The Annals of Statistics
1985, Vol. 13, No. 1, 403–412

ON THE ASYMPTOTIC BEHAVIOUR OF EMPIRICAL BAYES TESTS FOR THE CONTINUOUS ONE-PARAMETER EXPONENTIAL FAMILY

BY THEO STIJNEN

University of Leiden, Netherlands

Several authors have proposed empirical Bayes tests (EBT) for the continuous one-parameter exponential family for the case that the prior distribution is completely unspecified. They investigated the convergence rate of the (unconditional) Bayes risk, and gave upper bounds for this convergence rate. In this paper it is proposed to study the convergence of the conditional Bayes risk. A method is presented which makes it possible to derive the exact convergence rate of the conditional risk and its limit distribution. Several results are given. Also the question is considered whether monotonizing an empirical Bayes test influences its asymptotic properties.

1. Introduction. We consider empirical Bayes tests (EBT's) for testing H_0: $\theta \le \theta_0$ against H_1: $\theta > \theta_0$ in the exponential family

$$f(x \mid \theta) = m(x)e^{x\theta}h(\theta), \quad -\infty \le a < x < b \le \infty, \tag{1.1}$$

with m positive and continuously differentiable on (a, b). The loss function is $L(\theta, 0) = \max\{\theta - \theta_0, 0\}$ for accepting H_0 and $L(\theta, 1) = \max\{\theta_0 - \theta, 0\}$ for accepting H_1. The parameter θ is distributed according to a completely unknown prior distribution G on the natural parameter space Ω.

Let $X_1, \cdots, X_n$ denote the observations from n independent past experiments: $X_1, \cdots, X_n$ are i.i.d. r.v.'s with (marginal) density $f(x) = \int m(x)e^{x\theta}h(\theta)\, dG(\theta)$. Let X be the observation in the present experiment. Then the conditional Bayes risk of an EBT ϕ_n is defined as $\hat{r}(G, \phi_n) = E[L(\theta, \phi_n(X_1, \cdots, X_n; X)) \mid X_1, \cdots, X_n]$, and the (unconditional) Bayes risk is defined as $r(G, \phi_n) = E\hat{r}(G, \phi_n)$. The minimal attainable risk, which is achieved by a Bayes test w.r.t. G, is denoted by $r(G)$. For a more detailed introduction to this EB testing problem and for further references, see Johns and VanRyzin (J & VR) (1971) and Van Houwelingen (VH) (1976).

J & VR (1971) and VH (1976) introduced EBT's for the above problem. They investigated the convergence rate of $r(G, \phi_n) - r(G)$. In this paper our approach is to study the asymptotic behaviour of $\hat{r}(G, \phi_n) - r(G)$. Whereas J & VR and VH derive only bounds for the convergence rates, our approach leads to exact results.

J & VR (1971) constructed EBT's as follows. Let d denote the Bayes estimator for estimating θ under quadratic loss. Then a Bayes test w.r.t. G can be given by $\phi(x) = I[d(x) > \theta_0]$, where $I[A]$ denotes the indicator function of a set A, and $d(x) = f'(x)/f(x) - m'(x)/m(x)$. An estimate d_n of d can be obtained by

Received January 1983; revised April 1984.
AMS 1980 *subject classifications.* Primary 62C12, secondary 62F05.
Key words and phrases. Empirical Bayes, convergence rate, conditional Bayes risk, monotone test.

using, for instance, kernel estimators for f' and f. This leads to an EBT $\phi_n(x) = I[d_n(x) > \theta_0]$. Unfortunately, since d_n is not necessarily monotone, this EBT is not monotone and therefore not conditionally component admissible (CCA) (cf. VH, 1973, 1977; and Boyer and Gilliland, 1980, for a discussion of the CCA property). VH (1973) proposed to monotonize J & VR's EBT's as follows. Let $\alpha_n(X_1, \cdots, X_n) = \int I[\phi_n^{(x)} = 1] f(x \mid \theta_0)\, dx$, and let $F^{-1}(\cdot \mid \theta_0)$ be the inverse of the c.d.f. corresponding to $f(\cdot \mid \theta_0)$. Then

$$\phi_n^*(X_1, \cdots, X_n; X) = I[X > F^{-1}(1 - \alpha_n(X_1, \cdots, X_n) \mid \theta_0)] \tag{1.2}$$

is a monotone EBT with smaller risk than the original ϕ_n. A formal proof can be found e.g. in Brown et al. (1976).

