EMPIRICAL BAYES PROCEDURES WITH CENSORED DATA*

V. Susarla

SUNY-Binghamton

and

John Van Ryzin

Columbia University

This paper extends empirical Bayes estimators with squared error loss and tests with linear loss for two classes of exponential families when the observed data is randomly right censored. Sufficient conditions for proving asymptotic optimality of the procedures are given. Various extensions to multiple action problems and to rate of convergence results are indicated.

1. Introduction.

The empirical Bayes approach of Robbins (1955, 1963, 1964) is applicable to statistical situations when one is confronted with an independent (but not necessarily identical as in O'Bryan and Susarla (1977)) sequence of Bayes decision problems each having similar structure. The statistical similarity in these problems includes the assumption of an unknown prior distribution Λ on the parameter space involved. Robbins (1964) argues that much can be gained by using the empirical Bayes approach which uses the data available in the first n decision problems in the (n + 1)st Bayes decision problem. Since Robbins' initiation of this idea, many papers evolved

*Research supported by NIH Grant 1R01 GM 28405.
AMS 1980 Classifications: 62C12, 62F12.
Key words and phrases: empirical Bayes, asymptotic optimality, survival analysis with censored data.

on developing empirical Bayes procedures and their asymptotic properties as the number of problems, n, approaches ∞. Most of these empirical Bayes methods have treated situations in which the observed data is noncensored data. Work on empirical Bayes problems in which the observed data is randomly right censored data first appeared in Susarla and Van Ryzin (1978). They treat the empirical Bayes problem of estimating a distribution function with the unknown prior involved being a Dirichlet process (Ferguson (1973)). Here, we investigate the empirical Bayes approach to some squared error loss estimation and linear loss k-action problems when the data is randomly right censored. We further assume the family of densities (with respect to the Lebesgue measure on Borel σ-field in $R = (-\infty,\infty)$) is either of the form: (Case I). $f(x, \lambda) = c(\lambda)h(x) \exp(\lambda x)$, or of the form: (Case II). $f(x,\lambda) = c(\lambda)\lambda^x h(x)$. Here λ is a generic element in the natural parameter space Ω of the exponential family involved and $h(x) > 0$ if and only if $x > a$.

Section 2 describes the estimation problem and its empirical Bayes analogue with censored data. It also introduces notation used throughout the rest of the paper. Section 3 constructs empirical Bayes estimators under the two exponential families cited above, and obtains sufficient conditions under which the proposed empirical Bayes procedures are asymptotically optimal (a.o.). Section 4 treats the empirical Bayes linear loss two-action problem and points out its extension to linear loss k-action problems. Section 5 gives some examples in which conditions of the various theorems are satisfied. The final section contains remarks on how one can construct empirical Bayes procedures which are a.o. with rates.

For notational convenience, we let [A] denote the indicator function of A and the arguments of functions will not be exhibited whenever they are clear from the context. Integrals are over $(-\infty,\infty)$ unless otherwise stated. All the limits are as $n \to \infty$. Throughout the index i will range over $1,\ldots,n$.

2. The Empirical Bayes Problem With Censored Data.

Let (λ,X,Y) be a random vector where λ has a prior distribution Λ; given λ, X has density $f(x,\lambda)$ with respect to Lebesgue measure on the real line R, and Y is a random variable with distribution G which is independent of (λ,X). The triplet (λ,X,Y) is unobservable. Instead we observe the pair (δ,z), where $\delta = I[X \leq Y]$ and $Z = \min\{X,Y\}$. This is the situation one generally encounters in analyzing randomly right censored data. Consider the Bayes statistical decision problem for squared error loss estimation of λ using δ and $Z = z$. If $E[\lambda^2] < \infty$, the Bayes decision rule can be shown to be

$$(2.1) \qquad d_\Lambda(\delta,z) = \begin{cases} \dfrac{\int \lambda f(z,\lambda)d\Lambda(\lambda)}{\int f(z,\lambda)d\Lambda(\lambda)} & \text{if } \delta = 1 \\ \\ \dfrac{\int \lambda \overline{F}(z,\lambda)d\Lambda(\lambda)}{\int \overline{F}(z,\lambda)d\Lambda(\lambda)}, & \text{if } \delta = 0 \end{cases}$$

where $\overline{F}(x,\lambda) = \int_x^\infty f(t,\lambda)dt$ with any undefined ratios here and elsewhere taken to be zero. The Bayes estimator d_Λ minimizes the risk among all estimators. Its risk is denoted by R_Λ. If Λ is known, we can use d_Λ and attain R_Λ. But Λ is rarely known even if it assumed to exist.

