Empirical Bayes Estimation of Binomial Parameter with Symmetric Priors*

by

TaChen Liang
Department of Mathematics
Wayne State University
Detroit, MI 48202

Technical Report # 89–13

Department of Statistics
Purdue University

June 1989

* This research was supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grants DMS-8606964, DMS-8702620 at Purdue University.

△

EMPIRICAL BAYES ESTIMATION OF BINOMIAL PARAMETER WITH SYMMETRIC PRIORS

TaChen Liang
Department of Mathematics
Wayne State University
Detroit, MI 48202

ABSTRACT

This paper deals with the problem of estimating the binomial parameter via the nonparametric empirical Bayes approach. This estimation problem has the feature that estimators which are asymptotically optimal in the usual empirical Bayes sense do not exist (Robbins (1956, 1964)). However, as pointed out by Liang (1984) and Gupta and Liang (1986), it is possible to construct asymptotically optimal empirical Bayes estimators if the unknown prior is symmetric about the point 1/2. In this paper, assuming symmetric priors a monotone empirical Bayes estimator is constructed by using the isotonic regression method. This estimator is asymptotically optimal in the usual empirical Bayes sense. The corresponding rate of convergence is investigated and shown to be at least of order n^{-1}, where n is the number of past observations at hand.

Key Words and Phrases: Bayes estimator, empirical Bayes, asymptotically optimal, rate of convergence, isotonic regression, symmetric prior.

1. INTRODUCTION

Consider a sequence of N Bernoulli trials. Let p denote the probability of success for each trial and Y stand for the number of successes among the total N trials. Then Y follows a binomial distribution with probability function $f(y|p) = \binom{N}{y} p^y (1-p)^{N-y}, y = 0, 1, \ldots, N$. Suppose that the parameter p is a realization of a random variable P having a prior distribution G. Thus, under the squared error loss, given $Y = y$, the Bayes estimator of p is the posterior mean of P given by

$$\varphi_G(y) = \frac{\int_0^1 pf(x|p)dG(p)}{\int_0^1 f(x|p)dG(p)} = \frac{w(y)}{h(y)} \tag{1.1}$$

where $h(y) = \int_0^1 p^y (1-p)^{N-y} dG(p)$ and $w(y) = \int_0^1 p^{y+1}(1-p)^{N-y} dG(p)$. Also, $f_G(y) = \binom{N}{y} h(y)$ is the marginal probability function of Y. The minimum Bayes risk is $r(G) \equiv r(G, \varphi_G) = E[(\varphi_G(Y) - P)^2]$.

When the prior distribution G is unknown, many authors, based on the past observations, treated this estimation problem via the empirical Bayes approach of Robbins (1956, 1964). For details, the reader is referred to Liang and Huang (1988), Vardeman (1978) and the related references. However, as pointed out by Robbins (1956, 1964), this estimation problem has the feature that estimators which are asymptotically optimal in the usual empirical Bayes sense for all prior distributions do not exist. This is related to the fact that the function $w(y)$ cannot be consistently estimated when the prior distribution G is completely unknown. To remedy this deficiency, Robbins (1956) suggested observing one more trial and proposed an estimator which is asymptotically optimal in a modified sense. Gupta and Liang (1989) treated this estimation problem through the parametric empirical Bayes approach assuming the prior to be a member of beta distribution family with unknown hyperparameters and then using the past observations to estimate the unknown hyperparameters. Liang (1984) and Gupta and Liang (1986) have pointed out that if the unknown prior distribution is symmetric about the point $\frac{1}{2}$, it is possible to construct asymptotically optimal empirical Bayes estimators for the binomial parameter p. However, no estimators were proposed.

In this paper, we deal with this estimation problem through the nonparametric empirical Bayes approach assuming symmetric prior distributions. A monotone empirical Bayes estimator is constructed by using the isotonic regression method. This estimator is asymptotically optimal in the usual empirical Bayes sense for the class of symmetric prior distributions. The corresponding rate of convergence is investigated and shown to be at least of order n^{-1} where n is the number of past observations at hand.

2. CONSTRUCTION OF EMPIRICAL BAYES ESTIMATORS

For each $j = 1, 2, \ldots$, let (Y_j, P_j) be a pair of random variables where Y_j is observable but P_j is not. Conditional on $P_j = p_j, Y_j$ has a binomial distribution $B(N, p_j)$. assumed that $P_j, j = 1, 2, \ldots$, are independently distributed with common unknown prior distribution G. Therefore, $Y_j, j = 1, 2, \ldots$, are iid with marginal probability function $f_G(y)$. Let $\underset{\sim}{Y}_n = (Y_1, \ldots, Y_n)$ denote the n past observations and $Y_{n+1} = Y$ the current random observation. In the empirical Bayes estimation case, an estimation φ_n for the present problem is a function based on a sequence of past observations $\underset{\sim}{Y}_n$ and the present observation $Y = y$. We investigate this estimation problem under the following assumption.

Assumption A: The prior distribution G is symmetric about the point $\frac{1}{2}$, and N is an even number.

Under Assumption A, we have the following lemma which describes the relationship between $w(y)$ and $h(y)$.

Lemma 2.1. Under Assumption A, we have

(a) $w\left(\frac{N}{2}\right) = \frac{1}{2}h\left(\frac{N}{2}\right)$.

(b) $w(x) = w(N-x-1)$ for $x = 0, 1, \ldots, N-1$.