The EBT's studied in this article are essentially those of J & VR (1971) and the monotonized versions thereof. Our goal is to derive the convergence rate and limit distribution of the conditional Bayes risk.

In Section 2 we define the EBT's to be studied in this paper. In Section 3 we present our results, which are proved in Section 4. Finally, in Section 5, we discuss the results.

2. A class of empirical Bayes tests. The main class of EBT's considered in this paper is constructed as follows. Define $t(x) = \int h(\theta)\exp(x\theta)\, dG(\theta)$. By Theorem 2.9 of Lehmann (1959), t is analytic and its derivative is given by $t'(x) = \int \theta h(\theta)\exp(x\theta)\, dG(\theta)$. Therefore, d can be written as $d = t'/t$. Following Singh (1979), we estimate t' and t as follows. Let $\nu \geq 1$ be a fixed integer, and let K_i $(i = 0, 1)$ be kernel functions satisfying the following conditions.

(2.1)
(i) K_i has finite support and the closure of the convex hull $[A_i, B_i]$ of the support contains zero,

(ii) K_i is absolutely continuous on $[A_i, B_i]$ with bounded derivative,

(iii) $$\int y^j K_i(y)\, dy = 0 \quad \text{if} \quad j = 0, 1, \cdots, \nu - 1, \quad j \neq i,$$
$$= 1 \quad \text{if} \quad j = i.$$

Denote $t^{(0)} = t$ and $t^{(1)} = t'$, then the estimator $t_n^{(i)}$ of $t^{(i)}$ is defined as

$$t_n^{(i)}(x) = (nh_n^{1+i})^{-1} \textstyle\sum_{j=1}^n K_i((x - X_j)/h_n)/m(X_j), \quad i = 0, 1, \tag{2.2}$$

where $(h_n)_{n\geq 1}$ is a sequence of positive real numbers. In order to ensure that $t_n^{(i)}(x)$ is a consistent estimator of $t^{(i)}(x)$, we have to require that

$$h_n \to 0 \quad \text{and} \quad nh_n^3 \to \infty. \tag{2.3}$$

Usually one chooses K_0 and K_1 such that $t_n^{(1)}$ is the derivative of $t_h^{(0)}$. This choice yields automatically the same ν and h_n for both estimators.

Since $d = t'/t$, d can be estimated by $t_n^{(1)}/t_n^{(0)}$. In order to avoid difficulties when the denominator becomes zero, let a_n and b_n be sequences, which may depend on $X_1, \cdots, X_n$, such that

$$t_n(x) > 0 \quad \text{on} \quad [a_n, b_n], \qquad a_n \to_P a \quad \text{and} \quad b_n \to_P b. \tag{2.4}$$

Now we define

$$d_n(x) = t_n^{(1)}(x)/t_n^{(0)}(x) \quad \text{if} \quad a_n \le x \le b_n,$$
$$(2.5) \qquad = d_n(a_n) \quad \text{if} \quad x < a_n,$$
$$= d_n(b_n) \quad \text{if} \quad x > b_n.$$

We assume that a_n and b_n are chosen such that

$$(2.6) \qquad \sup\{|d_n(x) - d(x)|, x \in [a_n, b_n]\} \to_P 0.$$

In Section 5 we shall indicate a possibility to construct a_n and b_n such that (2.4) and (2.6) are fulfilled.

The main class of EBT's to be studied in this paper is given by

$$(2.7) \qquad \phi_n(x) = I[d_n(x) > \theta_0],$$

where d_n is defined in (2.5). According to (1.2) the monotonized version of ϕ_n is given by

$$(2.8) \qquad \phi_n^*(x) = I[x > c_n],$$

where $c_n = F^{-1}(1 - \alpha_n \mid \theta_0)$, with α_n defined as

$$(2.9) \qquad \alpha_n = \int_a^b I[d_n(x) > \theta_0] f(x \mid \theta_0)\, dx.$$

3. Main results. In this section we shall present two theorems that describe the asymptotic behaviour of the conditional Bayes risk of the EBT's introduced in the previous section. Before formulating the results, a theorem is mentioned which plays a central role in the derivation of the results.