Suppose now that the above decision problem occurs n+1 times leading to the random vectors $\{\lambda_i,\delta_i,Z_i)\}_{i=1}^{n+1}$, where each triplet has the same probability structure as (λ,δ,Z) given above. The $\{\lambda_i\}_{i=1}^{n+1}$ are unobservable and $\{\delta_i,Z_i)\}_{i=1}^{n+1}$ is the available observable data at the (n+1)st problem. Following Robbins (1955, 1964), we would like to construct empirical Bayes estimators not knowing Λ, which do as well as d_Λ in the (n+1)st problem as the number n of problems increases to ∞. In this empirical Bayes approach, one constructs estimators of the form $d_n(\{(\delta_i,Z_i)\}_{i=1}^{n}; (\delta_{n+1},Z_{n+1}))$ to estimate λ_{n+1}. The risk $R_n(d_n, \Lambda)$ of using d_n to estimate λ_{n+1} is at least as large as $R(\Lambda)$. This leads to the following definition (see Robbins (1964)).

Definition 2.1 $D = \{d_n\}$ is said to be asymptotically optimal (abbreviated by a.o. hereafter) relative to Λ if $R_n(d_n,\Lambda) - R(\lambda) \to 0$.

There has been a great deal of work on empirical Bayes problems in the case when there is no censoring, that is, situations in which censoring distribution G assigns all its mass to $+\infty$. In particular, empirical Bayes procedures have been constructed which satisfy Definition 2.1 with and without rates of convergence to zero of $R_n(d_n,\Lambda) - R(\Lambda)$ with several variations of the component problem involved. Here we consider the empirical Bayes squared error loss problem in the presence of random censoring when $f(x,\lambda)$ has the form either given by Case 1 or Case II noted above. Let Ω be the natural parameter space of the exponential family involved, that is, $\Omega = \{\lambda | \int_a^\infty h(x)\exp(\lambda x)dx < \infty\}$ in Case I. Until otherwise stated, it is assumed that

(A1) Support of Λ is contained in the finite closed interval $[\alpha,\beta]$, where α and β are known. Additional conditions on h will be needed and they will be introduced later. We assume that

(A2) $$\sup\{x:x \text{ is in the support of } G\} > a.$$

The implication of (A2) is that $P(\delta = 1) > 0$. That is, we will observe infinitely many uncensored observations among $\{(\delta_n,Z_n)\}_{n=1}^\infty$. Also, if (A2) is not satisfied, the Bayes estimator (2.1) reduces to $d_\Lambda(0,z) = E[Y]$ and $P(\delta = 0) = 1$. This makes the empirical Bayes problem meaningless.

Let

(2.2) $$\tau = \sup\{x:1 - G(x) < 1\}.$$

Then $\tau > a$, $P(Z_i < \tau) = 1$. We take $\tau = \infty$ if the l.h.s. of (A2) is infinite. If we need to estimate $F_\Lambda(z)$ for all z, we need $\tau = \infty$ (see Gill (1983)).

Since the support of $\Lambda \subseteq [\alpha,\beta]$, the inequality $\alpha \leq d_\Lambda \leq \beta$ leads us to consider empirical Bayes estimates $\{d_n(\{(\delta_i,Z_i)\}_{i=1}^n, (\delta_{n+1},Z_{n+1}))\}_{n=1}^\infty$ taking values in $[\alpha,\beta]^\infty$. Under the assumptions, we have

$$(2.3) \qquad 0 \leq R_n(d_n,\Lambda) - R(\Lambda) = E[(d_n - d_\Lambda(\delta_{n+1},Z_{n+1}))^2].$$

For $\{d_n\}$ to be a.o., by (2.3) it is enough to show that

$$(2.4) \qquad d_n(\{(\delta_i,Z_i)\}_{i=1}^n,x) \xrightarrow{P} d_\Lambda(x)$$

almost everywhere w.r.t. the distribution of (δ,Z), where $\xrightarrow{P}$ denotes convergence in probability as $n \to \infty$.

In Section 3, we provide $\{d_n\}$ satisfying (2.4) and hence a.o. for Case I. In Section 5, we consider some examples to which the results of Section 3 can be applied. Section 4 describes an empirical Bayes linear loss 2-action problem in which we need only assume that $E[|\lambda|] < \infty$. We use without further comment that if the r-th derivative $h^{(r)}$ of h exists and is continuous on (a,∞), then $f_\Lambda^{(r)}$ exists and is continuous on (a,∞), where $f_\Lambda(x) = \int f(x,\lambda)d\Lambda(\lambda)$. See Theorem 2.9 of Lehmann (1959).

