This is to certify that the

thesis entitled

On Asymptotic Optimality of
Bayes Empirical Bayes Estimators

presented by

Tze Fen Li

has been accepted towards fulfillment
of the requirements for

Doctoral degree in Statistics

Dennis C. Gilliland

Major professor

Date March 26, 1981

O-7639

ON ASYMPTOTIC OPTIMALITY OF
BAYES EMPIRICAL BAYES ESTIMATORS

By

Tze Fen Li

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

1981

ABSTRACT

ON ASYMPTOTIC OPTIMALITY OF
BAYES EMPIRICAL BAYES ESTIMATORS

By

Tze Fen Li

In an empirical Bayes decision problem, a prior distribution Λ is placed on a one-dimensional family G of priors $G_\omega, \omega \in \Omega$, to produce a Bayes empirical Bayes estimator. The asymptotic optimality of the Bayes estimator is established when the support of Λ is Ω and the marginal distributions H_ω have monotone likelihood ratio and continuous Kullback-Leibler information number.

For the normal case, a simple class of empirical Bayes estimators is constructed that dominate the James-Stein estimator. Here the Bayes estimator is smooth, admissible and asymptotically optimal on G. The rate of convergence to minimum risk is $O(n^{-1})$ uniformly on G. The results of a Monte Carlo study are presented to demonstrate the favorable risk behavior of the Bayes estimator in comparison with other competitors including the James-Stein estimator.

ACKNOWLEDGMENTS

I would like to take this opportunity to express my appreciation to my advisor, Professor Dennis C. Gilliland, and my guidance committee for invaluable guidance, constructive assistance and suggestions during the entire course of this study. The financial support provided by the Department of Statistics and Probability and National Science Foundation made my graduate studies possible. I wish to thank Clara Hanna who accurately typed the thesis with great patience and skill.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION TO BAYES EMPIRICAL BAYES ESTIMATION

1.1 Introduction

Consider the component decision problem consisting of estimation of θ based on X which has a distribution F_θ. Let $L(\theta,\cdot)$ denote a loss function and let $R(G,d)$ denote the risk of an estimator d when G is a prior distribution on θ, i.e.,

$$R(G,d) = \int L(\theta,d(x))\, dF_\theta(x)\, dG(\theta). \tag{1.1}$$

Let $\mathcal{D}$ denote the class of all component estimators d. The infimum risk,

$$R(G) = \inf_{d\in\mathcal{D}} R(G,d), \tag{1.2}$$

defines the Bayes envelope at G. An estimator $d_G \in \mathcal{D}$ such that $R(G,d_G) = R(G)$ is said to be a Bayes component rule versus G.

In the empirical Bayes (EB) decision problem with this component, (θ_i,X_i), $i = 1,2,\ldots$ are i.i.d. with $\theta_i \sim G$, and, conditional on $\theta_i, X_i \sim F_{\theta_i}$. The EB problem is to estimate θ_n based on observing $X_1,\ldots,X_n$. This can be construed as using $\underset{\sim}{X}_{n-1} = (X_1,\ldots,X_{n-1})$ to select a component decision rule $t_n(\underset{\sim}{X}_{n-1}) \in \mathcal{D}$

and evaluating it at X_n to estimate θ_n. (In what follows t_n will sometimes be used to abbreviate the evaluation $t_n(\underline{X}_{n-1})(X_n)$.)

The risk of t_n at G conditional on $\underline{X}_{n-1}$ is $R(G, t_n(\underline{X}_{n-1})) \geq R(G)$, and the overall risk is

$$R_n(G, t_n) = \int R(G, t_n(\underline{x}_{n-1}))dH_G^{n-1}(\underline{x}_{n-1}), \tag{1.3}$$

where H_G^{n-1} denotes the $(n-1)$-fold product of the G-mixture of the F_θ.

Let $\mathcal{G}$ be a specified family of distributions on θ.

<u>Definition 1.1</u>. (Robbins (1956)) A sequence of EB rules t_n is said to be asymptotically optimal (a.o.) on $\mathcal{G}$ if for each $G \in \mathcal{G}$,

$$\lim_n R_n(G, t_n) = R(G).$$

When the component loss is squared error, a Bayes component rule is $d_G(X) = E[\theta|X]$. Furthermore, EB risk (1.3) has the representation

$$R_n(G,t_n) = R(G) + E(t_n-d_G(X_n))^2 \tag{1.4}$$

provided $E\theta^2 < \infty$. This representation was noted by Johns (1956) and used as a starting point to prove the a.o. property of certain EB estimators. It follows from the L_2-orthogonality of $E[\theta_n|X_n] - \theta_n$ and $t_n - E[\theta_n|X_n]$.

In this thesis, $\mathcal{G}$ is assumed to be a parametric family of distributions of θ. Each $G \in \mathcal{G}$ is identified by an element ω of an indexing set Ω which is a subset of the reals, i.e.,

$$\mathcal{G} = \{G_\omega | \omega \in \Omega\}. \tag{1.5}$$

Let Λ be a prior distribution on Ω. An EB estimator t_n is said to be Bayes and called Bayes EB with respect to Λ if it minimizes

$$R_n(\Lambda, t_n) = \int R_n(G_\omega, t_n) d\Lambda(\omega). \tag{1.6}$$

Good (1965) refers to such priors on priors as Type III probabilities. Meeden (1972) illustrates the Bayes approach to empirical Bayes squared error loss estimation problems with several examples. Other literature discussing Bayes empirical Bayes includes Lindley (1971), Gilliland and Hannan (1974), Gilliland, Hannan and Huang (1976), Deely and Lindley (1979) and Gilliland and Boyer (1979).

In the next section we develop the EB and Bayes EB methods for an example. In Section 3 we give a brief review of the related literature. In Chapter II we consider the Bayes EB method and show that it produces a.o. procedures for a variety of EB decision problems.

1.2 Example - Normal Case

In the present section we consider the component consisting of squared error loss estimation of θ based on $X \sim F_\theta = N(\theta, 1)$. First consider the compound decision problem with this component. This consists of estimation of $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$ based on

$\underline{X} = (X_1, X_2, \ldots, X_n) \sim F_{\theta_1} \times F_{\theta_2} \times \ldots \times F_{\theta_n}$ with the compound risk being the average risk across the n components. James and Stein (1961) show that the estimator $\underline{t}^1 = (t_1^1, t_2^1, \ldots, t_n^1)$, where for $i = 1,2,\ldots,n$,

$$t_i^1(\underline{X}) = (1 - \frac{n-2}{S})X_i \quad \text{where} \quad S = \sum_{i=1}^{n} X_i^2, \tag{1.7}$$

has compound risk satisfying

$$\underline{R}\ (\underline{\theta}, \underline{t}^1) < 1 \quad \text{for all} \quad \underline{\theta},\ n \geq 3. \tag{1.8}$$

This demonstrates the inadmissibility of the compound estimator $\underline{X}$ if $n \geq 3$.