We begin by introducing a useful expression for the conditional Bayes risk. It is easy to see that for an arbitrary (non-EB) test ϕ the Bayes risk can be written as

$$(3.1) \qquad r(G, \phi) = \int_a^b \phi(x) g(x)\, dx + \int_\Omega L(\theta, 0)\, dG(\theta),$$

where $g(x) = f(x)(\theta_0 - d(x))$. To ensure that the Bayes risk (3.1) is finite we assume that

$$(3.2) \qquad E|\theta| < \infty.$$

Further, in order to avoid degeneracy, we shall suppose that

$$(3.3) \qquad \lim_{x \downarrow a} d(x) < \theta_0 < \lim_{x \uparrow b} d(x).$$

In particular, (3.3) implies that G is nondegenerate. Since $d'(x)$ can be written as $d'(x) = \mathrm{var}\{\theta \mid X = x\}$, this implies that d is strictly increasing, so $c = d^{-1}(\theta_0)$ is well defined and $c \in (a, b)$. Since $\phi(x) = I[d(x) > \theta_0]$, the following expression is obtained from (3.1):

$$(3.4) \qquad r(G, \phi) - r(G) = \int_a^c I[\phi(x) = 1] g(x)\, dx - \int_c^b I[\phi(x) = 0] g(x)\, dx.$$

Let $D_n = \hat{r}(G, \phi_n) - r(G)$, with ϕ_n the EBT that was defined by (2.7). Then according to (3.4) we have

$$D_n = \int_a^c I[d_n(x) > \theta_0]g(x)\,dx - \int_c^b I[d_n(x) \le \theta_0]g(x)\,dx. \tag{3.5}$$

For the monotonized version ϕ_n^* (see (2.8) and (2.9)) we define $D_n^* = \hat{r}(G, \phi_n^*) - r(G)$. To investigate the asymptotic behaviour of D_n^* the same method can be used as for D_n. This can be seen as follows. Define $\alpha = 1 - F(c \mid \theta_0)$, analogously to α_n (cf. (2.9)). Then we have

$$\alpha_n - \alpha = \int_a^c I[d_n(x) > \theta_0]f(x \mid \theta_0)\,dx - \int_c^b I[d_n(x) \le \theta_0]f(x \mid \theta_0)\,dx. \tag{3.6}$$

Notice that this expression is very similar to (3.5). Therefore the limit distribution of $\alpha_n - \alpha$ can be determined completely analogously to that of D_n. Once that is done, the limit distribution of D_n^* is derived as follows. According to the definition of c_n we have $c_n - c \sim_d -f(c \mid \theta_0)^{-1}(\alpha_n - \alpha)$, where $\sim_d$ means that both sides have identical limit distributions. From (2.8) and (3.4) one can easily see that D_n^* can be written as $D_n^* = \int_{c_n}^c g(x)\,dx$. Since g is continuously differentiable and $g(c) = 0$, it follows that $D_n^* \sim_d -\tfrac{1}{2}g'(c)(c - c_n)^2$, so that

$$D_n^* \sim_d -\tfrac{1}{2}g'(c)f(c \mid \theta_0)^{-2}(\alpha_n - \alpha)^2. \tag{3.7}$$

The investigation of the limit behaviour of D_n and $\alpha_n - \alpha$ is based on approximation of d_n by a suitable Gaussian process. The following theorem is used.

THEOREM 1 (Stijnen, 1982). *Let d_n be defined by* (2.2) *and* (2.5), *and suppose that* (2.1) *and* (2.3) *are satisfied. Then*

$$d_n(x) =_d d(x) + \beta_n f(x)^{-1/2} W(x/h_n) + \gamma_n u_n(x) + R_n(x), \quad x \in (a, b), \tag{3.8}$$

where $\beta_n = (nh_n^3)^{-1/2}$, $\gamma_n = h_n^\nu$ and $=_d$ means that the stochastic processes on both sides are identically distributed. Furthermore, the following holds:

(i) *W is a stationary zero mean Gaussian process with covariance function*

$$v(t) = \int K_1(t + z)K_1(z)\,dz. \tag{3.9}$$

(ii) *$(u_n)_{n\ge1}$ is a sequence of real functions uniformly converging on compact intervals to the function u defined by*

$$u(x) = C_1 \frac{t^{(\nu+1)}(x)}{t(x)} - C_0 \frac{t^{(\nu)}(x)t^{(1)}(x)}{t(x)^2}. \tag{3.10}$$

with $C_i = \int (-z)^\nu(\nu!)^{-1}K_1(z)\,dz$, $i = 0, 1$.