3. An Empirical Bayes Estimator and its Asymptotic Optimality.

Throughout we deal with Case I where $f(x,\lambda) = c(\lambda)h(x)\exp(\lambda,x)$. In this case, d_Λ given by (2.1) can be shown to be

$$(3.1) \quad d_\Lambda(\delta,z) + \delta h^{(1)}(z)/h(z) = \frac{\delta f_\Lambda^{(1)}(z)}{f_\Lambda(z)} + (1-\delta)\left\{\frac{f_\Lambda(z)}{\bar{F}_\Lambda(z)} - \frac{\int_z^\infty f_\Lambda d(\log h)}{\bar{F}_\Lambda(z)}\right\},$$

if $h^{(1)}$ exists. By our assumption on Λ, it follows that $\alpha \leq d_\Lambda \leq \beta$. In view of (2.2) and (3.1), we need estimators of f_Λ, $f_\Lambda^{(1)}$, $\bar{F}_\Lambda$, and the integral $\int_z^\infty f_\Lambda d(\log h)$ using the data $\{(\delta_i,Z_i)\}$. These are constructed below. Throughout (δ,z) will denote a generic vector with $\delta = 0$ or 1 and $-\infty < z < \tau$.

Let $N^+(x)$ be the number of $Z_i > x$. Then the product-limit estimator of $\bar{F}_\Lambda$ is defined by

$$\hat{\bar{F}}_\Lambda(x) = \Pi\{N^+(Z_i)/(1 + N^+(Z_i))\}^{[\delta_i, Z_i < x]}. \tag{3.2}$$

By the results of Gill (1983), it follows that

$$\hat{\bar{F}}_\Lambda(z) - \bar{F}_\Lambda(z) \to 0 \text{ almost surely} \tag{3.3}$$

for each $z < \tau$. We now construct estimators for the density and its derivative $f_\Lambda^{(1)}$ based on the observed data $\{(\delta_i, Z_i)\}_{i=1}^n$. There are several estimators for f_Λ based on censored data such as those in Blum and Susarla (1980) or Liu and Van Ryzin (1985). Since Blum and Susarla (1980) use the kernel method of estimation of f_Λ which can easily be extended to estimation of the derivatives of f_Λ, we use their method applied to the kernels introduced in the uncensored problem given by Johns and Van Ryzin (1972).

Let K_0 and K_1 be two real valued functions satisfying

(3.4) K_0, K_1 vanish off $(0,1)$, are bounded on $(0,1)$, have their derivatives bounded, and $\int K_0 = \int u K_1(u)du = 1$ and $\int K_1 = 0$.

We now estimate the derivatives $f_\Lambda^{(\ell)}$, $\ell = 0$ or 1, by

$$\hat{f}_\Lambda^{(\ell)}(x) = -(\varepsilon_n)^{-(\ell+1)} \int K_\ell((r - x)/\varepsilon_n) d\hat{\bar{F}}_\Lambda(r), \tag{3.5}$$

where $\varepsilon_n \downarrow 0$ and $\hat{\bar{F}}_\Lambda$ is defined by (3.2). Using $\hat{f}_\Lambda$ and $\hat{\bar{F}}_\Lambda$, we define our empirical Bayes estimator for use in the (n+1)st problem as

$$d_n(\{(\delta_i, Z_i)\}_{i=1}^n, (\delta_{n+1}, Z_{n+1}))$$

(3.6)
$$= R_{[\alpha,\beta]}\left[\left\{\frac{\hat f_\Lambda^{(1)}(Z_{n+1})}{\hat f_\Lambda(Z_{n+1})} - \frac{h^{(1)}(Z_{n+1})}{h(Z_{n+1})}\right\}\delta_{n+1}\right.$$

$$\left. + \left\{\frac{\hat f_\Lambda(Z_{n+1})}{\hat F_\Lambda(Z_{n+1})} - \frac{\int_{Z_{n+1}}^{Z_{(n)}-\varepsilon_n} \hat f_\Lambda d\log h}{\hat{\bar F}_\Lambda(Z_{n+1})}\right\}(1-\delta_{n+1})\right]$$

where $R_{[\alpha,\beta]}(c) = \alpha[c < \alpha] + c[\alpha \leq c \leq \beta] + \beta[c > \beta]$. If $Z_{n+1} > Z_{(n)} - \varepsilon_n$ in the integral, the integral is taken to be zero. With d_n so defined, we have the following asymptotic result for $\{d_n\}$.