Efron and Morris (1972) point out that t_n^1 is a natural EB estimator for the EB problem with this component and a class of normal prior distributions, here parameterized as

$$G = \{N(0, \frac{1-\omega}{\omega}) \mid 0 < \omega \leq 1\}. \tag{1.9}$$

This is seen to be the case since a Bayes component rule versus $G_\omega = N(0, \frac{1-\omega}{\omega})$ is

$$d_\omega(X) = (1-\omega)X, \tag{1.10}$$

and the shrinking factor $(n-2)/S$ in the James-Stein rule is an unbiased estimator of ω based on $\underline{X} \sim H_\omega^n$. H_ω denotes the marginal of the X_i in this EB problem, i.e., $N(0,\omega^{-1})$.

Let G_ω be abbreviated by ω. The Bayes envelope on G is

$$R(\omega) = 1-\omega,\ 0 < \omega \leq 1. \tag{1.11}$$

From (1.4) and (1.10) it follows that for any EB estimator of the form

$$t_n = (1-\phi(S))X_n, \tag{1.12}$$

$$R_n(\omega,t_n) = R(\omega) + E[(\omega-\phi(S))X_n]^2. \tag{1.13}$$

The evaluation of the RHS(1.13) is simple if ϕ is a linear combination of powers (including negative powers) of S since, conditional on ω, $\omega S \sim \chi^2(n)$ and $X_n^2/S \sim$ Beta $(\frac{1}{2}, \frac{n-1}{2})$ are independent and the moments (including negative moments) of S are easily calculated. Efron and Morris (1972) use this calculation for t_n^1 with shrinking factor $\phi_1(S) = (n-2)/S$ to show that

$$R_n(\omega,t_n^1) = R(\omega) + \frac{2\omega}{n}, \quad 0 < \omega \leq 1. \tag{1.14}$$

From (1.14), the James-Stein EB estimator t_n^1 is seen to be a.o. on G with a rate $O(n^{-1})$ uniform in ω. (That t_n^1 is not a.o. on the class of all priors G is pointed out by Susarla (1976).)

We now demonstrate a simple class of a.o. EB estimators each dominating on G the James-Stein estimator t_n^1 for large n. Consider

$$t_n^2 = t_n^1 + \frac{2(n-6)}{S^2} X_n \tag{1.15}$$

Using (1.13) and letting H denote the marginal distribution of ωS, i.e., $\chi^2(n)$,

$$R_n(\omega,t_n^2)-R(\omega) = \int[\omega - \frac{n-2}{S} + \frac{2(n-6)}{S^2}]^2 \, x_n^2 \, dH_\omega^n \, (\underline{x}_n)$$

$$= \int \frac{x_n^2}{S} dH_\omega^n(\underline{x}_n) \int [\omega - \frac{n-2}{S} + \frac{2(n-6)}{S^2}]^2 S \, dH_\omega^n(\underline{x}_n)$$

$$= \frac{\omega}{n} \int [1 - \frac{n-2}{y} + \frac{2(n-6)}{y^2} \omega]^2 \, y \, dH(y)$$

$$= \frac{\omega}{n(n-2)} [2(n-2) + \frac{4\omega^2(n-6)}{n-4} - \frac{8\omega(n-6)}{n-4}]. \quad (1.16)$$

Using the fact that $\omega^2 \leq \omega$, (1.16) and (1.14) combine to show

$$R_n(\omega, t_n^2) < R_n(\omega, t_n^1) - \frac{4\omega^2(n-6)}{n(n-2)(n-4)}, \; n > 6. \quad (1.17)$$

The coefficient $2(n-6)$ of S^{-2} in the definition (1.15) of t_n^2 was chosen from among all constants to produce an adjustment for the James-Stein estimator t_n^1 which results in domination (1.17). Continuing this method of construction leads to EB estimators with nested risk functions. The next two estimators in the construction are

$$t_n^3 = t_n^2 + \frac{2(n-10)^2}{S^3} X_n \quad (1.18)$$

and

$$t_n^4 = t_n^3 + \frac{2(n-14)(n^2 - 28n + 188)}{S^4} X_n \quad (1.19)$$

with

$$R_n(\omega, t_n^3) < R_n(\omega, t_n^2) - \frac{4\omega^3(n-10)^3}{n(n-2)(n-4)(n-6)(n-8)}, \; n > 10 \quad (1.20)$$

and

$$R(\omega, t_n^4) < R_n(\omega, t_n^3) - \frac{4\omega^4(n-14)(n^2-28n+188)}{n(n-2)\ldots(n-12)}, \; n > 14. \tag{1.21}$$

From (1.13) and the fact $\omega \in (0,1]$, each estimator t_n^j can be improved by retracting its shrinking factor $\phi_j(S)$ to the interval $[0,1]$. Each t_n^j is dominated on G by

$$t_n^{j+} = (1 - \phi_j^*(S))X_n \tag{1.22}$$

where $a^* = \max\{\min\{a,1\}, 0\}$.

We now turn to the study of Bayes empirical Bayes rules with respect to a prior Λ on $\omega \in (0,1]$. The estimators are of the same general form as the t_n^j but with more complicated shrinking factors. Moreover, they are monotone in X_n, and, for suitable Λ, admissible on $(0,1]$ and a.o. on $(0,1]$ with a rate $O(n^{-1})$. The results of a Monte Carlo study will demonstrate the favorable risk behavior of the Bayes EB estimator based on uniform Λ.

Let Λ be a prior distribution on $\omega \in (0,1]$. By (1.4) and (1.10)

$$R_n(\Lambda, t_n) = \int R(\omega)d\Lambda(\omega)$$
$$+ \int (t_n(\underline{x}_{n-1})(x_n) - (1-\omega)x_n)^2 dH_\omega^n(\underline{x}_n)d\Lambda(\omega). \tag{1.23}$$

For (1.23), a minimizer is

$$t_n^\Lambda = (1-\hat{\omega})X_n, \tag{1.24}$$

where

$$\hat{\omega} = \frac{\int_0^1 \omega^{\frac{n}{2}+1} e^{\frac{-\omega S}{2}} d\Lambda(\omega)}{\int_0^1 \omega^{\frac{n}{2}} e^{\frac{-\omega S}{2}} d\Lambda(\omega)} . \tag{1.25}$$

Here $\hat{\omega}$ is conditional expectation of ω given $\underset{\sim}{X}_n$ in the model $\omega \sim \Lambda$ and conditional on ω, $X_1,\ldots,X_{n-1}, X_n$ i.i.d. $H_\omega = N(0,\omega^{-1})$. Note that t_n^Λ is unique a.e. Lebesgue on R^n-space. Also note that t_n^Λ is of the same form as the t_n^j introduced earlier but with a more complicated shrinking factor $\phi(S) = \hat{\omega}$.

Whereas the EB estimators t_n^j are not monotone in X_n for fixed $\underset{\sim}{X}_{n-1}$, we have

<u>Remark 1.1</u>. For any prior Λ on $(0,1]$, the Bayes EB rule t_n^Λ is monotonically increasing in X_n for fixed $\underset{\sim}{X}_{n-1}$.

<u>Proof</u>. From (1.24) and the fact $S = \sum_1^n X_i^2$, it suffices to prove $\hat{\omega} \downarrow$ wrt S. But $\hat{\omega}$ is the mean of an exponential family with parameter $-S/2$ and is therefore $\uparrow$ wrt $-S/2$. □

<u>Remark 1.2</u>. For every prior Λ on $(0,1]$, the Bayes EB rule t_n^Λ is admissible on $(0,1]$.

