The Annals of Statistics
1977, Vol. 5, No. 1, 164–171

COMPLETE CLASS THEOREMS FOR THE SIMPLEST EMPIRICAL BAYES DECISION PROBLEMS

BY TOM SNIJDERS

University of Groningen

For the problem of empirical Bayes classification into two known probability distributions on a finite outcome space, an essentially complete class of procedures is determined. This class is proven to be minimal essentially complete if there are only two possible outcomes.

1. Introduction. Consider a population of people who are Russian with unknown probability Θ. The probability that a Russian from this population likes caviar is known to be p_1; for a non-Russian person from this population, this probability is known to be $p_0 \neq p_1$. For a sample of n people from this population it has been observed only whether they like caviar. A statement must be made whether the nth person is a Russian or not. This amounts to the empirical Bayes classification into either of two completely known populations after observing the outcome of a Bernoulli random variable. For this decision problem, the simplest possible empirical Bayes problem, a minimal essentially complete class will be obtained (Corollary 5, characterized by (6)). More generally, if the random variable has a finite distribution, an essentially complete class is obtained (Theorem 3). The proof uses a monotone likelihood-ratio property, generalized to a partially ordered outcome space.

The empirical Bayes approach was introduced by Robbins (1955); a good introduction is also supplied by Robbins (1964). Ordinarily, it is treated asymptotically. This paper considers an empirical Bayes problem in the setting of decision theory with fixed sample size. Then it becomes a prediction problem, where the prior distribution has the role of an unknown parameter.

In the literature about empirical Bayes procedures and the closely related compound decision problem (see Robbins (1951) or Samuel (1967) for a general discussion about the latter), various remarks and results have appeared concerning the (in) admissibility of proposed procedures in the finite sample case. Research into this was done, e.g., by Maritz (see Maritz (1968) and references cited there), Meeden (1972), van Houwelingen (1973), Copas (1974) and Inglis (1976). Meeden and van Houwelingen improved on empirical Bayes procedures which had appeared in the literature. Meeden constructed some admissible empirical Bayes procedures. The decision problems treated by these authors are much more complex than the problem dealt with in this paper.

2. Formulation and preliminaries. The random variable X has the outcome

Received February 1975; revised May 1976.
AMS 1970 *subject classifications.* Primary 62C07, 62C10; Secondary 62F05.
Key words and phrases. Empirical Bayes classification, complete class, monotone procedures.

space $\{0, 1, \cdots, k-1\}$ and either of the known probability distributions P_t, $t \in \{0, 1\}$. These distributions satisfy $P_t\{x\} > 0$ for all x and t, and have increasing likelihood ratio; that is,

$$0 < r_0 < r_1 < \cdots < r_{k-1} < \infty \qquad \text{where} \quad r_x = P_1\{x\}/P_0\{x\} .$$

The parameter t is the outcome of a random variable T with prior distribution $Q_\Theta\{T = 1\} = \Theta$, $Q_\Theta\{T = 0\} = 1 - \Theta$ for a $\Theta \in [0, 1]$. In the empirical Bayes set-up this experiment is repeated independently with the same unknown Θ. Thus a sequence

$$(X_1, T_1), (X_2, T_2), \cdots, (X_n, T_n)$$

is realized, consisting of i.i.d. pairs of random variables with the described distribution. The marginal distribution of $X_1, \cdots, X_n$ is that of a sequence of i.i.d. variables with distribution P_Θ where

$$p_\Theta(x) = P_\Theta\{X_i = x\} = \Theta P_1\{x\} + (1 - \Theta)P_0\{x\} . \tag{1}$$