THEOREM 3.1. Let (A1) and (A2) hold; K_0, K_1 satisfy (3.4) and $\varepsilon_n \downarrow 0$ such that $n\varepsilon_n^2 \to \infty$. Let $\tau = \infty$. Let $h^{(2)}$ exist and be continuous. Additionally for $T \in (-\infty,\infty)$, assume

(A3)
$$\overline{\lim} \int_T^\infty \bar F_\Lambda(t) \frac{|h^{(1)}(t)|}{h(t)} \{\int_{-\infty}^{t+\varepsilon_n} \frac{1}{\bar G}\, d(-\frac{1}{\bar F_\Lambda})\}^{(1/2)+\gamma} dt < \infty,$$

for a $0 < 2\gamma < 1$, and for a $\eta > 0$,

(A4)
$$\int_T^\infty \sup\{|f_\Lambda(t + \eta_1) - f_\Lambda(t)|\,|0 < \eta_1 < \nu\}|h^{(1)}(t)|/h(t)dt < \infty.$$

Then, $\{d_n\}$ is a.o. in the sense of Definition 2.1.

REMARK 3.1 The interpretation that $\tau = \infty$ (see 2.2)) is that the supremum of the points in the support of $G = \infty$. This appears to be a reasonable assumption since, for d_n to be a good estimator of d_Λ, we need to estimate f_Λ over the entire line. To do this, as has been observed in almost every paper on random censoring models, we need that the supremum of the support of $1 - \bar F_\Lambda$ (which is infinite in our case) to be at most τ. Hence $\tau = \infty$.

REMARK 3.2 The γ in condition in (A3) could depend on T, but is always less than 1/2.

The proof of the theorem follows from the lemma given below, which is proved under the conditions of Theorem 3.1.

LEMMA 3.1

(i) $\hat{\overline{F}}_\Lambda(z) \xrightarrow{P} \overline{F}_\Lambda(z)$

(ii) $\hat{f}_\Lambda^{(\ell)}(z) \xrightarrow{P} f_\Lambda^{(\ell)}(z)$ for $\ell = 0,1$

(iii) $\int_z^{Z(n)-\varepsilon_n} \hat{f}_\Lambda \, d \log h \xrightarrow{P} \int_z^\infty f_\Lambda \, d \log h$ is $|\int_z^\infty f_\Lambda h^{(1)}/h| < \infty$.

Proof. (i) follows from Theorem 1.1 of Gill (1983) which implies that the random process $\{\sqrt{n}\,(F_\Lambda - \hat{F}_\Lambda)/\, \overline{F}_\Lambda : -\infty < x \leq \tau^*\}$ converges in distribution to a continuous Gaussian process with mean = 0 and covariance function $c(s,t) = c(\min(s,t) = -\int_{-\infty}^{\min(s,t)} (1/\,\overline{G})d(1/\,\overline{F}_\Lambda)$, $s,t \leq \tau^* < \infty$.

To prove (ii), we write $\hat{f}_\Lambda^{(\ell)}(z) - f_\Lambda^{(\ell)}(z)$ as

$$(3.7) \qquad \hat{f}_\Lambda^{(\ell)}(z) - f_\Lambda^{(\ell)}(z) = [\varepsilon_n^{-(\ell+1)} \int K_\ell((t-z)/\varepsilon_n)d(\overline{F}_\Lambda - \hat{\overline{F}}_\Lambda)] +$$

$$[-\varepsilon_n^{-(\ell+1)} \int K_\ell((t-z)/\varepsilon_n)d\overline{F}_\Lambda - f_\Lambda^{(\ell)}(z)]$$

$$= I + II \text{ (say)}.$$

The non-stochastic term II is easy to deal with. Rewrite II as

$$II = \frac{1}{\varepsilon_n^\ell} \int K_\ell(u)[f_\Lambda(z + \varepsilon_n u)]du - f_\Lambda^{(\ell)}(z)$$

For $\ell = 0$, write II as $\int K_\ell(u)[f_\Lambda(z + \varepsilon_n u) - f_\Lambda(z)]du$ which converges to zero since h is assumed to be continuous. For $\ell = 1$, use a first order Taylor expansion to obtain $|II| \leq \sup \{f_\Lambda^{(1)}(z + \varepsilon_n t) - f_\Lambda^{(1)}(z)|: 0 < t < 1\} \times \int u|K_1(u)|du \rightarrow 0$ since $h^{(1)}$ is assumed to be continuous. Therefore,