<u>Proof</u>. The uniqueness a.e. Lebesque of a Bayes rule implies it admissibility. □

The Bayes EB estimator t_n^Λ is easily evaluated and the rate of a.o. determined for the prior $\Lambda = \text{Beta}(\alpha,1)$, $\alpha > 0$. Let

$t_n^{(\alpha)}$ be Bayes EB versus Beta $(\alpha, 1)$. (This class of Beta priors was used by Strawderman (1971) in constructing a prior for the compound decision problem. He shows that the compound version $\underline{t}^{(\alpha)}$ of $t_n^{(\alpha)}$ satisfies (1.8) provided $n \geq 4 + 2\alpha$.) With the Beta $(\alpha, 1)$ prior, (1.25) simplifies using integration by parts to

$$\hat{\omega} = \frac{n+2\alpha}{S} - 2(S \int_0^1 u^{\frac{1}{2}n+\alpha-1} e^{\frac{1}{2}(1-u)S} du)^{-1}. \tag{1.26}$$

In case $\frac{n}{2} + \alpha - 1$ is an integer, say m, (1.26) further simplifies through repeated integration by parts to the closed form

$$\hat{\omega} = \frac{2(m+1)}{S} - (\tfrac{1}{2}S)^m [m!(e^{\frac{1}{2}S} - \sum_0^m \frac{(\frac{1}{2}S)^k}{k!})]^{-1}. \tag{1.27}$$

Using (1.27), $\hat{\omega}$ can be easily calculated to any degree of precision.

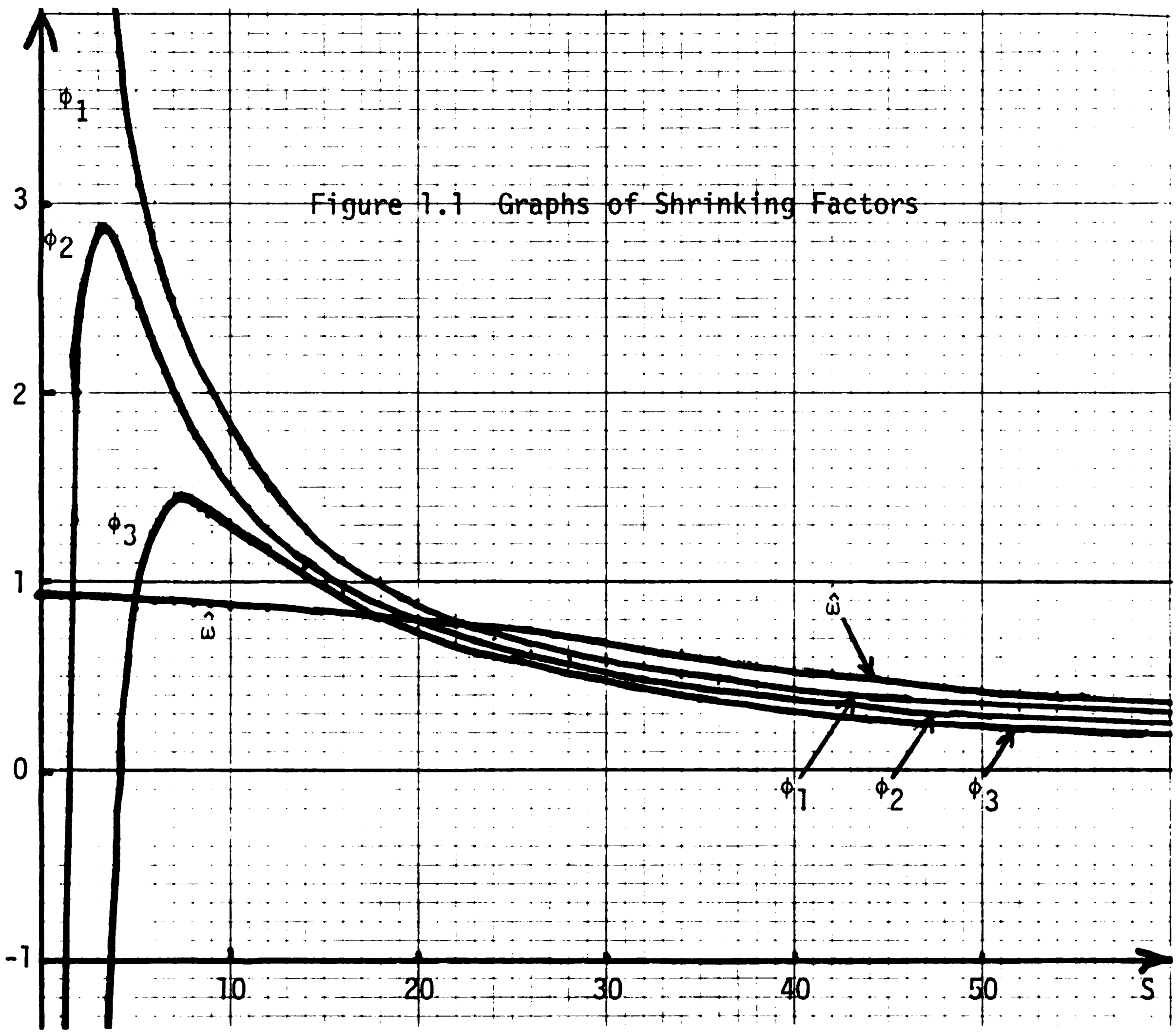
Since the EB risk $R_n(\omega, t_n^{(\alpha)})$ is the expectation of the compound risk $R(\underline{\theta}, \underline{t}^{(\alpha)})$, the aforementioned Strawderman (1971) result implies

<u>Remark 1.3</u>. If $n \geq 4 + 2\alpha$, $R_n(\omega, t_n^{(\alpha)}) < 1$, $0 < \omega \leq 1$. (1.28)

The conditional expectation $\hat{\omega}$ of (1.26) satisfies

$$\lim_{S \downarrow 0} \hat{\omega} = \frac{n+2\alpha}{n+2\alpha+2}, \quad \lim_{S \uparrow \infty} S\hat{\omega} = n+2\alpha. \tag{1.29}$$

Figure 1.1 displays for $n = 20$ the graphs of $\phi_j(S)$ that are part of the EB estimators t_{20}^j, $j = 1,2,3$ as well as $\hat{\omega}$ of (1.27) with $\alpha = 1$ (uniform prior).



