\psi_{jn_k}(W)g(W)] \to E_0[\lambda_j(W)g(W)]$$

for every integrable function g on $\tilde{C}$. Let,

(A.11) $$\delta_j = \lambda_j/\psi_j.$$

Since,

$$E_0[\lambda_j(W)g(W)] \leqq E_0[\psi_j(W)g(W)] \tag{A.12}$$

for every integrable g, we can select δ_j so that $0 \leqq \delta_j \leqq 1$ and, of course, δ_j is $\mathscr{B}_{t_j}$ measurable. Evidently, $((\psi_1, \cdots, \psi_s), (\delta_1, \cdots, \delta_s))$ a policy in the second form and $\{n_k\}$ satisfy (A.1) and (A.2). To obtain the procedure in form I simply define (following Wald and Wolfowitz [23]),

$$\begin{aligned} \tau(x, z) &= t_r \quad \text{if} \quad \sum_{j=1}^{t_r - 1} \psi_j(x) < z \leqq \sum_{j=1}^{t_r} \psi_j(x). \\ \delta(x, z) &= \delta_j(x) \quad \text{on the set} \quad [\tau(x, z) = t_j]. \end{aligned} \tag{A.13}$$

Since $\tau_n \leqq \tau_n'$ the statement of the lemma leads to limiting times (in the second form) with $\Sigma_{j=1}^{\ell}\psi_j'(x) \geqq \Sigma_{j=1}^{\ell}\psi_j(x)$ for every ℓ and x and our second assertion follows from (A.13). The lemma is proved.

LEMMA A.2. *The theorem is valid if it is true that there exists a T such that $Q_0[\tau_n \leqq T] = 1$ for all n. Furthermore, order is preserved in the limit as in Lemma* A.1.

PROOF. Consider a grid $0, T/2^m, 2T/2^m, \cdots, T$. Define $\tau_n^{(m)} = kT/2^m$ if $(k-1)T/2^m < \tau_n \leqq kT/2^m$ for $k = 0, 1, \cdots, 2^m$.

Let $\pi_n^{(m)} = (\tau_n^{(m)}\delta_n)$. (Note that δ_n is $\tilde{\mathscr{B}}_{\tau_n^{(m)}}$ measurable.) Then,

$$E_\theta(\tau_n^{(m)}) - E_\theta(\tau_n) \leqq T/2^m \tag{A.14}$$

and

$$R_\theta(\pi_n^{(m)}) - R_\theta(\pi_n) \leqq T/2^m. \tag{A.15}$$

Extract a subsequence $\{n_k\}$ and limits in the sense of Lemma A.1 $\tau^{(m)}, \delta^{(m)}$ for each of the sequences $\pi_{n_k}^{(m)}$. Since $\tau_n^{(m)} \geqq \tau_n^{(m+1)}$ for every n, we may suppose that $\tau^{(m)} \geqq \tau^{(m+1)}$ for every m. Let $\tau = \lim_m \tau^{(m)}$. Note that $\tilde{\mathscr{B}}_{\tau^{(m)}} \subset \tilde{\mathscr{B}}_{\tau^{(m-1)}}$ for every m and,

$$\tilde{\mathscr{B}}_\tau = \bigcap_m \tilde{\mathscr{B}}_{\tau^{(m)}}. \tag{A.16}$$

Consider the functions $\{\delta^{(m)}\}$. These are $\tilde{\mathscr{B}}_{\tau^{(j)}}$ measurable for $m \geqq j$. Extract a subsequence $\{m_k\}$ by the diagonal process and $\tilde{\mathscr{B}}_{\tau^{(j)}}$ measurable functions $\tilde{\delta}^{(j)}$ such that

$$E_\theta[\delta^{(m)}(W, U)g_j(W, U)] \to E_\theta[\tilde{\delta}^{(j)}(W, U)g_j(W, U)], \tag{A.17}$$