$$|II| \longrightarrow 0 \text{ for } \ell = 0,1. \tag{3.8}$$

We now deal with I. Since $1 - \overline{F}_\Lambda$, $1 - \hat{\overline{F}}_\Lambda$ are probability measures, and K_ℓ has a bounded derivative, $\varepsilon_n I = \varepsilon_n^{-1} \int (\overline{F}_\Lambda - \hat{\overline{F}}_\Lambda) d\, K_\ell((x - \gamma)\varepsilon_n)$ which, after a change of variables, value can be bounded in absolute value by a constant multiple of $\sup \{|\overline{F}_\Lambda(t) - \hat{\overline{F}}_\Lambda(t)|: z < t < z + \varepsilon_n\}$. This, in turn, is $O_p(1/\sqrt{n})$ by the results of Gill (1983). Since $n\varepsilon_n^2 \rightarrow 0$ by assumption, we thus have $I = o_p(1)$. This together with (3.8) implies (ii).

To prove (iii), recall that by assumption $|\int_z^\infty f_\Lambda h^{(1)}/h| < \infty$ for each z. Since $Z_{(n)} \rightarrow \tau = \infty$ with probability one, (iii) follows if we show that

$$\int_z^{Z_{(n)} - \varepsilon_n} (\hat{f}_\Lambda - f_\Lambda) d(\log h) = o_p(1). \tag{3.9}$$

To obtain (3.9), write $\hat{f}_\Lambda - f_\Lambda = I + II$ as in (3.7). Now by condition (A4) of the theorem, we obtain that $\int_z^\infty II\, d\,(\log h) = 0$. Hence it suffices to show that $\int_z^\infty I\, d(\log h) = o_p(1)$. First, one obtains that

$$I(t) = -\int (\hat{\overline{F}}_\Lambda(t + \varepsilon_n u) - \overline{F}_\Lambda(t + \varepsilon_n u)) dK_0(u). \tag{3.10}$$

Now let $Q^2(t) = (t(1 - t))^{1-2\gamma}$ for $0 \leq t \leq 1$, and $0 < 2\gamma < 1$ with γ as in condition (A3) and $M(t) = c(t)\,(1 + c(t))^{-1}$ with $C(T) = -\int_{-\infty}^T (1/\overline{G}) d(1/\overline{F}_\Lambda)$. Using Q and M, one can bound the integrand of (3.10) by

$$[\sup \{\sqrt{n}\, \frac{|\hat{\overline{F}}_\Lambda - \overline{F}_\Lambda|}{\overline{F}_\Lambda} \frac{(1 - M)}{Q(M)} : -\infty < x \leq Z_{(n)}\}] \times$$

(3.11)

$$\{n^{-1/2}\ \overline{F}_\Lambda(t) \int \frac{Q(M(t + \varepsilon_n u))}{1 - M(t + \varepsilon_n u)}\ |K_0^{(1)}(u)|du\}$$

But the random variable inside the square brackets in (3.11) is $O_p(1)$ by Theorem 3.4 of Gill (1983) since Q is a symmetric function on [0,1], $\int Q^{-2}(t)dt < \infty$, and Q is nondecreasing on [0,½]. Moreover,
as $t \to \infty$, for $0 < u < 1$, $Q(M(t + \varepsilon_n u))/1 - M(t + \varepsilon_n u))$ is bounded by a constant multiple of

$$\{-\int_{-\infty}^{t+\varepsilon_n} \frac{1}{\bar{G}}\, d(1/\bar{F}_\Lambda)\}^{\gamma+1/2} .$$

This bound together with condition (A3) and (3.9) through (3.11) imply that $\sqrt{n}\int_z^\infty I\, d(\log h) = O_p(1)$ showing that $\int_z^\infty I\, d(\log h) = o_p(1)$. This completes the proof of the lemma.

Proof of Theorem 3.1. Theorem 3.1 follows from realizing that both d_Λ and d_n are bounded and that for each (δ,z), $d_n(\{\delta_i,z_i\}_{i=1}^n, (\delta,z)) \xrightarrow{P} d_\Lambda(\delta,z)$.