The conditional expectation $\hat{\omega}$ of (1.26) is close to the corresponding factor $\phi_1^*(S)$ in the modified James-Stein EB rule t_n^{1+}. Also t_n^1 and $t_n^{(\alpha)}$ have the same rate of a.o., namely, $O(n^{-1})$ uniformly in ω. To establish these results we begin with

<u>Lemma 1.1</u>. Let $\omega_n = \min\{1, (n+2\alpha-2)/S\}$ and let $\hat{\omega}$ be given by (1.26). If $n+2\alpha-2 > 0$, then

$$(\hat{\omega}-\omega_n)^2 \leq \frac{4\pi}{n+2\alpha-2} \quad . \tag{1.30}$$

<u>Proof</u>. Let $\hat{\Lambda}$ denote the conditional distribution of ω in the model $\omega \sim \Lambda$ and conditional on ω, $X_1, X_2, \ldots, X_n$ are i.i.d. $H_\omega = N(0,\omega^{-1})$. Then $\hat{\Lambda}$ has a density with respect to Lebesgue measure which is proportional to

$$\hat{\lambda}(\omega) = \omega^{\frac{n}{2}+\alpha-1} e^{-\frac{\omega S}{2}}, \quad 0 < \omega \leq 1. \tag{1.31}$$

Assume that $n+2\alpha-2 > 0$. Examination of $g(\omega) = \ln \hat{\lambda}(\omega)$ shows that ω_n is the maximizer of $g(\omega)$. Note that

$$\begin{aligned} |\hat{\omega} - \omega_n| &= \left| \int (\omega-\omega_n) d\hat{\Lambda}(\omega) \right| \\ &\leq \int |\omega-\omega_n| d\hat{\Lambda}(\omega) \\ &= \int_0^\infty \hat{\Lambda}[\omega > \omega_n+t]dt + \int_0^\infty \hat{\Lambda}[\omega < \omega_n-t]dt. \end{aligned} \tag{1.32}$$

The density $\hat{\lambda}$ is $\ln$ concave from which it follows that

$$\frac{\hat{\lambda}(\omega)}{\hat{\Lambda}[\omega,1]} \uparrow \text{ in } \omega \in (0,1]. \tag{1.33}$$

(cf. Gilliland, Hannan and Huang (1976), Lemma B)). Also, $\hat{\lambda}(1-\omega)$ is $\ln$ concave from which it follows that

$$\frac{\hat{\lambda}(\omega)}{\hat{\Lambda}(0,\omega]} \downarrow \text{ in } \omega \in (0,1]. \tag{1.34}$$

Thus,

$$\begin{aligned} \hat{\Lambda}\,[\omega > \omega_n+t] &\leq \hat{\lambda}(\omega_n+t)/\hat{\lambda}(\omega_n) \\ \hat{\Lambda}\,[\omega < \omega_n-t] &\leq \hat{\lambda}(\omega_n-t)/\hat{\lambda}(\omega_n). \end{aligned} \tag{1.35}$$

The Taylor series expansion of $g = \ell n\, \hat{\lambda}$ about $\omega = \omega_n$ shows that

$$g(\omega_n+t)-g(\omega_n) \le tg'(\omega_n) - \frac{1}{4}(n+2\alpha-2)t^2$$

$$g(\omega_n-t)-g(\omega_n) \le -tg'(\omega_n) - \frac{1}{4}(n+2\alpha-2)t^2 \tag{1.36}$$

for t such that $\omega_n \pm t \in (0,1]$. Since $g'(\omega_n) = 0$ if $\omega_n \in (0,1)$ and $g'(\omega_n) \ge 0$ if $\omega_n = 1$, it follows from (1.36) that for t such that $\omega_n \pm t \in (0,1]$,

$$g(\omega_n \pm t)-g(\omega_n) \le -\frac{1}{4}(n+2\alpha-2)t^2. \tag{1.37}$$

Using $\exp\{-\frac{1}{4}(n+2\alpha-2)t^2\}$ to bound RHS (1.35), (1.32) and the inequality $\int_0^\infty \exp\{-\frac{1}{2}at^2\}dt \le \sqrt{\pi/2a}$, the proof of (1.30) is complete. □

Remark 1.4. If $n+2\alpha-2 > 0$,

$$R_n(\omega,t_n^{(\alpha)})-R(\omega) \le \frac{8\pi}{n+2\alpha-2}\,\omega^{-1} + \frac{4n+8(\alpha^2-1)}{n(n-2)}\,\omega. \tag{1.38}$$

Proof. Using (1.13), and triangulation about ω_n of Lemma 1.1,

$$R_n(\omega,t_n^{(\alpha)})-R(\omega) \le 2\,E[(\hat{\omega}-\omega_n)^2X_n^2] + 2\,E[(\omega_n-\omega)^2X_n^2]. \tag{1.39}$$

By Lemma 1.1 and the fact $E\,X_n^2 = \omega^{-1}$,

$$E[(\hat{\omega}-\omega_n)^2X_n^2] \le \frac{\pi 4}{n+2\alpha-2}\,\omega^{-1}. \tag{1.40}$$

Also

$$E[(\omega_n-\omega)^2 X_n^2] \leq E\,[(\frac{n+2\alpha-2}{S}-\omega)^2 X_n^2]. \tag{1.41}$$

Expanding the square in (1.41) and using the fact $\omega S \sim \chi^2(n)$ independent of X_n^2/S, it follows that

$$\text{RHS}(1.41) \leq \frac{2n+4(\alpha^2-1)}{n(n-2)}\,\omega, \tag{1.42}$$

which together with (1.39) - (1.41) completes the proof. □

The bound in (1.38) fails to demonstrate the uniform $O(n^{-1})$ rate on $(0,1]$. The following theorem does establish this uniform rate by combining (1.38) with a bound designed for risk in a neighborhood of $\omega = 0$.

Theorem 1.1. The Bayes EB estimator $t_n^{(\alpha)}$ is a.o. with a rate $O(n^{-1})$ uniform in $\omega \in (0,1]$.

Proof. By (1.13), (1.26), and the fact $n\,E\,[(\hat{\omega}-\omega)^2 X_n^2] = E\,[(\hat{\omega}-\omega)^2 S]$ (by the symmetry of $\hat{\omega}$ in $X_1, X_2, \ldots, X_n$),

$$n[R_n(\omega, t_n^{(\alpha)}) - R(\omega)] = E\,[\tfrac{1}{S}\,\{n-2-\omega S + 2(1+\alpha) - 2f(S)\}^2] \tag{1.43}$$

where

$$f(S) = (\int_0^1 u^{\frac{1}{2}n+\alpha-1}\, e^{\frac{1}{2}(1-u)S}\, du)^{-1}. \tag{1.44}$$

Let $n > 2$. Note that $E\,[S^{-1} g(S)] = \omega(n-2)^{-1}\,E\,g(Y)$ where $\omega S \sim \chi^2(n)$ and $\omega Y \sim \chi^2(n-2)$ by a change of variable argument.

Now the variance of ωY is

$$E\{(n-2)-\omega Y\}^2 = 2(n-2) \tag{1.45}$$

and

$$\mathrm{Cov}(Y,f(Y)) < 0 \tag{1.46}$$

since $f(Y) \downarrow$ with respect to Y, so that (1.43) shows that

$$\mathrm{LHS}(1.43) \leq 2\,\omega + \frac{4\omega E[\{(1+\alpha)-f(Y)\}^2]}{n-2}\,. \tag{1.47}$$

The proof is completed by showing that $E\,f^2(Y)$ is $O(n)$ uniformly in $\omega \in (0,\frac{1}{4}]$ and using (1.38) to establish the uniform rate on $[\frac{1}{4}, 1]$. From (1.44), for any $b > 0$

$$E\,f^2(Y) \leq \left(\int_0^1 u^{\frac{1}{2}n+\alpha-1}du\right)^{-2} P[\omega Y \leq b] + \left(\int_0^1 u^{\frac{1}{2}n+\alpha-1}e^{(1-u)b/2\omega}du\right)^{-2} \tag{1.48}$$

where use is made of the fact that the integrand in (1.44) is increasing in S. The choice $b = \frac{1}{2}n+\alpha-1$ ensures that the $P[\omega Y \leq b] \leq 2(n-2)$ $(\frac{1}{2}n-\alpha-1)^{-2}$ (use the Chebyshev inequality and the fact $\omega\,Y \sim \chi^2(n-2)$). With this choice of b,

$$u^{\frac{1}{2}n+\alpha-1}e^{(1-u)b/2\omega} \geq 1, \quad 2\omega \leq u \leq 1 \tag{1.49}$$

so that the last term on RHS(1.48) is bounded by $(1-2\omega)^{-2}$. Hence, $E\,f^2(Y) = O(n)$ uniformly in $\omega \in (0, \frac{1}{4}]$. □

Maritz (1970, Chapter 3) proposes several methods of obtaining "smooth" EB estimators. He illustrates two of these methods for the normal case EB example of this section.