for every g_j which is $\tilde{\mathscr{B}}_{\tau^{(j)}}$ measurable and bounded for every θ. This follows by the weak compactness theorem for test functions applied to $\tilde{\mathscr{B}}_{\tau^{(j)}}$ successively since the Q_θ are all equivalent on $\tilde{\mathscr{B}}_{\tau^{(j)}}$ and the space Ω is complete separable metric. By construction for every θ the $\tilde{\delta}^{(j)}$ form a martingale and in view of (A.16) and by the martingale convergence theorem,

$$\tilde{\delta}^{(j)} \to E_\theta[\tilde{\delta}^{(1)}|\tilde{\mathscr{B}}_\tau] \tag{A.18}$$

a.s. Q_θ for every θ. Let,

$$\delta = E_0[\tilde{\delta}^{(1)}|\tilde{\mathscr{B}}_\tau]. \tag{A.19}$$

Then δ is $\tilde{\mathscr{B}}_\tau$ measurable and

$$E_\theta(\delta) = E_\theta(\tilde{\delta}^{(1)}) = \lim_k E_\theta(\delta^{(m_k)}) = \lim E_\theta(\delta_{n_r}) \tag{A.20}$$

while

$$\begin{aligned} E_\theta(\tau) &= \lim_k E(\tau^{(m_k)}) = \lim_k \lim_r E_\theta(\tau_{n_r}^{(m_k)}) \\ &\leqq \lim_r E_\theta(\tau_{n_r}) \end{aligned} \tag{A.21}$$

by (A.4). The lemma follows.

We complete the proof of the theorem. Given τ_n let $(\tau^{(T)}, \delta^{(T)})$ be the limits guaranteed by Lemma A.2 for a subsequence of the procedures $\pi_n^{(T)} = (\tau_n^{(T)}, \delta_n^{(T)})$ given by

$$\begin{aligned} \tau_n^{(T)} &= \min(\tau_n, T), \\ \delta_n^{(T)} &= \begin{cases} \delta & \text{if} \quad \tau \leqq T \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{A.22}$$

By Lemma A.2 we can find a subsequence $\{n_k\}$ which works for every $T = 1, 2, \cdots$ and such that $\tau^{(j)} \leqq \tau^{(j+1)}$ for every j. Let $\tau = \lim_j \tau^{(j)}$. By the monotone convergence theorem,

$$E_\theta(\tau) = \lim_j E_\theta(\tau^{(j)}) \leqq \lim_k \inf E_\theta(\tau_{n_k}). \tag{A.23}$$

Consider the sequence $\delta^{(j)}$. Tracing back its construction via Lemmas A.1 and A.2 it is easy to see that the ordering $\delta_n^{(j)} \leqq \delta_n^{(j+1)}$ is preserved with Q_θ probability one in the limit. Let $\delta = \sup_j \delta^{(j)}$. Clearly δ is $\tilde{\mathscr{B}}_\tau$ measurable and by the monotone convergence theorem.

$$E_\theta(\delta) = \lim_j \lim_k E_\theta(\delta_{n_k}^{(j)}). \tag{A.24}$$

Therefore,

$$\begin{aligned} &\lim_k \sup |E_\theta(\delta) - E_\theta(\delta_{n_k})| \\ &\leqq \lim_j \sup \lim_k \sup Q_\theta[\tau_{n_k} > j] \\ &\leqq \lim_j \sup \frac{1}{j} \lim_k \sup E_\theta(\tau_{n_k}) = 0, \end{aligned} \tag{A.25}$$

if $\lim\sup E_\theta(\tau_{n_k}) < \infty$. The theorem follows.