A result similar to Theorem 3.1 holds also for Case II for which $f(x,\lambda) = c(\lambda)\lambda^x h(x)$. Here, the Bayes estimator is of the form

$$(3.12)\quad d_\Lambda(\delta,z) = \delta\{ \frac{f_\Lambda(z+1)h(z)}{f_\Lambda(z)h(z+1)} \} + (1-\delta)\{\int_z^\infty f_\Lambda(t)(\frac{h(t)}{h(t+1)})dt\} \frac{1}{\bar{F}_\Lambda(z)} .$$

One advantage in this empirical Bayes formulation is that one need not estimate $f_\Lambda^{(1)}$ as in Case 1. This fact and details analogous to those that lead to Theorem 3.1 give the following result concerning the empirical Bayes estimator given below. As in (3.6), define the empirical Bayes estimator as

$$d_n(\{(\delta_i, Z_i)\}_{i=1}^n, (\delta_{n+1}, Z_{n+1}))$$

$$(3.13)\qquad = R_{[\alpha,\beta]}\{\delta_{n+1}(\frac{\hat{f}_\Lambda(Z_{n+1}+1)}{\hat{f}_\Lambda(Z_{n+1})} \frac{h_n(Z_{n+1})}{h(Z_{n+1}+1)})$$

$$+ (1-\delta_{n+1}) (\int_{Z_{n+1}}^{Z_{(n)}^{-\varepsilon_n}} \hat{f}_\Lambda(t)(h(t)/h(t+1))dt)(\hat{\bar{F}}_\Lambda(Z_{n+1}))^{-1}\} .$$

THEOREM 3.2. Let (A1) hold, K_0 satisfies (3.4) and $\varepsilon_n \to 0$ such that $n\varepsilon_n^2 \to \infty$. Let $\tau = \infty$. Let h be continuous. Also, let (A3) and (A4) hold with $h^{(1)}(x)$ there replaced by $h(x + 1)$. Then, $\{d_n\}$, with d_n defined by (3.13), is a.o.

4. Empirical Bayes Linear Loss k-Action Problems.

Here we consider the statistical decision problem involving k-actions as in Johns and Van Ryzin (1972). In the notation of the previous sections, this statistical decision problem when k = 2 can be given as testing the hypothesis $H_1: \lambda \leq c$ against the hypothesis $H_2: \lambda > c$, with c known. The loss is measured by $L(a_1,\lambda) = (\lambda - c)^+$ and $L(a_2,\lambda) = (\lambda - c)^-$, where a_j is the action of deciding in favor of H_j, j = 1,2. In this problem, the Bayes rule is given by a probability distribution γ on $\{a_1, a_2\}$ for each observable data point (δ,Z) with $\gamma(\delta,Z)$ representing the probability of taking action a_1. Explicitly,

$$\gamma(\delta,z) = \delta[\alpha_1(z) \leq 0] + (1 - \delta)[\alpha_2(z) \leq 0] \tag{4.1}$$

where

$$\alpha_1(z) = \int\lambda\, f(z,\lambda)d\Lambda(\lambda) - c\, f_\Lambda(z) \quad \text{and}$$

(4.2)

$$\alpha_2(z) = \int\lambda\, \overline{F}(z,\lambda)d\Lambda(\lambda) - c\, \overline{F}_\Lambda(z).$$

The corresponding minimum Bayes risk is given by

$$R(\Lambda) = \int [\alpha_1(z) \leq 0]\overline{G}(z)\, \alpha_1(z)dz + \int [\alpha_2(z) \leq 0]\, \alpha_2(z)dG(z) + E[(c - \lambda)^+]. \tag{4.3}$$

If $E\{|\lambda|\} < \infty$ (assumed throughout this section), then $R(\Lambda)$ is finite. In view of (4.1) and (4.2), the empirical Bayes procedure to be used in the (n + 1) set problem is given by

$$\gamma_n(\delta_{n+1}, Z_{n+1}) = \delta_{n+1}[\hat{\alpha}_1(Z_{n+1}) \leq 0] + (1 - \delta_{n+1})\ [\hat{\alpha}_2(Z_{n+1}) \leq 0] \tag{4.4}$$

where $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are estimators of α_1 and α_2 based on (δ_i, Z_i), $1 \leq i \leq n$. Using the estimators given in the previous section for Case I, we choose $\hat{\alpha}_1$ and $\hat{\alpha}_2$ to be

$$\hat{\alpha}_1(z) = \hat{f}_\Lambda^{(1)}(z) - \left(\frac{h^{(1)}(z)}{h(z)} \right) \hat{f}_\Lambda(z) - c\hat{f}_\Lambda(z) \text{ and}$$

(4.5)

$$\hat{\alpha}_2(z) = \hat{f}_\Lambda(z) - \int_z^{Z(n)^{-\varepsilon_n}} \hat{f}_\Lambda \; d \log h - c\hat{\bar{F}}_\Lambda(z),$$

where $\hat{f}_\Lambda$, $\hat{f}_\Lambda^{(1)}$, $\hat{\bar{F}}_\Lambda$ and ε_n are as defined in Theorem 3.1. The asymptotic optimality of the procedure defined by (4.4) and (4.5) follows from Theorem 3.1 and the following inequality which is a generalization of the fundamental inequality in Johns and Van Ryzin (1972).