An estimator $\hat{\delta}_G$ is described in Maritz Examples 3.4.4 and 2.14.2 which is

$$\hat{\delta}_G = (1-\hat{\omega}_M^*)X_n, \tag{1.50}$$

where $\hat{\omega}_M^*$ is the retraction of the shrinking factor

$$\hat{\omega}_M = \frac{n-1}{\check{S}} \tag{1.51}$$

and $\check{S} = \sum_{i=1}^{n-1} X_i^2$.

The estimator $\hat{\delta}_G$ is seen to be a delete version of t_n^{1+}, where delete refers to the fact that only the initial observations $\underline{X}_{n-1}$ are used to estimate ω. By (1.13),

$$\begin{aligned} R_n(\omega,\hat{\delta}_G)-R(\omega) &= \omega^{-1}\ E\ (\omega-\hat{\omega}_M^*)^2 \\ &\le \omega^{-1}\ E\ (\omega-\frac{n-1}{\check{S}})^2 \\ &= \frac{2(n+3)\omega}{(n-3)(n-5)}\ , \end{aligned} \tag{1.52}$$

so $\hat{\delta}_G$ is seen to be a.o. with a rate $O(n^{-1})$ uniform in $\omega \in (0,1]$. For comparison, the James-Stein estimator, which uses the untruncated $\phi_1(S) = (n-2)/S$ to estimate ω, has excess risk $2\omega/n$.

The Maritz estimator $\hat{\emptyset}_3$ is constructed by finding the 3-point uniform distribution $\hat{G}$ on $\theta \in (-\infty,\infty)$ among all such distributions so that the $\hat{G}$-mixture of $N(\theta,1)$ minimizes a distance between the mixture and the empirical distribution of $X_1,X_2,\ldots,X_{n-1}$.

(See Maritz (1970, pp. 54-55).) The EB estimator of θ_n is then taken to be

$$\hat{\phi}_3 = d_{\hat{G}}(X_n) \ . \tag{1.53}$$

Both Maritz EB estimators are "smooth" in the sense of being monotone in X_n for fixed $\underline{X}_{n-1}$. However, as the following table shows, both estimators perform poorly relative to $t_n^{(1)}$ and even t_n^{1+} and t_n^1 for selected values of n and ω.

Table 1.1. $R_n(\omega,t_n)-R(\omega)$ Values

σ^2	ω	$R(\omega)$		n = 4	n = 10	n = 20
			$t_n^{(1)}$	.13	.05	.02
5	.167	.833	t_n^{1+}	.09	.03	.02
			t_n^1	.08	.03	.02
			$t_n^{(1)}$	.09	.06	.04
2	.333	.667	t_n^{1+}	.18	.06	.03
			t_n^1	.17	.07	.03
			$t_n^{(1)}$	.06	.04	.04
			t_n^{1+}	.25	.09	.05
1	.500	.500	t_n^1	.25	.10	.05
			$\hat{\delta}_G$	—	.17	.10
			$\hat{\phi}_3$	—	.32	.19
			$t_n^{(1)}$	.07	.03	.03
			t_n^{1+}	.33	.10	.06
.5	.667	.333	t_n^1	.33	.13	.07
			$\hat{\delta}_G$	—	.16	.09
			$\hat{\phi}_3$	—	.22	.14

Table 1.1. (Continuation)

			$t_n^{(1)}$	.16	.07	.04
			t_n^{1+}	.44	.12	.07
.1	.909	.091	t_n^1	.46	.18	.09
			$\hat{\delta}_G$	___	.12	.07
			$\hat{\emptyset}_3$	___	.16	.07
			$t_n^{(1)}$	.20	.10	.06
.01	.990	.010	t_n^{1+}	.48	.13	.07
			t_n^1	.50	.20	.10

The values for the James-Stein estimator t_n^1 are exact $2\omega/n$. The values for the modified James-Stein estimator t_n^{1+} and the Bayes EB rule $t_n^{(1)}$ are estimates based on 1000 Monte Carlo trials. The estimated standard deviations are generally about 8% of the estimated excess risks. The values for the Maritz estimators $\hat{\delta}_G$ and $\hat{\emptyset}_3$ are taken from his Table 3.14 (1970) where the margins of errors of the Monte Carlo estimates are reported to be less than $\pm$ .02. Also the Maritz risks are based on EB decision problems with $n = 11$ (not 10) and $n = 21$ (not 20) observations.

Figure 1.1, Lemma 1.1 and Theorem 1.1 suggest that $t_n^{(1)}$ and the modified James-Stein estimator t_n^{1+} should have very similar EB risk behavior for moderate to large n. Also t_n^{1+} and t_n^1 should have very similar EB risk when ω is small since $\omega S \dot{\sim} \chi^2(n)$ and $t_n^{1+} = t_n^1$ if $S \geq n\text{-}2$. Table 1.1 illustrates these facts. Furthermore, it shows that $\hat{\emptyset}_3$ (and $\hat{\delta}_G$) are poor estimators in the tested combinations of ω and n contrary to the conclusion Maritz (1970, p. 72) reaches by comparing $\hat{\emptyset}_3$ and $\hat{\delta}_G$ with the simple estimator X_n.

1.3 Literature Review.

Gilliland and Hannan (1974) and Gilliland, Hannan and Huang (1976) discuss Bayes procedures for the finite state compound decision problem. Bayes compound procedures with respect to mixtures of product distributions are shown to be Bayes EB procedures. Implications for the EB problem are discussed including asymptotic optimality. Gilliland and Boyer (1979) demonstrate that a.o. in the finite state EB problem is an easy consequence of classical results on the consistency of posterior distributions. Tsao (1980a, 1980b) gives an algorithm to efficiently compute Bayes EB procedures and uses a Monte Carlo simulation to develop small n risk behavior for selected Λ and the Robbins two state component.

In the finite state case, the unrestricted family G of distributions on θ is finite-dimensional. Otherwise, G is infinite-dimensional and the process of placing priors Λ on G is itself a technical problem. Meeden (1972) for certain components, with θ restricted to the unit interval, places a prior Λ on G through the moment sequence and demonstrates the a.o. of the resulting Bayes EB estimator. Kuo (1980) proposes a way to compute the Bayes EB estimators when Λ is a Dirichlet process. The a.o. property is not established in this general setting. Deely and Lindley (1979) illustrate the use of Bayes procedures in EB decision making but do not consider a.o..