(iii) *$(nh_n^2)^{1/2} \mid R_n(x) \mid$ is bounded in probability uniformly in x on every compact interval contained in (a, b).*

From (3.8) it follows that the convergence rate of d_n is fastest if $(nh_n^3)^{-1/2} \sim h_n^\nu$. Although our method is capable of managing the general case in which only (2.3) is assumed, we restrict ourselves for the sake of simplicity to this important particular case. Therefore we assume that

$$h_n \sim Cn^{-1/(2\nu+3)}, \tag{3.11}$$

where C is an arbitrary positive constant. Since the constant C may be incorporated in the kernel function K, we assume without loss of generality that $C = 1$.

Before giving our main results, we introduce the following random variables. Let $\lambda < 0$ and η be arbitrary constants, and let W be the Gaussian process defined in Theorem 1. Then we define

$$D(\eta, \lambda) = \int_0^\infty yI[W(y) < \lambda y + \eta]\, dy - \int_{-\infty}^0 yI[W(y) > \lambda y + \eta]\, dy$$

and

$$D^*(\eta, \lambda) = \int_0^\infty I[W(y) < \lambda y + \eta]\, dy - \int_{-\infty}^0 I[W(y) > \lambda y + \eta]\, dy.$$

Since the sample paths of W are continuous (Cramér and Leadbetter, 1967, page 170), the above integrals are well defined as sample path integrals in the ordinary sense. Since $\max\{|W(x)|,\ 0 \le x \le t\}(2 \log t)^{-1/2}$ tends to 1 almost surely as $t \to \infty$ (Pickands, 1967), each of the four integrals is finite with probability one. Therefore D and D^* are well defined random variables for all $\lambda < 0$ and η.

The following theorems give the asymptotic behaviour of D_n and D_n^*.

THEOREM 2. *Let ϕ_n be defined by* (2.2), (2.5) *and* (2.7). *Suppose that* (2.1), (2.3), (2.4), (2.6), (3.2), (3.3) *and* (3.11) *hold.*

(i) *If $\nu = 1$ then*

$$nh_n^3 D_n \to_d d'(c)f(c)D(\eta, \lambda), \tag{3.12}$$

where $\lambda = -d'(c)f(c)^{1/2}$ and $\eta = -u(c)f(c)^{1/2}$.

(ii) *If $\nu > 1$ then*

$$nh_n^3 D_n \to_d \tfrac{1}{2} d'(c)^{-1}\{U + f(c)^{1/2}u(c)\}^2, \tag{3.13}$$

where U is a normally distributed stochastic variable with expectation zero and variance equal to $\int K_1(z)^2\, dz$.

THEOREM 3. *Suppose that the conditions of Theorem 2 are fulfilled. Let ϕ_n^* be the monotonized version of ϕ_n defined by* (2.8) *and* (2.9).

(i) *If $\nu = 1$, then, with η and λ as in Theorem 2,*

$$nh_n^3 D_n^* \to_d \tfrac{1}{2} d'(c)f(c)D^*(\eta, \lambda)^2. \tag{3.14}$$

(ii) *If $\nu > 1$ then $nh_n^3 D_n^*$ converges in distribution to the r.h.s. of* (3.13).

In the next section Theorem 2 will be proved. Since the limit behaviour of D_n^* easily follows from that of $\alpha_n - \alpha$ (see (3.7)), and the expressions for D_n and $\alpha_n - \alpha$ are very similar (cf. (3.5) and (3.6)), the proof of theorem 3 is essentially the same and is therefore omitted.

4. Proof of Theorem 2. Let $\delta > 0$ be such that $[c - \delta, c + \delta] \in (a, b)$, and define (cf. (3.5))

$$D_n^\delta = \int_{c-\delta}^{c} I[d_n(x) > \theta_0] g(x)\, dx - \int_{c}^{c+\delta} I[d_n(x) \le \theta_0] g(x)\, dx, \tag{4.1}$$

then it follows from (2.6) and the fact that d is strictly increasing that $D_n^\delta - D_n = o_p(\beta_n^2)$. According to Theorem 1, d_n can be approximated in distribution by $\hat{d}_n(x) = d(x) + \beta_n f(x)^{-1/2} W(x/h_n) + \gamma_n u_n(x)$. Let $\hat{D}_n^\delta$ be defined by substituting $\hat{d}_n$ for d_n in (4.1). In Lemma 1 below it is proved that $D_n^\delta =_d \hat{D}_n^\delta + o_p(\beta_n^2)$. Therefore it is sufficient to derive the limit distribution of $\hat{D}_n^\delta$.