Only X_1 through X_n are observable. Either of the statements "$t_n = 0$" or "$t_n = 1$" is to be made. If the statement made is untrue, a loss of β or α, respectively is incurred. Both α and β are positive numbers. A true statement causes a zero loss. Represent randomized actions by the probability $a \in [0, 1]$ of stating "$t_n = 1$." The loss function is then $L(0, a) = \alpha a$, $L(1, a) = \beta(1 - a)$. The decision procedure δ: $\{0, 1, \cdots, k-1\}^n \to [0, 1]$ has the risk

$$R(\Theta, \delta) = E_\Theta L(T_n, \delta(X_1, \cdots, X_n)) .$$

It will be shown that as for the outcome of $(X_1, \cdots, X_{n-1})$ one only needs to be concerned about the number of times that the outcomes 0 through $k - 1$ occur in the sequence $(x_1, \cdots, x_{n-1})$. Let

$$s^{(h)}(\cdot) = (s_0^{(h)}(\cdot), \cdots, s_{k-1}^{(h)}(\cdot))$$

where

$$s_j^{(h)}(x_1, \cdots, x_h) = \text{number of components of} \quad (x_1, \cdots, x_h) \quad \text{equal to} \quad j .$$

Let $S^{(h)} = s^{(h)}(X_1, \cdots, X_h)$, and denote its outcome space by $\mathscr{S}^{(h)}$. As the loss is a function of the unobservable random variable T_n, statements about sufficiency must be made with care. The distribution of $(X_1, \cdots, X_n)$ conditional on $(S^{(n-1)}, X_n, T_n)$ does not depend on T_n or Θ. From this may be deduced easily that the rule $\delta_0 = E_\Theta\{\delta(X_1, \cdots, X_n) \mid S^{(n-1)}, X_n, T_n\}$ does not involve T_n or Θ and has the same risk function as δ. Therefore the class D of procedures depending on $(x_1, \cdots, x_n)$ through $(s^{(n-1)}, x_n)$ is essentially complete. Only the class D will be considered. A complete subclass of D is an essentially complete class in the original problem.

The statistics $(S^{(n-1)}, X_n)$ and $(S^{(n)}, X_n)$ generate the same partition of the outcome space $\{0, \cdots, k-1\}^n$. The latter form of the sufficient statistic has the technical disadvantage that its two components are stochastically dependent. It is however more pleasant conceptually, as $S^{(n)}$ contains all available information concerning Θ, and X_n all information concerning t_n. Therefore, after this reduction by sufficiency the outcome space

$$\mathscr{U} = \{(s, x) \in \mathscr{S}^{(n)} \times \{0, \cdots, k-1\} \mid s_x \geqq 1\}$$

is considered, and procedures $\delta : \mathscr{U} \to [0, 1]$.

The probability distribution of $(S^{(n)}, X_n, T_n)$ is given by

$$P_\Theta\{S^{(n)} = s, X_n = x\} = h(s, x)p_\Theta{}^s \qquad \text{for} \quad (s, x) \in \mathscr{U}$$
$$P_\Theta\{T_n = t \mid S^{(n)} = s, X_n = x\} = p_t(x)/p_\Theta(x)$$

where

$$p_\Theta{}^s = \prod_{j=1}^k [p_\Theta(j)]^{s_j}$$

$h(s, x)$ = number of outcomes $(x_1, \cdots, x_n)$ corresponding to outcome (s, x). The risk of a procedure $\delta : \mathscr{U} \to [0, 1]$ is

$$\begin{aligned} R(\Theta, \delta) &= E_\Theta L(T_n, \delta(S^{(n)}, X_n)) \\ &= \alpha(1-\Theta) \textstyle\sum_{(s,x)\in\mathscr{U}} \delta(s, x)h(s, x)p_0(x)p_\Theta{}^s/p_\Theta(x) \\ &\quad + \beta\Theta \textstyle\sum_{(s,x)\in\mathscr{U}} (1 - \delta(s, x))h(s, x)p_1(x)p_\Theta{}^s/p_\Theta(x) \\ &= \beta\Theta + (\alpha + \beta) \textstyle\sum_{(s,x)\in\mathscr{U}} \delta(s, x)h(s, x)p_\Theta{}^s\{f_\Theta(x) - c\} \end{aligned}$$

where

$$f_\Theta(x) = (1-\Theta)p_0(x)/p_\Theta(x)\,, \qquad c = \beta/(\alpha+\beta)\,.$$