CHAPTER II

ASYMPTOTIC OPTIMALITY OF BAYES EMPIRICAL BAYES RULES

Consider the EB decision problem of Section 1.1 with the one-dimensional family G of component priors of (1.5). We let F_θ denote the distribution of X given θ and when $\theta \sim G_\omega$ let H_ω denote the marginal distribution of X. Throughout this chapter we assume that the family $\{H_\omega | \omega \in \Omega\}$ is identifiable and dominated and let $h(x|\omega)$ denote a density for H_ω. It is assumed that $h(x|\omega)$ is jointly measurable in x and ω. This assumption is part of the hypothesis of a Schwartz theorem on consistency of posterior distributions, which will be applied in the Bayes EB approach where a probability measure Λ is placed on ω.

From (1.3) and (1.6),

$$R_n(\Lambda,t_n) = \iint R(\omega,t_n(\underline{x}_{n-1}))d\hat{\Lambda}_n(\omega)dP(\underline{x}_{n-1}) \tag{2.1}$$

where, here and throughout, $\hat{\Lambda}_n(\cdot) = \Lambda(\cdot|\underline{x}_{n-1})$ is the conditional distribution of ω given $\underline{X}_{n-1}$ in the model where $\omega \sim \Lambda$ and conditional on ω, $X_1,X_2,\ldots,X_{n-1}$ are i.i.d. H_ω. As Gilliland and Boyer (1979) point out, it follows from (2.1) that a Bayes EB rule is provided by

$$t_n^\Lambda(\underline{x}_{n-1}) = d_{\hat{G}_n} \tag{2.2}$$

where $d_{\hat{G}_n}$ is component Bayes versus the mixture

$$\hat{G}_n(\circ) = \int G_\omega(\circ)\, d\hat{\Lambda}_n(\omega). \tag{2.3}$$

Of course, the random measure $\hat{G}_n$ need not be $\mathcal{G}$-valued where $\mathcal{G} = \{G_\omega | \omega \in \Omega\}$. In the normal example of Section 1.2, $G_\omega = N(0, (1-\omega)/\omega)$ and nondegenerate Λ-mixtures of G_ω are not normal. (cf. Teicher (1960), Corollary, p. 67.)

A basic condition critical for our proofs of a.o. for Bayes EB procedures is

$$\hat{\Lambda}_n(V^c) \to 0 \quad \text{a.s.} \quad H^\infty_{\omega_o} \tag{A}$$

for every $\omega_o \in \Omega$ and every neighborhood V of ω_o

where V^c denotes the complement of V in Ω. The condition (A) is easily seen to imply

$$\int \psi(\omega) d\hat{\Lambda}_n(\omega) \to \psi(\omega_o) \quad \text{a.s.} \quad H^\infty_{\omega_o} \tag{A-}$$

for every $\omega_o \in \Omega$ and every bounded continuous function ψ.

Theorems 2.1, 2.2, and 2.3 give conditions which together with (A-) imply the a.o. property for the Bayes EB procedure t_n^Λ. Theorem 2.4 states conditons on the family $\{H_\omega | \omega \in \Omega\}$ sufficient for (A). The chapter is completed with four examples of Bayes EB procedures whose a.o. property follows from the theorems.

Theorem 2.1. Suppose that the component risk functions $R(\theta,d)$ are bounded (by $M < \infty$). Suppose each G_ω has a density with respect to Lebesgue measure given by $g(\theta|\omega)$ which is continuous in $\omega \in \Omega$ for each fixed $\theta \in \Theta$. Then if (A-) obtains, t_n^Λ of (2.2) is a.o. on Ω, that is,

$$R_n(\omega_\circ, t_n^\Lambda) \to R(\omega_\circ) \quad \text{for all} \quad \omega_\circ \in \Omega. \tag{2.4}$$

Proof. Fix $\omega_\circ \in \Omega$. Note that for each prior $\hat{G}$ on θ,

$$0 \leq R(\omega_\circ, d_{\hat{G}}) - R(\omega_\circ) = R(\omega_\circ, d_{\hat{G}}) - R(\hat{G}, d_{\hat{G}}) + R(\hat{G}, d_{\hat{G}}) - R(\omega_\circ) \leq [R(\omega_\circ, d_{\hat{G}}) - R(\hat{G}, d_{\hat{G}})] + [R(\hat{G}, d_{\omega_\circ}) - R(\omega_\circ, d_{\omega_\circ})]. \tag{2.5}$$

Let $\hat{G}_n$ be given by (2.3) and note that $\hat{G}_n$ has density

$$\hat{g}_n(\theta) = \int g(\theta|\omega) d\hat{\Lambda}_n(\omega). \tag{2.6}$$

For any $d \in \mathcal{D}$,

$$R(\omega_\circ, d) - R(\hat{G}_n, d) = \int R(\theta, d) \int \{g(\theta|\omega_\circ) - g(\theta|\omega)\} d\hat{\Lambda}_n(\omega) d\theta$$

from which

$$|R(\omega_\circ, d) - R(\hat{G}_n, d)| \leq M \iint |g(\theta|\omega_\circ) - g(\theta|\omega)| d\theta d\hat{\Lambda}_n(\omega). \tag{2.7}$$

Note that $|g(\theta|\omega_\circ) - g(\theta|\omega)| \leq g(\theta|\omega_\circ) + g(\theta|\omega)$ and $\int \{g(\theta|\omega_\circ) + g(\theta|\omega)\} d\theta = 2$. It follows from the assumed continuity of $g(\theta|\omega)$ in ω and the general dominated convergence theorem (DCT) (Royden (1968, p. 89)), that $\int |g(\theta|\omega_\circ) - g(\theta|\omega)| d\theta$ is a continuous as well as bounded function of ω. Therefore, (A-) implies

$$\text{RHS}(2.7) \to 0 \quad \text{a.s.} \quad H_{\omega_\circ}^\infty . \tag{2.8}$$

Letting $\hat{G} = \hat{G}_n$ in (2.5) and using the bound RHS (2.7) for each of the square bracket terms, one obtain

$$R(\omega_\circ, t_n^{\Lambda}(\underline{X}_{n-1})) - R(\omega_\circ) \to 0 \quad \text{a.s.} \quad H_{\omega_\circ}^{\infty}. \tag{2.9}$$

Since all risks are bounded, the a.o. of t_n^{Λ} follows from (2.9) and the DCT. □

The Bayes EB rule t_n^{Λ} defined by (2.2) has a useful alternative representation in the special case of squared error loss estimation and when the family $\{F_\theta | \theta \in \Theta\}$ is dominated. Let $f(x|\theta)$ denote a density for F_θ. Since $d_G(x)$ is the point estimator

$$d_G(x) = \frac{\int \theta\, f(x|\theta)dG(\theta)}{\int f(x|\theta)dG(\theta)} \tag{2.10}$$

and $\hat{G}_n$ is a $\hat{\Lambda}_n$-mixture of G_ω, (2.2) implies that

$$t_n^{\Lambda}(\underline{x}_{n-1})(x_n) = \frac{\iint \theta\, f(x_n|\theta)dG_\omega(\theta)d\hat{\Lambda}_n(\omega)}{\iint f(x_n|\theta)dG_\omega(\theta)d\hat{\Lambda}_n(\omega)}$$

$$= \frac{\int d_\omega(x_n)h(x_n|\omega)d\hat{\Lambda}_n(\omega)}{\int h(x_n|\omega)d\hat{\Lambda}_n(\omega)}$$

$$= \frac{\int d_\omega(x_n) \prod_{i=1}^{n} h(x_i|\omega)d\Lambda(\omega)}{\int \prod_{i=1}^{n} h(x_i|\omega)d\Lambda(\omega)}$$

Thus, t_n^{Λ} is the point estimator

$$t_n^{\Lambda}(\underline{x}_{n-1})(x_n) = \int d_\omega(x_n)d\hat{\Lambda}_{n+1}(\omega). \tag{2.11}$$

Theorem 2.2. Suppose that the component loss function is squared error loss estimation of θ. Suppose that for each $\omega \in \Omega$, the component Bayes rule d_ω with respect to G_ω is of the form

$$d_\omega(X) = \sum_{i=1}^{p} \varphi_i(X)\psi_i(\omega) \tag{2.12}$$

for some integer p, some square integrable (H_ω) functions φ_i and some bounded, continuous functions ψ_i. Then if (A-) obtains, t_n^Λ of (2.11) is a.o. on Ω.