PROOF OF LEMMA 3.3. We proceed as in [18]

$$\|U_{f,n} - \Phi\| = \frac{1}{2} \int_{-\infty}^{\infty} |f_n(t) - \phi(t)|\, dt \tag{A.26}$$

where $f_n = dU_{f,n}(t)/dt$ and ϕ is the standard normal density. By the Schwarz and Minkowski inequalities,

$$\|U_{f,n} - \phi\| \leqq \frac{1}{2}\left[\int (1+x)^{-2}\,dx\right]^{1/2} \left\{\int (1+x)^2[f_n(x) - \phi(x)]^2\,dx\right\}^{1/2}$$
$$\leqq C_1\left(\left\{\int [f_n(x) - \phi(x)]^2\,dx\right\}^{1/2} + \left\{\int [xf_n(x) - x\phi(x)]^2\,dx\right\}^{1/2}\right) \quad \text{(A.27)}$$

where C is a numerical constant. Since $C_1(\mathscr{F}) < \infty$ we may apply the Plancherel theorem to obtain

$$\int [f_n(x) - \phi(x)]^2\,dx = 2\pi \int\left[\lambda^n\left(\frac{t}{\sqrt{n}}\right) - e^{-t^2/2}\right]^2 dt, \quad \text{(A.28)}$$

where $\lambda(t) = \int_{-\infty}^{\infty} e^{itx}f(x)\,dx$. Similarly,

$$\int [xf_n(x) - x\phi(x)]^2\,dx = 2\pi\int\left\{\left[\lambda^2\left(\frac{t}{\sqrt{n}}\right)\right]' - \left[e^{-t^2/2}\right]'\right\}^2 dt. \quad \text{(A.29)}$$

It is well known that

$$\left|\lambda_n\left(\frac{t}{\sqrt{n}}\right) - e^{-t^2/2}\right| \leqq \frac{C_3\{C_2(\mathscr{F})\}}{\sqrt{n}}\{|t|^3 + |t|^2\}e^{-t^2/4}, \quad \text{(A.30)}$$

$$\left|\left[\lambda_n\left(\frac{t}{\sqrt{n}}\right)\right]' - [e^{-t^2/2}]'\right| \leqq \frac{C_3\{C_2(\mathscr{F})\}}{\sqrt{n}}\{|t|^3 + |t|^4\}e^{-t^2/4}, \quad \text{(A.31)}$$

for

$$|t| \leqq \frac{C_4\sqrt{n}}{C_2^{1/2}(\mathscr{F})}, \quad \text{(A.32)}$$

and

$$\left|\left[\lambda^n\left(\frac{t}{\sqrt{n}}\right)\right]'\right| \leqq C_5 n^{1/2} C_2^{1/3}(\mathscr{F})\left|\lambda^{n-1}\left(\frac{t}{\sqrt{n}}\right)\right| \quad \text{(A.33)}$$

where $C_1 - C_5$ are numerical constants.

Finally note that since the Riemann Lebesgue lemma holds uniformly on compact sets of L_1, we have

$$\sup\{|\lambda(t)|:|t| \geqq C_4/C_2^{1/2}(\mathscr{F}), f \in \mathscr{F}\} = C_3(\mathscr{F}) < 1. \quad \text{(A.34)}$$

(To prove this note that the map $(f, t) \to |\lambda(t)|$ is continuous on $L_1 \times [-\infty, \infty]$ with $\lambda(-\infty) = \lambda(+\infty) = 0$. Since $|\lambda(t)| < \int |f(t)|\, dt$ for every $t \neq 0$, (A.31) follows.) Now,

$$\text{(A.35)} \qquad \int \left| \lambda^n\left(\frac{t}{\sqrt{n}}\right) - e^{-t^2/2} \right|^2 dt$$

$$\leqq (C_3^2/n) C_2^2(\mathscr{F}) \int_{|t| > C_4 n^{1/2} C^{-1/2}(\mathscr{F})} \{|t|^3 + |t|^2\}^2 e^{-t^2/2}\, dt$$

$$+ \int_{|t| > C_4 n^{1/2} C^{-1/2}(\mathscr{F})} e^{-t^2/2}\, dt$$

$$+ C_3^{n-2}(\mathscr{F}) \int_{|t| > C_4 n^{1/2} C^{-1/2}(\mathscr{F})} \left| \lambda^n \left(\frac{t}{\sqrt{n}}\right) \right| dt$$

$$\leqq C_7 \{C_2^2(\mathscr{F})/n + C_3^{n-2}(\mathscr{F})\} \leqq C^2(\mathscr{F})/n$$

since $C_3 < 1$. A similar estimate can be given for the second term on the right of (A.27). The result follows.