$$0 \leq R(\gamma_n, \Lambda) - R(\Lambda)$$

$$\leq \int P(|\hat{\alpha}_1(z) - \alpha_1(z)| > |\alpha_1(z)|)|\bar{G}(z)|\alpha_1(z)|dz \tag{4.6}$$

$$+ \int P(|\hat{\alpha}_2(z) - \alpha_2(z)| > |\alpha_2(z)|)|\alpha_2(z)|dG(z).$$

As a consequence of this inequality, the dominated convergence theorem, and Theorem 3.1, we have

THEOREM 4.1. If

$$E[|\lambda|] < \infty \tag{A5}$$

and the other conditions in Theorem 3.1 (excluding A1) are satisfied, then the empirical Bayes testing procedure defined by (4.4) and (4.5) is a.o.

A Theorem 4.2, say, similar to the Theorem 4.1 can be stated for Case II with the obvious modifications in the definitions of $\hat{\alpha}_1$ and $\hat{\alpha}_2$. Now $\hat{\alpha}_1$ and $\hat{\alpha}_2$ will correspond to those in Theorem 3.2; and as was noted earlier, we need not estimate $f_\Lambda^{(1)}$.

An extension of the above empirical Bayes results in the 2-action case to k-action cases is possible in the following way. Let the action a_j correspond to deciding that the hypothesis $H_j : [\lambda_{j-1} \leq \lambda \leq \lambda_j]$ is true for $j=1,\ldots,k+1$ in the statistical decision problem where $\lambda_0 < \lambda_1 < \ldots < \lambda_k$. As a loss function, we take $L(\lambda,a_j)$ to be such that for $j = 1,\ldots,k$, $L(\lambda,a_j) - L(\lambda,a_{j-1}) = (\lambda_{j-1} - \lambda)$ and $L(\lambda,a_1) = 0$ if $\lambda \leq \lambda_0$ and $= \sum_{\ell=1}^{j-1}(\lambda - \lambda_\ell)$ if $\lambda_{j-1} < \lambda \leq \lambda_j$. Here the Bayes decision rule is given by the probability vector $\underset{\sim}{\gamma}(\delta,Z) = (\gamma_1(\delta,Z),\ldots,\gamma_k(\delta,Z))$, where $\gamma_j(\delta,Z)$ represents the probability of deciding in favor of action a_j given that the observed data point is (δ,Z). It can then be shown that for $1 \leq j < k$,

$$
\begin{aligned}
\gamma_j(\delta,Z) = {} & \delta[\lambda_{j-1}f_\Lambda^{(z)} \leq \int\lambda f(z,\lambda)d\Lambda(\lambda) < \lambda_j f_\Lambda(z)] \\
& + (1-\delta)[\lambda_{j-1}\overline{F}_\Lambda(z) \leq \int\lambda\overline{F}(z,\lambda)d\Lambda(\lambda) < \lambda_j\overline{F}_\Lambda(z)]
\end{aligned}
\tag{4.7}
$$

To get a natural empirical Bayes procedure for use in the $(n+1)$th problem, we replace in (4.7) the argument (δ,Z) by (δ_{n+1},Z_{n+1}) and replace f_Λ, $\int\lambda f(x,\lambda)d\Lambda(\lambda)$, $\overline{F}_\Lambda$ and $\int\lambda\overline{F}(x,\lambda)d\Lambda(\lambda)$ by their estimators given in Section 3. The resulting empirical Bayes testing procedure can be shown to be a.o. if $E[|\lambda|] < \infty$ and the other conditions of Theorems 4.1 or 4.2 are satisfied.

The extension of the above results to any multiple decision problem can be carried out as in Section 2 of Susarla and Van Ryzin (1977).

5. Examples: In this section, we present several examples in which the conditions of Theorems 3.1 or 3.2 are satisfied. Obviously, the same examples can be used to illustrate Theorems 4.1 and 4.2.

Example 5a (Case I). Let $\sqrt{2\pi}f(x,\lambda) = \exp(-(x-\lambda)^2/2)$. Then $a = -\infty$, $h(x) = \exp(-x^2/2)$, $c(\lambda) = (2\pi)^{-1/2}\exp(-\lambda^2/2)$, $\Omega = (-\infty,\infty)$, and $h^{(1)}(x)/h(x) = -x$. By an obvious weakening, (A3) is implied by $\overline{\lim} \int_T^\infty F_\Lambda(t)(\overline{G}(t + \varepsilon_n)\overline{F}_\Lambda(t + \varepsilon_n))^{-(1+2\gamma)/2}dt < \infty$ with $\overline{\lim}$ taken as $\eta \to \infty$. This condition is satisfied if for $\eta > 2$, $\overline{\lim}\ \{\exp(-t^2/\eta)/\overline{G}(t)\} < \infty$ with $\overline{\lim}$ taken as $t \to \infty$. For such an η, $2\gamma < (\eta - 2)/(\eta + 2)$. This condition on $\overline{G}$ amounts to assuming that $\overline{G}$ has a heavier right tail than that of the standard normal distribution. Since Λ has compact support, it can also be verified that (A4) holds here.