The function $p_\Theta{}^s\{f_\Theta(x) - c\}$ will be denoted also by $l_\Theta(s, x)$. The procedure δ_1 dominates δ_2 if $R(\Theta, \delta_1) \leqq R(\Theta, \delta_2)$ for all $\Theta \in [0, 1]$, which can be expressed by

$$\textstyle\sum_{u\in\mathscr{U}} \{\delta_1(u) - \delta_2(u)\}h(u)l_\Theta(u) \leqq 0 \qquad \text{for all} \quad \Theta \in [0, 1]\,. \tag{2}$$

When a complete subclass of D is to be obtained, it is handy to know that a minimal complete subclass exists. Then, any subclass of D which is left over after throwing out some inadmissible procedures is complete. From Theorem 2.22 of Wald (1950) can be concluded that a minimal complete subclass exists in this decision problem. A complete class will be obtained with the aid of properties resembling monotone likelihood ratio.

3. Ordering the outcome space. The outcome space $\mathscr{U}$ is to be ordered in such a way that a larger outcome suggests larger values of Θ and of t_n. The ordering (2) of procedures leads towards the following definition.

DEFINITION. For $u, v \in \mathscr{U}$, $u \preceq v$ if and only if a constant $\mu > 0$ exist such that

$$-\mu l_\Theta(u) + l_\Theta(v) \leqq 0 \qquad \text{for all} \quad \Theta \in [0, 1] \tag{3}$$

and $u \prec v$ if and only if $u \preceq v$ and not $v \preceq u$.

For each pair $u, v \in \mathscr{U}$ with $u \neq v$, $pl_\Theta(u) + ql_\Theta(v) \equiv 0$ implies $p = q = 0$. Hence, $u \prec v$ is equivalent to $u \preceq v$ and $u \neq v$. It can be verified easily that $\preceq$ is a partial ordering in $\mathscr{U}$.

PROPOSITION 1. *If* $x \leqq y$ *and* $\eta(h) = \sum_{j=h}^{k-1} (s_j - t_j) \leqq 0$ *for* $h = 1, \cdots, k-1$ *then* $(s, x) \preceq (t, y)$.

PROOF. Take $(s, x), (t, y) \in \mathscr{U}$ for which these conditions hold. A $\mu > 0$ is to be found so that (3) holds, which means

$$(4) \qquad -\mu p_\Theta{}^s\{f_\Theta(x) - c\} + p_\Theta{}^t\{f_\Theta(y) - c\} \leqq 0 \qquad \text{for all} \quad \Theta \in [0, 1] .$$

Divide by the positive function $p_\Theta{}^s$. Since

$$p_\Theta{}^{t-s} = p_\Theta{}^t/p_\Theta{}^s = p_0{}^{t-s} \prod_{h=1}^{k-1} \{(\Theta r_h + 1 - \Theta)/(\Theta r_{h-1} + 1 - \Theta)\}^{-\eta(h)}$$

is a product of positive nondecreasing functions, $p_\Theta{}^{t-s}$ is nondecreasing. The functions $f_\Theta(x)$ and $f_\Theta(y)$ are decreasing in Θ, and equal to c in $\Theta(x)$ and $\Theta(y)$, respectively, with $0 < \Theta(y) \leqq \Theta(x) < 1$. A positive μ exists with

$$p_{\Theta(y)}^{t-s} \leqq \mu \leqq p_{\Theta(x)}^{t-s}$$

which now will be shown to satisfy (4). For $\Theta(y) \leqq \Theta \leqq \Theta(x)$, both terms are nonpositive so that the inequality holds. For $0 \leqq \Theta \leqq \Theta(y)$, from $0 < f_\Theta(y) - c \leqq f_\Theta(x) - c$ and $p_\Theta{}^{t-s} \leqq p_{\Theta(y)}^{t-s}$ can be concluded

$$\begin{aligned} -\mu p_\Theta{}^s\{f_\Theta(x) - c\} + p_\Theta{}^t\{f_\Theta(y) - c\} &\leqq p_\Theta{}^s(-\mu + p_\Theta{}^{t-s})\{f_\Theta(x) - c\} \\ &\leqq p_\Theta{}^s(-\mu + p_{\Theta(y)}^{t-s})\{f_\Theta(x) - c\} \leqq 0 . \end{aligned}$$

For $\Theta(x) < \Theta \leqq 1$, the situation is symmetric. □

As is clear from the proof, the condition "$x \leqq y$ and $p_\Theta{}^{t-s}$ is nondecreasing in Θ" is also sufficient for $(s, x) \preceq (t, y)$. For $k \geqq 3$ this condition is strictly weaker than the condition of Proposition 1. For $k = 2$ these conditions are equivalent, which is seen from the following.