Proof. Fix $\omega_\circ \in \Omega$. By (2.11) and (2.12),

$$t_n^\Lambda(\underline{x}_{n-1})(x_n) - d_{\omega_\circ}(x_n) = \sum_{i=1}^{p} \varphi_i(x_n)\{\hat{\psi}_i - \psi_i(\omega_\circ)\} \tag{2.13}$$

where

$$\hat{\psi}_i = \int \psi_i(\omega)d\hat{\Lambda}_{n+1}(\omega),\ i=1,2,\ldots,p. \tag{2.14}$$

By (2.13),

$$[t_n^\Lambda(\underline{x}_{n-1})(x_n) - d_{\omega_\circ}(x_n)]^2 \leq p \sum_{i=1}^{p} \varphi_i^2(x_n)\{\hat{\psi}_i - \psi_i(\omega_\circ)\}^2. \tag{2.15}$$

Using (1.13) and (2.15) and using the invariance of $\hat{\psi}_i$ and the distribution H_ω^n under permutation of $x_1, x_2, \ldots, x_n$ to permute x_1 and x_n,

$$0 \leq R_n(\omega_\circ, t_n^\Lambda) - R(\omega_\circ) \leq p \sum_{i=1}^{p} \int \varphi_i^2(x_1)\{\hat{\psi}_i - \psi_i(\omega_\circ)\}^2 dH_{\omega_\circ}^n(\underline{x}_n). \tag{2.16}$$

By (A-), $\hat{\psi}_i \to \psi_i(\omega_\circ)$ a.s. $H_{\omega_\circ}^\infty$. Therefore, by the assumed integrability of the φ_i^2 and the DCT, RHS (2.16) $\to 0$. □

Consider the linear loss multiple decision problem of Van Ryzin and Susarla (1977). From (33) of Van Ryzin and Susarla (1977) or (14) of Gilliland and Hannan (1977), it follows that the excess risk of a Bayes EB procedure can be bounded by a multiple of the mean error of estimator of $d_\omega(X_n)$ by RHS (2.11) rather than the mean square error as with the squared error loss estimation component.

Theorem 2.3. Consider the linear loss k-action multiple decision problem of Van Ryzin and Susarla (1977). Suppose that for each $\omega \in \Omega$, the component conditional expectation $d_\omega(X)$ is given by (2.12) for some integer p, some integrable (H_ω) function φ_i and some bounded, continuous function ψ_i. Then if (A-) obtains, t_n^Λ as a.o. on Ω.

Proof. From the $\varepsilon = 1$ case of (14) of Gilliland and Hannan (1977),

$$0 \leq R_n(\omega_\circ, t_n^\Lambda) - R(\omega_\circ) \leq (k-1)\sum_{i=1}^{p} \int |\varphi_i(x_1)\{\hat{\psi}_i - \psi_i(\omega_\circ)\}| dH_{\omega_\circ}^n(\underline{x}_n) \quad (2.17)$$

where use is made of (2.11) and (2.12) and the invariance under permutations used to reach (2.16). By (A-), $\hat{\psi}_i \to \psi_i(\omega_\circ)$ a.s. $H_{\omega_\circ}^\infty$. Therefore, by the assumed integrability of the φ_i and the DCT, RHS (2.17) $\to 0$. □

The next theorem gives a set of conditions sufficient for (A) and hence (A-). It depends for its proof upon Theorem 6.1 of Schwartz (1965) which we state in a notation appropriate to the application at hand.

Theorem 6.1 (Schwartz). Suppose that (i) the denisities $h(x|\omega)$ are jointly measurable (ii) V is a neighborhood of ω_o and there is a uniformly consistent test of the hypothesis $\omega = \omega_o$ against the alternative $\omega \in V^c$, and (iii) for every $\varepsilon > 0$ V contains a subset W such that $\Lambda(W) > 0$ and the Kullback-Leibler information number

$$K(\omega,\omega_o) = \int h(x|\omega)dH_{\omega_o}(x) \tag{2.18}$$

satisfies $K(\omega,\omega_o) > K(\omega_o,\omega_o) - \varepsilon$ on W. Then $\hat{\Lambda}_n(V^c) \to 0$ a.s. $H^{\infty}_{\omega_o}$.

Theorem 2.4. Suppose that for each n, the joint density $h(\underline{x}_n|\omega) = \pi\, h(x_i|\omega)$ has a monotone likelihood ratio (MLR) in $T_n(\underline{x}_n)$. Suppose that the Kullback-Leibler information number $K(\omega,\omega_o)$ is finite and is continuous in ω for each $\omega_o \in \Omega$. Then if Λ has support equal to Ω, (A) obtains.

Proof. Let $\omega_o,\omega_1 \in \Omega$, $\omega_1 < \omega_o$. There exists a consistent test δ_n of H_{ω_o} versus H_{ω_1} since $H_{\omega_1} \neq H_{\omega_o}$ by identifiability. Let δ_n^* be the Neyman-Pearson test based on T_n of the same size as δ_n. Then δ_n^* is uniformly consistent for H_{ω_o} versus $\{H_\omega|\omega \in \Omega$ and $\omega \leq \omega_1\}$ by the MLR property. Similarly, there exists a uniformly consistent test of H_{ω_o} versus $\{H_\omega|\omega \in \Omega$ and $\omega \geq \omega_2\}$ where $\omega_2 \in \Omega$, $\omega_2 > \omega_o$. It follows that there exists a uniformly consistent test of H_{ω_o} versus $\{H_{\omega_o}|\omega \in \Omega$ and $\omega \notin (\omega_1,\omega_2)\}$. Moreover, the continuity of K implies that for every Ω-neighborhood V of ω, and every $\varepsilon > 0$ there exists a subset W of V such that $\Lambda(W) > 0$ and $K(\omega,\omega_o)-K(\omega_o,\omega_o) > -\varepsilon$ on W. Theorem 6.1 establishes that (A) holds. □

The family of densities

$$h(x|\omega) = c(\omega)e^{Q(\omega)T(x)}h(x),\ \omega \in \Omega \tag{2.19}$$

where Q is a monotone function, results in the family $h(\underline{x}_n|\omega)$ being MLR in $T_n(\underline{x}_n)= \sum T(x_i)$. In each of the examples to follow, the mixtures $\{H_\omega|\omega \in \Omega\}$ form a one-parameter exponential family.