Let $\psi_n(x) = f(x)^{1/2}(\theta_0 - d(x) - \gamma_n u_n(x))$, then $\hat{d}_n(x) > \theta_0$ is equivalent to $W(x/h_n) > \beta_n^{-1}\psi_n(x)$. With $x = c + \beta_n y$, $\hat{D}_n^\delta$ can be written as

$$\begin{aligned} \hat{D}_n^\delta = \beta_n \Bigg\{ &\int_{-\delta/\beta_n}^{0} I\left[W\left(\frac{c + \beta_n y}{h_n}\right) > \psi_n(c + \beta_n y)/\beta_n \right] g(c + \beta_n y)\, dy \\ &- \int_0^{\delta/\beta_n} I\left[W\left(\frac{c + \beta_n y}{h_n}\right) \le \psi_n(c + \beta_n y)/\beta_n \right] g(c + \beta_n y)\, dy \Bigg\}. \end{aligned} \tag{4.2}$$

Since W is a stationary process, the distribution of the r.h.s. of (4.2) does not change if $W((c + \beta_n y)/h_n)$ is replaced by $W(\beta_n y/h_n)$. From the definition of β_n (see Theorem 1) and (3.11) one can see that $W(\beta_n y/h_n) \to_d W(C_0 y)$, with $C_0 = 1$ if $\nu = 1$, and $C_0 = 0$ if $\nu > 1$. Further, since $g(c) = 0$, $g(c + \beta_n y)$ behaves asymptotically as $\beta_n y g'(c)$. Finally one can easily verify that $\psi_n(c + \beta_n y)/\beta_n \to -f(c)^{1/2}\{d'(c)y + u(c)\}$.

Taking these limits in (4.2) yields

$$\begin{aligned} \beta_n^{-2}\hat{D}_n^\delta \to_d g'(c) \Bigg\{ &\int_{-\infty}^{0} I[W(C_0 y) > -f(c)^{1/2}(d'(c)y + u(c))] y\, dy \\ &- \int_0^{\infty} I[W(C_0 y) < -f(c)^{1/2}(d'(c)y + u(c))] y\, dy \Bigg\}. \end{aligned} \tag{4.3}$$

A formal proof of (4.3) is given in Lemma 2. For $\nu = 1$, (4.3) yields the first part of Theorem 2 (observe that $g'(c) = -d'(c)f(c)$). If $\nu > 1$, then $C_0 = 0$ and the r.h.s. of (4.3) reduces to

$$\tfrac{1}{2} g'(c)\, d'(c)^{-2}(W(0)f(c)^{-1/2} - u(c))^2.$$

Since $W(0) \simeq \mathcal{N}(0, \int K_1(z)^2\, dz)$, this yields the second part of Theorem 2. □

LEMMA 1. *Under the conditions of Theorem* 2, $D_n^\delta =_d \hat{D}_n^\delta + o_p(\beta_n^2)$.

PROOF. Let $\tilde{D}_n^\delta$ be defined by substituting $\hat{d}_n + R_n$ for d_n in (4.1). Then, according to Theorem 1, $\tilde{D}_n^\delta =_d D_n^\delta$, so we have to prove that $|\tilde{D}_n^\delta - \hat{D}_n^\delta| = o_p(\beta_n^2)$. With $S_n = \sup\{|R_n(x)|, |x - c| \le \delta\}$ we have

$$|\tilde{D}_n^\delta - \hat{D}_n^\delta| \le \int_{c-\delta}^{c+\delta} I[|\hat{d}_n(x) - \theta_0| < S_n]|g(x)|\,dx,$$

and we can write

$$|\tilde{D}_n^\delta - \hat{D}_n^\delta| \le \int_{c-\delta}^{c+\delta} I\left[\left|W\left(\frac{x}{h_n}\right) - \frac{\psi_n(x)}{\beta_n}\right| < S_n f(x)^{1/2}\beta_n^{-1}\right]|g(x)|\,dx.$$

Since $S_n = O_p(\beta_n h_n^{1/2})$ and f is bounded on $[c - \delta, c + \delta]$, it is sufficient to prove that

$$\beta_n^{-2} \int_{c-\delta}^{c+\delta} I\left[\left|W\left(\frac{x}{h_n}\right) - \frac{\psi_n(x)}{\beta_n}\right| < h_n^{1/4}\right]|g(x)|\,dx \to_P 0.$$