PROPOSITION 2. *Consider the case* $k = 2$, *and represent outcome* $s = (n - s_1, s_1)$ *of* $S^{(n)}$ *by* $s_1 \in \{0, \cdots, n\}$. *Then* $(s, x) \preceq (t, y)$ *if and only if* $s_1 \leqq t_1$ *and* $x \leqq y$.

PROOF. The if-statement follows from Proposition 1. If not $s_1 \leqq t_1$, then $s_1 > t_1$ and according to the proof of Proposition 1, $p_\Theta{}^{t-s}$ is strictly decreasing. From $-\mu l_\Theta(s, x) + l_\Theta(t, y) \leqq 0$ for $\Theta = 0, 1$ follows $p_0{}^{t-s} \leqq \mu \leqq p_1{}^{t-s}$ which is impossible. If not $x \leqq y$, then $x > y$. A $\Theta(x) \in (0, 1)$ exists with $f_{\Theta(x)}(x) = c$ so that $f_{\Theta(x)}(y) > c$. For any $\mu > 0$ holds $-\mu l_{\Theta(x)}(s, x) + l_{\Theta(x)}(t, y) = l_{\Theta(x)}(t, y) > 0$. This completes the proof of the only if-statement. □

4. Complete class theorems.

DEFINITION. A procedure $\delta \in D$ is called monotone if

$$\delta(u) > 0 , \quad u \prec v \quad \text{implies} \quad \delta(v) = 1 .$$

The class of monotone procedures in D is denoted by D_m.

THEOREM 3. D_m *is a complete subclass of* D, *and hence an essentially complete class in the original problem.*

PROOF. If $\delta \in D \sim D_m$, there exist $u, v \in \mathscr{U}$ with $u \prec v$, $\delta(u) > 0$ and $\delta(v) < 1$. A $\mu > 0$ exists with $-\mu l_\Theta(u) + l_\Theta(v) \leqq 0$ with strict inequality for some Θ. Define δ_ε for $\varepsilon > 0$ by

$$\begin{aligned} \delta_\varepsilon(u) &= \delta(u) - \varepsilon\mu h(v) \\ \delta_\varepsilon(v) &= \delta(v) + \varepsilon h(u) \\ \delta_\varepsilon(w) &= \delta(w) \quad w \neq u, v\,. \end{aligned}$$

For small enough ε, $\delta_\varepsilon \in D$. Equation (2) shows that such a δ_ε strictly dominates δ, so that δ is inadmissible. Since a minimal complete subclass of D exists and $D \sim D_m$ contains only inadmissible procedures, D_m is complete. □

In the following, a totally ordered subset of a partially ordered set will be called a chain.

THEOREM 4. *If $\mathscr{U}$ can be partitioned into two chains, D_m is the minimal complete subclass of D, and is minimal essentially complete in the original problem.*

PROOF. According to Theorem 3, D_m is complete. Showing that D_m contains only admissible elements suffices to prove its minimal completeness. Argue by contradiction, and suppose $\delta \in D_m$ is inadmissible. As D_m is complete, a dominating $\delta' \neq \delta$ can be found in D_m. Defining $\xi(u) = (\delta'(u) - \delta(u))h(u)$ and applying equation (2) gives

$$\sum_{u \in \mathscr{U}} \xi(u) l_\Theta(u) \leqq 0 \qquad \text{for all} \quad \Theta \in [0, 1]\,. \tag{5}$$