REFERENCES

[1] A. E. ALBERT, "The sequential design of experiments for infinitely many states of nature," *Ann. Math. Statist.*, Vol. 32 (1961), pp. 774–799.

[2] J. A. BATHER, "Bayes procedures for deciding the sign of a normal mean," *Proc. Cambridge Philos. Soc.*, Vol. 58 (1962), pp. 599–620.

[3] S. A. BESSLER, "Theory and applications of the sequential design of experiments k-actions, and infinitely many experiments," Department of Statistics, Stanford University, Technical Report No. 55 (1960).

[4] P. J. BICKEL and J. A. YAHAV, "Asymptotically pointwise optimal procedures in sequential analysis," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1967, Vol. 1, pp. 401–413.

[5] ———, "On testing sequentially the mean of a normal distribution," Stanford Technical Report N.S.F. No. 26 (1967).

[6] H. CHERNOFF, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Statist.* Vol. 23 (1952), pp. 493–507.

[7] ———, "Sequential design of experiments," *Ann. Math. Statist.*, Vol. 30 (1959), pp. 755–770.

[8] ———, "Sequential tests for the mean of a normal distribution," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 1, pp. 79–91.

[9] ———, "Sequential test for the mean of a normal distribution III (small t)," *Ann. Math. Statist.*, Vol. 36 (1965), pp. 28–54.

[10] ———, "Sequential test for the mean of a normal distribution IV (discrete case)," *Ann. Math. Statist.*, Vol. 36 (1965), pp. 55–68.

[11] A. DVORETZKY, A. WALD, and J. WOLFOWITZ, "Elimination of randomization in certain statistical decision procedures and zero-sum two-person games," *Ann. Math. Statist.*, Vol. 22 (1951), pp. 1–21.

[12] T. S. FERGUSON, *Mathematical Statistics*, New York, Academic Press, 1967.

[13] J. KIEFER and J. SACKS, "Asymptotically optimum sequential inference and design," *Ann. Math. Statist.*, Vol. 34 (1963), pp. 705–750.

[14] L. LECAM. "On the asymptotic theory of estimation and testing hypotheses." *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1956, Vol. 1, pp. 129–156.

[15] E. L. LEHMANN, *Testing Statistical Hypotheses*, New York, Wiley, 1959.

[16] D. V. LINDLEY and B. N. BARNETT, "Sequential sampling: two decision problems with linear losses for binomial and normal random variables," *Biometrika*, Vol. 52 (1965), pp. 507–532.

[17] G. LORDEN, "Integrated risk of asymptotically Bayes sequential tests," *Ann. Math. Statist.*, Vol. 38 (1967), pp. 1399–1422.

[18] S. MORIGUTI and H. ROBBINS, "A Bayes test of '$p \leqq 1/2$' versus '$p > 1/2$'," *Rep. Statist. Appl. Res. Un. Japan Sci. Engrs.*, Vol. 9 (1962), pp. 39–60.

[19] V. V. PETROV, "Asymptotic analysis of some limit theorems in probability," *Vestnik Leningrad Univ.*, Vol. 16 (1961), pp. 51–61. (Also in *Selected Transl. Math. Statist. and Prob.*, Vol. 5 (1965), pp. 179–190.)

[20] G. SCHWARZ, "Asymptotic shapes of Bayes sequential testing regions," *Ann. Math. Statist.*, Vol. 33 (1962), pp. 224–236.

[21] A. WALD, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Trans. Amer. Math. Soc.*, Vol. 54 (1943), pp. 426–482.

[22] ———, *Statistical Decision Functions*, New York, Wiley, 1950.

[23] A. WALD and J. WOLFOWITZ, "Two methods of randomization in statistics and the theory of games," *Ann. Math.*, Vol. 22 (1951), pp. 581–586.