We shall prove this by showing that the expectation of the l.h.s. tends to zero. With $x = c + \beta_n z$ this expectation can be written as

$$\int_{-\delta/\beta_n}^{\delta/\beta_n} \left|\Phi\left(\frac{\psi_n(c + \beta_n z)}{\beta_n} + h_n^{1/4}\right) - \Phi\left(\frac{\psi_n(c + \beta_n z)}{\beta_n} - h_n^{1/4}\right)\right| \beta_n^{-1}|g(c + \beta_n z)|\,dz,$$

where Φ denotes the distribution function of $\mathcal{N}(0, \int K_1^2(z)\,dz)$ distribution. Since $\beta_n^{-1}g(c + \beta_n z) = zg'(c + \beta_n z^*)$ for some $0 \le |z^*| \le |z|$ and g' is bounded on $[c - \delta, c + \delta]$, it is sufficient to show that (use the Mean Value Theorem)

$$\int_{-\delta/\beta_n}^{\delta/\beta_n} |z|\,\Phi'(\psi_n(c + \beta_n z)/\beta_n + h_n^{*1/4})\,dx, \tag{4.4}$$

with $0 < |h_n^*| < h_n$, is bounded. One can easily check that there exist positive constants M_1 and M_2 such that $|\psi_n(c + \beta_n z)/\beta_n + h_n^{*1/4}| \ge M_1|z| + M_2$ for n and $|z|$ large enough. Therefore (4.4) is uniformly bounded in n. This completes the proof. □

LEMMA 2. *Under the conditions of Theorem* 2, (4.3) *holds.*

PROOF. Define for arbitrary $k > 0$ (cf. (4.2))

$$A_n^k = \int_{c-k\beta_n}^{c} I[\hat{d}_n(x) > \theta_0]g(x)\,dx - \int_{c}^{c+k\beta_n} I[\hat{d}_n(x) \le \theta_0]g(x)\,dx. \tag{4.5}$$

Since (i) W is a stationary process, (ii) $\psi_n(c + \beta_n y)/y \to \lambda y + \eta$ uniformly in y and (iii) $\beta_n^{-1} g(c + \beta_n y) \to yg'(c)$ uniformly in y, it follows that

$$\beta_n^{-2}A_n^k \to_d g'(c) \cdot \left\{\int_{-k}^{0} I[W(C_0 y) > \lambda y + \eta]y\,dy - \int_{0}^{k} I[W(C_0 y) < \lambda y + \eta]y\,dy\right\} \tag{4.6}$$

where C_0 is a constant equal to 1 if $\nu = 1$, else $C_0 = 0$. Notice that the r.h.s. of

(4.6) tends to the r.h.s. of (4.3) as $k \to \infty$. Therefore the proof is completed by showing that for arbitrary $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ there exist n_0 and k_0 such that

$$(4.7) \qquad P(\beta_n^{-2} \mid B_n^k \mid > \varepsilon_0) < \varepsilon_1 \quad \text{for} \quad n \ge n_0 \quad \text{and} \quad k \ge k_0,$$

where $B_n^k = \hat{D}_n^\delta - A_n^k$. From the definition of B_n^k it is easy to see that $\beta_n^{-2} E \mid B_n^k \mid$ is bounded by

$$(4.8) \qquad \int_{-\delta/\beta_n}^{-k} \Phi\left(-\frac{\psi_n(c+\beta_n y)}{\beta_n}\right) \mid \beta_n^{-1} g(c+\beta_n y) \mid dy + \int_k^{\delta/\beta_n} \Phi\left(\frac{\psi_n(c+\beta_n y)}{\beta_n}\right) \mid \beta_n^{-1} g(c+\beta_n y) \mid dy.$$

Since $\mid \beta_n^{-1} g(c + \beta_n y) \mid < M_1 \mid y \mid$ and $-\psi_n(c + \beta_n y) < M_2 y + M_3$ for some positive constants M_1, M_2 and M_3, the integrand of the first term of (4.8) is bounded on $(-\infty, -k]$ by the integrable function $M_1 \mid y \mid \Phi(M_2 y + M_3)$. A similar statement holds for the second term of (4.8). Taking limits, one can easily see that (4.8) tends to a constant that depends on k and converges to zero as $k \to \infty$. Now (4.7) follows. This completes the proof. □