Since $\delta \neq \delta'$ and $l_1(u) < 0 < l_0(u)$ for all u, this implies that u' and u'' exist with $\xi(u') < 0 < \xi(u'')$. Let $\mathscr{U} = \mathscr{U}_1 \cup \mathscr{U}_2$, where $\mathscr{U}_1$ and $\mathscr{U}_2$ are disjoint chains. Since δ and δ' are monotone and $\mathscr{U}_1$ is a chain, either $\xi(u) \leqq 0$ holds for all $u \in \mathscr{U}_1$, or $\xi(u) \geqq 0$; similarly for $\mathscr{U}_2$. It may be supposed that $\xi(u) \leqq 0$ for $u \in \mathscr{U}_1$ and $\xi(u) \geqq 0$ for $u \in \mathscr{U}_2$. Let $u_1 = \min\{u \in \mathscr{U}_1 \mid \delta(u) > 0\}$ and $u_2 = \max\{u \in \mathscr{U}_2 \mid \delta(u) < 1\}$. Then $\delta(u_1) > 0$ and $\delta(u_2) < 1$. Since δ is monotone $u_1 \not\preccurlyeq u_2$. Also $\xi(u) = 0$ for $u \in \mathscr{U}_1$, $u \prec u_1$ and for $u \in \mathscr{U}_2$, $u \succ u_2$. Equation (5) gives

$$\begin{aligned} 0 &\geqq \sum_{u \in \mathscr{U}_1; u \succcurlyeq u_1} \xi(u) l_\Theta(u) + \sum_{u \in \mathscr{U}_2; u \preccurlyeq u_2} \xi(u) l_\Theta(u) \\ &\geqq \sum_{u \in \mathscr{U}_1; u \succcurlyeq u_1} \xi(u)\mu(u) l_\Theta(u_1) + \sum_{u \in \mathscr{U}_2; u \preccurlyeq u_2} \xi(u)\mu^{-1}(u) l_\Theta(u_2) \end{aligned}$$

where the last step uses inequalities of the form (3) with $\mu(u) > 0$ for all u. Since

$$\sum_{u \in \mathscr{U}_1; u \succcurlyeq u_1} \xi(u)\mu(u) < 0 < \sum_{u \in \mathscr{U}_2; u \preccurlyeq u_2} \xi(u)\mu^{-1}(u)\,,$$

this implies that $u_1 \preceq u_2$; the contradiction is established. This argument shows also that D_m does not contain two different procedures with the same risk function. Therefore, it is minimal essentially complete in the original problem. □

(A suggestion made by a referee and by Dr. J. C. van Houwelingen made it possible to shorten this proof.)

COROLLARY 5. *For $k = 2$, where a Bernoulli experiment is performed, D_m is a minimal essentially complete class.*

PROOF. Proposition 2 shows that here, $\mathscr{U}$ can be partitioned into the chains $\{(s, x) \mid s = 0\}$ and $\{(s, x) \mid s = 1\}$. □

In the case of a Bernoulli experiment, D_m can be characterized in the following way. As in Proposition 2, represent outcome $u = ((n - s, s), x) \in \mathscr{U}$ by (s, x) with $x = 0, 1$ and $0 \leqq s - x \leqq n - 1$. Proposition 2 shows that D_m consists of all procedures δ of the form

$$\delta(s, x) = 0 \quad s < \sigma_x$$
$$(6) \qquad = \gamma_x \quad s = \sigma_x$$
$$= 1 \quad s > \sigma_x$$

for numbers $\gamma_0, \gamma_1 \in [0, 1]$, σ_0 and σ_1 with $0 \leqq \sigma_x - x \leqq n - 1$, for which either $\sigma_0 \geqq \sigma_1 + 1$ or $\sigma_1 = 0$, $\gamma_1 = 1$ (which means $\delta(s, 1) \equiv 1$) or $\sigma_0 = n$, $\gamma_0 = 0$ (which means $\delta(s, 0) \equiv 0$). For example, nonrandomized procedures which are monotone functions only of s, or only of x, are admissible; but nontrivial procedures that take only the number of successes in the first $n - 1$ trials, that is $s - x$, into account are inadmissible.