<u>Example 2.1 (Normal case of §1.2)</u>. Let $F_\theta = N(\theta,1), \theta \in \Theta = (-\infty,\infty)$ and $G_\omega = N(0,(1-\omega)/\omega)$, $\omega \in \Omega = (0,1]$. Then

$$h(x|\omega) = \sqrt{\frac{\omega}{2\pi}}\ e^{-\frac{1}{2}\omega x^2},\ -\infty < x < \infty.$$

Moreover,

$$K(\omega,\omega_o) = -\tfrac{1}{2}[\ln(2\pi) - \ln \omega + \frac{\omega}{\omega_o}].$$

Here the component conditional expectation is $(1-\omega)X$ and note that $\psi(\omega) = 1-\omega$ is bounded and continuous and X is square integrable (H_ω). The density $g(\theta|\omega)$ is continuous in ω for each fixed θ.

<u>Example 2.2</u>. Let F_θ = Poisson (θ), $\theta \in (0,\infty)$ and let G_ω = Gamma $(\alpha_o,\omega/(1-\omega))$, $\omega \in (0,1)$, where $\alpha_o > 0$ is fixed. Then

$$h(x|\omega) = \frac{\Gamma(\alpha_o+x)}{\Gamma(\alpha_o)x!}\ \omega^x\ (1-\omega)^{\alpha_o},\ x = 0,1,\ldots\ .$$

Here $K(\omega,\omega_o)$ is continuous in ω and component conditional expectation is $(\alpha_o + X)\omega$. Also, ω is bounded and continuous and

$\alpha_o + X$ is square integrable (H_ω). The density $g(\theta|\omega)$ is continuous in ω for fixed θ.

Example 2.3. Let F_θ be Binomial $(n = 1,\theta)$, $\theta \in [0,1]$ and let G_ω = Beta $(\alpha_o,\alpha_o(1-\omega)/\omega)$, $\omega \in (0,1)$, where $\alpha_o > 0$ is fixed. Then

$$h(x|\omega) = \omega^x (1-\omega)^{1-x}, \; x = 0,1.$$

Here $K(\omega,\omega_o)$ is continuous in ω and the component conditional expectation is $(\alpha_o + X)\omega/(\alpha_o + \omega)$. Also, $\omega/(\alpha_o + \omega)$ is bounded and continuous and $\alpha_o + X$ is square integrable (H_ω). The density $g(\theta|\omega)$ is continuous in ω for fixed θ.

Example 2.4. Let F_θ = Uniform $(0,\theta)$, $\theta \in (0,\infty)$ and G_ω = Gamma $(2,\omega^{-1})$, $\omega \in [a,\infty)$ where $a > 0$. Then

$$h(x|\omega) = \omega\, e^{-\omega x}, \; x > 0.$$

Here $K(\omega,\omega_o)$ is continuous in ω and the conditional expectation in $X + \omega^{-1} = X\cdot 1 + 1\cdot\omega^{-1}$ and note that X and 1 are square integrable (H_ω) and 1 and ω^{-1} are bounded and continuous on $[a,\infty)$. The density $g(\theta|\omega)$ is continuous in ω for fixed θ.

Consider any EB decision problem with component probability structure that of any of the Examples. If the loss structure is either bounded, or is squared error loss estimation, or is linear loss multiple decision. Then the Theorems show that a Bayes EB procedure t_n^Λ will be a.o. on the support of Λ, and, therefore, on $G = \{G_\omega | \omega \in \Omega\}$ if the support of Λ is equal to Ω.

The results of this thesis pertain to EB problems with one-dimensional families G. In generalizing to higher (finite) dimensional families, there is no problem invoking the Schwartz consistency theorem except in the demonstration of a uniformly consistent test. For infinite dimensional families, the technical problems are much greater and only a few examples of a.o. Bayes EB procedures have been given (c.f. Meeden (1972)).

BIBLIOGRAPHY

BIBLIOGRAPHY

Deely, J.J. and Lindley, D.V. (1979). Bayes empirical Bayes. University of Canterbury, Christ Church and University College, London. (private communication).

Efron, Bradley and Morris, Carl (1972). Limiting the risk of Bayes and empirical Bayes estimators, Part II: The empirical Bayes case. J. Amer. Statist. Assoc. 67, 130-139.

Efron, Bradley and Morris, Carl (1973). Stein's estimation rule and its competitors -- An empirical Bayes approach. J. Amer. Statist. Assoc. 68, 117-130.

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

Gilliland, Dennis C. and Boyer, John E., Jr. (1979). Bayes empirical Bayes. Dept. of Statistics and Probability, MSU. (submitted for publication).

Gilliland, Dennis C. and Hannan, James (1974). The finite state compound decision problem, equivariance and restricted risk components. RM-317, Statistics and Probability, MSU.

Gilliland, Dennis C. and Hannan, James (1977). Improved rates in the empirical Bayes monotone mulitple decision problem with MLR family. Ann. Statist. 5, 516-521.

Gilliland, Dennis C., Hannan, James and Huang, J.S. (1976). Asymptotic solutions to the two state component compound decision problem, Bayes versus diffuse priors on proportions. Ann. Statist. 4, 1101-1112.

Good, I.J. (1965). The estimation of probabilities: An essay on modern Bayesian methods. Research Monograph No. 30, M.I.T. Press.

James, W. and Stein, C. (1961). Estimation with quadratic loss. Proc. Fourth Berkeley Symp. Math. Statist. Prob., 361-379. University of California Press.

Johns, M.V., Jr. (1965). Contributions to the theory of non-parametric empirical Bayes procedures in statistics. Ph.D. Dissertation, Columbia.

Kuo, Lynn (1980). Computations of mixtures of Dirichlet processes. Technical Report No. 96, Dept. Stat., University of Michigan.

Lindley, D.V. (1971). Bayesian statistics, a review. Regional Conference Series in Applied Mathematics No. 2, SIAM, Philadelphia.

Maritz, J.S. (1970). Empirical Bayes methods. Methuen and Co. Ltd, London.

Meeden, Glen (1972). Some admissible empirical Bayes procedures. Ann. Math. Statist. 43, 96-101.

Robbins, H. (1956). An empirical Bayes approach to statistics. Proc. Third Berkeley Symp. Math. Statist. Prob. 1, University of California Press, 157-163.

Royden, H.L. (1968). Real Analysis, 2nd Edition, Macmillan, New York.

Schwartz, Lorraine (1965). On Bayes' procedures. Z Wahrscheinlichkeitstheorie und verw. Gebiete 4, 10-26.

Strawderman, William E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. Ann. Statist. 42, 385-388.

Susarla, V. (1976). A property of Stein's estimator for the mean of a multivariate normal distribution. Statistica Neerlandica 30, 1-5.

Teicher, Henry (1960). On the mixture of distributions. Ann. Math. Statist. 31, 55-73.

Tsao, How Jan (1980a). On the risk performance of Bayes empirical Bayes procedures for classification between N(-1,1) and N(1,1). Statistic. Neerlandica 34.

Tsao, How Jan (1980b). On the risk performance of Bayes empirical Bayes procedures in the finite state component case. Ph.D. Dissertation, Dept. Stat. Prob., Michigan State University.

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