5. Discussion. Theorems 2 and 3 enable us to compare the asymptotic properties of ϕ_n and its monotonized version ϕ_n^*. Of course, since ϕ_n is generally not monotone, the small sample behaviour of ϕ_n^* will be better than that of ϕ_n. It turns out that the two EBT's are asymptotically equivalent when $\nu > 1$. However, when $\nu = 1$, the monotonized version is more efficient, because it holds that $P(\frac{1}{2} D^*(\eta, \lambda)^2 < D(\eta, \lambda)) = 1$ for all $\lambda < 0$ and η. This can be verified as follows. Let $A = \{y > 0 : W(y) < \lambda y + \eta\}$ and $B = \{y < 0 : W(y) > \lambda y + \eta\}$. Then we can write

$$\begin{aligned} D^*(\eta, \lambda)^2 &= \left(\int_A dy - \int_B dy\right)^2 < \left(\int_A dy\right)^2 + \left(\int_B dy\right)^2 \\ &= 2 \int_A \int_A I[x < y]\, dx\, dy + 2 \int_B \int_B I[x > y]\, dx\, dy \\ &\le 2 \int_A y\, dy - 2 \int_B y\, dy = 2D(\eta, \lambda). \end{aligned}$$

Theorems 2 and 3 give the limit distributions of the conditional risk of ϕ_n and ϕ_n^* for the special case $h_n = Cn^{-1/(2\nu+3)}$. The more general case $h_n = Cn^{-\delta}$, with $0 < \delta < \frac{1}{3}$, can be handled by the same methods. For a detailed discussion, the reader is referred to Stijnen (1980). It is interesting to compare the asymptotic behavior of ϕ_n with that of ϕ_n^* in the general case $h_n = Cn^{-\delta}$. Therefore we present in Figure 1 the graphs of the convergence rate of D_n and D_n^* as a function of δ. Let $\delta_0(\delta)$ and $\delta_1(\delta)$ be functions of δ such that $n^{\delta_0} D_n$ and $n^{\delta_1} D_n^*$ have a limit distribution that is not degenerate in zero. Then δ_0 and δ_1 turn out to be piecewise linear. Their graphs are shown in Figure 1. From this figure one can see that

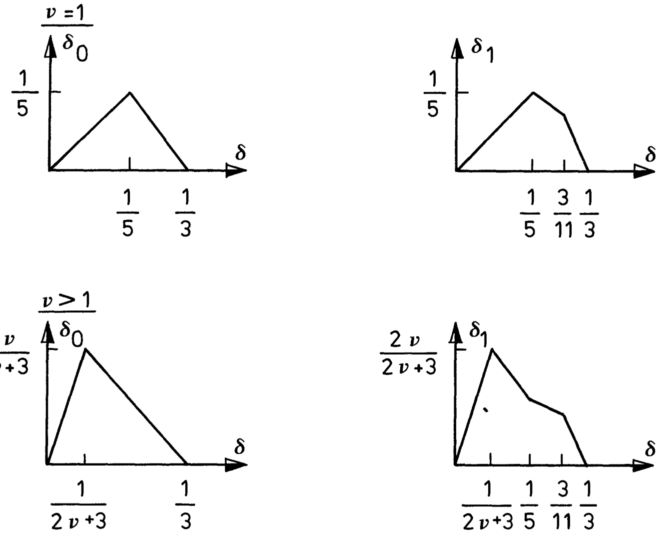

FIG. 1. *Convergence rates of D_n and D_n^*.*

the convergence of D_n^* is faster than that of D_n if $\delta > 1/5$. For $\delta < 1/5$ the convergence rates are equal and also the limit distributions are identical. If $\delta = 1/5$, convergence rates are equal, but the limit distributions are not. This situation is already considered above for the case $\nu = 1$. If $\nu > 1$ an analogous phenomenon occurs.

If in the definition of u (see (3.10)) t is replaced by f, then theorem 2 holds for the class of EBT's introduced by J & VR (1972) (they used the usual kernel density estimators instead of (2.2)). For the monotonized versions of these tests, Theorem 3 holds. The proof is completely analogous to the proof in Section 4. For details the reader is again referred to Stijnen (1980). Our results are not directly comparable with those of J & VR, because they investigated the unconditional risk and gave upper bounds for its convergence rate. However, since usually convergence in distribution indicates convergence in absolute mean, our results provide at least a good guess for what the unknown exact convergence rate and limit constant of the unconditional risk will be.