If $k \geqq 3$ and $n \geqq 3$, $\mathscr{U}$ contains at least three pairwise incomparable outcomes and hence cannot be split up into two chains. When the minimal number of chains into which $\mathscr{U}$ can be split up is $l \geqq 3$, "relations" in $\mathscr{U}$ are to be considered involving l elements, because D_m will in general contain inadmissible elements. As an example of an inadmissible procedure in D_m, consider the case $k = n = 3$. By the same reasoning as in Proposition 2 can be proved that the outcomes $u = ((1, 0, 2), 1)$, $v = ((1, 1, 1), 2)$ and $w = ((2, 0, 1), 3)$ are pairwise incomparable. When $c = \frac{1}{2}$ and $r_1 - 2r_2 + r_3 \leqq 0$ some calculations show that $-l_\Theta(u) + 2l_\Theta(v) - l_\Theta(w) \leqq 0$ for all $\Theta \in [0, 1]$. Hence any δ with $\delta(u) > 0$, $\delta(v) < 1$, $\delta(w) > 0$ is inadmissible; and D_m does contain such elements δ.

5. Remarks.

5.1. *Minimax procedures*. In the case without empirical Bayes information about Θ, we might say $n = 1$, an equalizer minimax rule δ exists. If we consider this rule as a procedure in the empirical Bayes case where $n \geqq 2$, depending on $(s^{(n)}, x)$ only through x, it still is a minimax procedure. However, δ is admissible if and only if it is nonrandomized. If δ is randomized, say $\delta(x_0) \in (0, 1)$, an admissible minimax procedure $\delta_0 \in D_m$ exists for which $\delta_0(s^{(n)}, x_0)$ is "monotone in $s^{(n)}$". If $\{(s, x_0) \in \mathscr{U}\}$ is a chain, in particular if $k = 2$, the admissible minimax procedure is unique.

5.2. *The classical empirical Bayes procedure*. Consider the case $k = 2$ and now call $P_0\{1\} = p$, $P_1\{1\} = q$ with $p < q$, $S^{(n)} = (n - S_n, S_n)$. In the classical empirical Bayes approach, Θ will be estimated by (a truncated version of) $\hat{\Theta}_n = (S_n/n - p)/(q - p)$ and the Bayes procedure with respect to the outcome of $\hat{\Theta}_n$ will be used. This gives the rule δ which as a function of (S_n, X_n) is given by

$$\delta(s, x) = 0 \quad s < n\tau_x$$
$$= 1 \quad s > n\tau_x$$

where

$$\begin{aligned} (7) \qquad \tau_0 &= \{\alpha q(1-p) + \beta p(1-q)\}/\{\alpha(1-p) + \beta(1-q)\} \\ \tau_1 &= (\alpha + \beta)pq/(\alpha p + \beta q) . \end{aligned}$$

Since $\tau_0 > \tau_1$, this is a monotone and therefore an admissible procedure. However, sometimes Θ is estimated by $\hat{\Theta}_{n-1}$ instead of $\hat{\Theta}_n$ in order to avoid stochastic dependence between $\hat{\Theta}$ and X_n.

In that case the critical points in terms of S_n become $(n-1)\tau_0$ and $(n-1)\tau_1 + 1$. If an integer s' exists with $(n-1)\tau_0 < s' < (n-1)\tau_1 + 1$ (at most one such s' exists), the classical empirical Bayes procedure δ has $\delta(s, 0) = 1$ for $s \geqq s'$, $\delta(s, 1) = 1$ for $s > s'$ and $\delta(s, x) = 0$ elsewhere. This δ is inadmissible. A referee suggested exhibiting a dominating procedure. In order to find this, note that $(s', 0) \prec (s', 1)$ so that equation (3) holds for this pair of outcomes with a certain μ. The proof of Proposition 1 shows that one must take $\mu = 1$. The proof of Theorem 3 gives a dominating procedure δ_ε. Using $h(s', 0)/h(s', 1) = (n - s')/s'$ it is found that taking ε as large as possible gives

$$\begin{aligned} &\delta_\varepsilon(s, x) = 0 \quad \text{for} \quad s < s' , \qquad \delta_\varepsilon(s, x) = 1 \quad \text{for} \quad s > s' \\ &\text{if} \quad s' \geqq n/2 : \quad \delta_\varepsilon(s', 0) = 1 - (n - s')/s' \\ &\qquad\qquad\qquad\quad\ \delta_\varepsilon(s', 1) = 1 \\ &\text{if} \quad s' \leqq n/2 : \quad \delta_\varepsilon(s', 0) = 0 \\ &\qquad\qquad\qquad\quad\ \delta_\varepsilon(s', 1) = s'/(n - s') . \end{aligned}$$