Example 5b (Case I). $f(x,\lambda) = \lambda \exp(-\lambda x)$ if $x > a = 0$. $h \equiv 1$. $c(\lambda) = -\lambda$, and $\Omega = (-\infty,0)$. Here we estimate $-\lambda$ instead of λ. Since $h^{(1)}/h \equiv 0$ on $(0,\infty)$, (A3) and (A4) are automatically satisfied.

Example 5c (Case I). $f(x,\lambda) = \lambda^\alpha x^{\alpha-1} \exp(-\lambda x)$ with $x > a = 0$ and with $\alpha > 1$ a known positive constant. $h^{(1)}(x)/h(x) = (\alpha - 1)/x$. As in example (5a), condition (A4) is implied by the condition $\overline{\lim}\ \{1/(\exp(\eta t)\ \overline{G}(t)\} < \infty$ for a η in $(0, \alpha)$. The interpretation of this condition is that $\overline{G}$ has a heavier tail than $\exp(-\alpha t)$ where $\alpha(> 0)$ is a lower bound on the support of prior Λ.

Example 5d (Case II). $\sqrt{2\pi}\ f(x,\theta) = \exp(-(x - \theta)^2/2)$. Now $\lambda = \exp(\theta)$, $c(\lambda) = (2\pi)^{-1/2}\exp(-(\ln \lambda)^2/2)$, $h(x) = \exp(-x^2/2)$. Then $f(x,\lambda) = \lambda^x c(\lambda)h(x)$. The same condition as derived as Example 5a will suffice here. Examples 5b and 5c can similarly be put in the form for application of Theorem 3.2.

6. Concluding Remarks: In this paper we have introduced the empirical Bayes problem with censored data and obtained empirical Bayes estimators or multiple decision procedures which are a.o. In the latter case, we need only assume that Λ has first absolute moment finite. In all the problems considered here, we can obtain rates of convergence for $R_n(t_n,\Lambda) - R(\Lambda) \to 0$. However, the techniques involved in obtaining such rate results are markedly different from the weak convergence approach of Gill (1983) used here. We believe that the

approach of Gill will not give any moment convergence results for the product-limit process. Consequently, Gill's results can not be used to obtain the needed (for empirical Bayes results rates) moment convergence results for the product-limit process.

REFERENCES

Blum, J. and Susarla, V. (1980). Maximal deviation theory of density and failure rate function estimates based on censored data. Proc. of the Fifth International Symp. on Multivariate Analysis - V, 213-222.

Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. Ann. Statist. **1** 209-230.

Gill, R.D. (1983). Large sample behavior of the product-limit estimator on the whole line. Ann. Statist. **11** 49-58.

Johns, M.V. Jr. and Van Ryzin, J. (1972). Convergence rates for empirical Bayes two-action problems II. Continuous case. Ann. Math. Statist. **43** 934-947.

Liu, R. and Van Ryzin, J. (1985). A histogram estimator of the hazard rate with censored data. Ann. Statist. **13** 592-605.

O'Bryan T. and Susarla, V. (1977). Empirical Bayes estimation with nonidentical components, continuous case. Aus. J. Statist. **19** 105-115.

Robbins, H. (1955). An empirical Bayes approach to statistics. Proc. Third Berkeley Symp. Math. Statist. Prob. 157-163. Univ. of California Press, Berkeley.

Robbins, H. (1963). The empirical Bayes approach to testing statistical hypotheses. Rev. Inst. Internat. Statist. **31** 195-208.

Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. Ann. Math. Statist. **35** 1-20.

Susarla, V. (1982). Empirical Bayes theory. Encyclopedia of Statistical Sciences. 2, 490-503.

Susarla, V. and Van Ryzin, J. (1977). On the empirical Bayes approach to multiple decision problems. Ann. Statist. 5, 172-181.

Susarla, V. and Van Ryzin, J. (1978). Empirical Bayes estimation of a distribution (survival) function from right censored observations. *Ann. Stastist.* *6*, 740-754.