Sequences a_n and b_n that satisfy (2.4) and (2.6) are easily constructed if some knowledge is available about the rate of convergence of $t_n^{(i)}$ ($i = 0, 1$) in the supremum norm. Suppose for instance that for some sequences $a_n' \downarrow a$ and $b_n' \uparrow b$ it holds that

$$\sup_{[a_n', b_n']} | t_n^{(i)}(x) - t^{(i)}(x) | = o_p(k_n), \tag{5.1}$$

for some sequence $k_n \downarrow 0$. Such results are widely available for the usual kernel density estimators (see Singh (1978), e.g.). These results are easily adapted for the special kernel estimators used here. If (5.1) holds, we can define $[a_n, b_n]$ as the longest interval contained in $[a_n', b_n']$ on which $t_n(x) \geq k_n^{1/3}$ and $t_n'(x) \leq k_n^{-1/3}$. Since t is analytic and strictly positive on (a, b), it follows easily that (2.4) holds.

In order to show that (2.6) holds, we notate $\| f \| = \sup\{ | d(x) |, x \in [a_n, b_n]\}$. Then we can write

$$\sup_{[a_n,b_n]} | d_n(x) - d(x) | = \left\| \frac{t'_n}{t_n} - \frac{t'}{t} \right\|$$

$$\leq \| 1/t_n \| \; \| t'_n - t' \| + \| 1/t_n \| \; \| t' \| \; \| 1/t \| \; \| t_n - t \|$$

$$= k_n^{-1/3} o_p(k_n) + k_n^{-1/3} O_p(k_n^{-1/3}) O_p(k_n^{-1/3}) o_p(k_n) = o_p(1).$$

Therefore, (2.6) is satisfied for $[a_n, b_n]$ defined as above.

Acknowledgement. This paper is based on a part of the author's doctoral thesis written at the University of Utrecht. I would like to thank my thesis advisor, Dr. J. C. van Houwelingen, for his help throughout this work.

REFERENCES

BOYER, J. E. and GILLILAND, C. (1980). Admissibility considerations in the finite state compound and empirical Bayes decision problems. *Statist. Neerlandica* **34** 151–159.

BROWN, L. D., COHEN, A. and STRAWDERMAN, W. E. (1976). A complete class theorem for strict monotone likelihood ratio with applications. *Ann. Statist.* **4** 712–722.

CRAMÉR, H. and LEADBETTER, M. R. (1967). *Stationary and Related Stochastic Processes.* Wiley, New York.

JOHNS, JR., M. V. and VANRYZIN, J. R. (1972). Convergence rates for empirical Bayes two-action problems II. Continuous case. *Ann. Math. Statist.* **43** 934–947.

PICKANDS, J. S. III (1967). Maxima of stationary Gaussian processes. *Z. Wahrsch. verw. Gebiete* **24** 190–223.

SINGH, R. S. (1978). Improvements on some known non-parametric consistent estimators of derivatives of a density. *Ann. Statist.* **5** 394–399.

SINGH, R. S. (1979). Empirical Bayes estimation in Lebesgue-exponential families with rates near the best possible rate. *Ann. Statist.* **7** 890–902.

STIJNEN, TH. (1980). On the asymptotic behaviour of monotonized empirical Bayes rules. Doctoral thesis at the University of Utrecht, Netherlands.

STIJNEN, TH. (1982). Empirical Bayes rules and Gaussian processes. *J. Statist. Plann. Infer.* **6** 363–372.

VAN HOUWELINGEN, J. C. (1973). On empirical Bayes rules for the continuous one-parameter exponential family. Doctoral thesis, University of Utrecht, Netherlands.

VAN HOUWELINGEN, J. C. (1976). Monotone empirical Bayes tests for the continuous one-parameter exponential family. *Ann. Statist.* **4** 981–989.

VAN HOUWELINGEN, J. C. (1977). Monotonizing empirical Bayes estimators for a class of discrete distributions with monotone likelihood ratio. *Statist. Neerlandica* **31** 95–104.

DEPARTMENT OF MEDICAL STATISTICS
STATE UNIVERSITY OF LEIDEN
80 WASSENAARSEWEG
P.O. BOX 9512
2300 RA LEIDEN, THE NETHERLANDS