This δ_ε dominates δ, and it is admissible because it is monotone. If an integer s' exists with $(n-1)\tau_0 \leqq s' \leqq (n-1)\tau_1 + 1$ a similar remark can be made, taking account of the possibility of a randomized δ. If no such s' exists, estimating Θ by $\hat{\Theta}_{n-1}$ yields an admissible procedure.

5.3. *Asymptotic optimality.* As was noted in the introduction, empirical Bayes problems are commonly treated as asymptotic problems. A sequence $\{\delta_n\}$ of procedures is called asymptotically optimal with respect to the prior distribution Q, if its Bayes risk with respect to Q approaches the minimum Bayes risk for Q, as $n \to \infty$. In the case treated in this paper with $k = 2$, a sequence of procedures of the form (6) with critical points $\sigma_0(n)$ and $\sigma_1(n)$ is asymptotically optimal with respect to all priors Q_Θ, $\Theta \in [0, 1]$ if and only if $\sigma_0(n)/n \to \tau_0$ and $\sigma_1(n)/n \to \tau_1$. Here τ_0 and τ_1 are defined by (7). This is shown by straightforward computations with the risk function. The procedures δ_n are admissible if and only if $\sigma_0(n) \geqq \sigma_1(n) + 1$. Hence, many sequences consisting of admissible procedures are not asymptotically optimal. On the other hand, if a sequence of procedures of the form (6) is asymptotically optimal, as $\tau_0 > \tau_1$ its elements are admissible for sufficiently large n.

It is now easy to construct asymptotically optimal sequences consisting of admissible procedures. The classical empirical Bayes approach (see Remark 5.2) yields such a sequence, if Q is estimated by $\hat{\Theta}(n)$. If Θ is estimated by $\hat{\Theta}(n-1)$

this approach yields an asymptotically optimal sequence of which the elements are admissible for sufficiently large n.

Acknowledgment. I want to thank the referees and professor Willem Schaafsma for their careful reading of the manuscript and valuable comments.

REFERENCES

[1] COPAS, J. B. (1974). On symmetric compound decision rules for dichotomies. *Ann. Statist.* **2** 199-204.

[2] HOUWELINGEN, J. C. VAN (1973). On empirical Bayes rules for the continuous one-parameter exponential family. Ph.D. dissertation, University of Utrecht.

[3] INGLIS, J. (1976). Admissible decision rules for the compound decision problem. *Ann. Statist.* To appear.

[4] MARITZ, J. S. (1968). On the smooth empirical Bayes approach to testing of hypotheses and the compound decision problem. *Biometrika* **55** 83-100.

[5] MEEDEN, GLEN (1972). Some admissible empirical Bayes procedures. *Ann. Math. Statist.* **43** 96-101.

[6] ROBBINS, HERBERT (1951). Asymptotically subminimax solutions of compound statistical decision problems. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 131-148, Univ. of California Press.

[7] ROBBINS, HERBERT (1955). An empirical Bayes approach to statistics. *Proc. Third Berkeley Symp. Math. Statist. Prob.* 157-163, Univ. of California Press.

[8] ROBBINS, HERBERT (1964). The empirical Bayes approach to statistical problems. *Ann. Math. Statist.* **35** 1-20.

[9] SAMUEL, ESTER (1967). The compound statistical decision problem. *Sankhyā Ser. A* **29** 123-140.

[10] WALD, ABRAHAM (1950). *Statistical Decision Functions.* Wiley, New York.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GRONINGEN
G.P.O. 800
GRONINGEN
THE NETHERLANDS