(c) $w(x) + w(N-x) = h(x) = h(N-x), x = 0, 1, \ldots, N$.

(d) $w(x) + w(x+1) = h(x+1), x = 0, 1, \ldots, N-1$.

(e) $\varphi_G(x) = 1 - \varphi_G(N-x), x = 0, 1, \ldots, N$, and

(f) $w(x) = \sum_{i=0}^{x-\frac{N}{2}} h(x-i)(-1)^i - (-1)^{x-\frac{N}{2}} h\left(\frac{N}{2}\right)/2, x = \frac{N}{2}, \ldots, N$.

Proof: Parts (a)-(e) can be obtained directly through straightforward computation by noting the fact that $h(x) = \int_0^1 p^x(1-p)^{N-x}dG(p)$ and $w(x) = \int_0^1 p^{x+1}(1-p)^{N-x}dG(p)$ and by the symmetric property of the prior distribution G about the point 1/2, which is guaranteed under Assumption A. Part (f) is a consequence of parts (a) and (d). □

For each $y = 0, 1, \ldots, N$, define

$$f_n(y) = f_n(N-y) = \begin{cases} \frac{1}{2n} \sum_{j=1}^{n} I_{\{y, N-y\}}(Y_j) & \text{if } y \neq \frac{N}{2}, \\ \frac{1}{n} \sum_{j=1}^{n} I_{\{y\}}(Y_j) & \text{if } y = \frac{N}{2}. \end{cases}$$

$$h_n(y) = f_n(y) \Big/ \binom{N}{y}.$$

Also, for each $y = N/2, \ldots, N$, define

$$w_n(y) = \sum_{i=0}^{y-\frac{N}{2}} h_n(y-i)(-1)^i - (-1)^{y-\frac{N}{2}} h_n\left(\frac{N}{2}\right)/2,$$

and for each $y = 0, \ldots, N/2-1$, let

$$w_n(y) = h_n(y) - w_n(N-y).$$

Both $h_n(y)$ and $w_n(y)$ are unbiased estimators of $h(y)$ and $w(y)$, respectively, $y = 0, 1, \ldots, N$. Thus, it is intuitive to use $\frac{w_n(y)}{h_n(y)}$ as an estimator for $\varphi_G(y) \equiv \frac{w(y)}{h(y)}$. However, this naive estimator may have serious deficiencies. First, $h_n(y)$ may be equal to zero and thus, the function $\frac{w_n(y)}{h_n(y)}$ is not well defined. Second, it is possible that the value of $\frac{w_n(y)}{h_n(y)}$ may be greater than 1 or less than 0, while $0 \le \varphi_G(y) \le 1$ for all $y = 0, 1, \ldots, N$. Third, the estimator $\frac{w_n(y)}{h_n(y)}$ may not possess the increasing property as $\varphi_G(y)$ does. Hence, in the following, we seek a smoothed estimator. The following lemma states the monotone properties of the functions $\varphi_G(y), h(y)$ and $w(y)$. These properties may suggest a way how to construct reasonable empirical Bayes estimators. We note that, in the literature, Van Houwelingen (1977) has studied a method to monotonize empirical Bayes estimators for discrete exponential family distributions. Also, Stijnen (1982) and, Van Houwelingen and Stijnen (1983) have studied monotone empirical Bayes estimators for the continuous one-parameter exponential family distributions.

Lemma 2.2. (a) For any prior distribution $G, \varphi_G(y)$ is an increasing function of $y, y = 0, 1, \ldots, N$.

(b) Under Assumption A, both $h(y)$ and $w(y)$ are increasing in y for $y = \frac{N}{2}, \ldots, N$.

Based on Lemmas 2.1 and 2.2, we construct a monotone empirical Bayes estimator as follows. Let $\{\tilde{h}_n(x)\}_{x=\frac{N}{2}}^{N}$ be the isotonic regression of $\{h_n(x)\}_{x=\frac{N}{2}}^{N}$ with equal weights. Define $\tilde{w}_n(x) = \sum_{i=0}^{x-\frac{N}{2}} \tilde{h}_n(x-i)(-1)^i - (-1)^{x-\frac{N}{2}} \tilde{h}_n\left(\frac{N}{2}\right)/2$, for $\frac{N}{2} \le x \le N$. Thus, $\tilde{h}_n(x)$ is nondecreasing in x for $\frac{N}{2} \le x \le N$, and by this nondecreasing property, $\tilde{w}_n(x) \ge 0$ for $\frac{N}{2} \le x \le N$. However, $\tilde{w}_n(x)$ may still not possess the nondecreasing property. Thus, we let $\{w_n^*(x)\}_{x=\frac{N}{2}}^{N}$ be the isotonic regression of $\{\tilde{w}_n(x)\}_{x=\frac{N}{2}}^{N}$ with equal weights. Then define $h_n^*(x) = w_n^*(x-1) + w_n^*(x)$ for $\frac{N}{2} + 1 \le x \le N$ and $h_n^*\left(\frac{N}{2}\right) = 2w_n^*\left(\frac{N}{2}\right)$. By the nondecreasing property of $w_n^*(x), h_n^*(x)$ is nondecreasing in x for $\frac{N}{2} \le x \le N$. Now, for

$\frac{N}{2} \le x \le N$, define

$$\varphi_n(x) = \begin{cases} \frac{w_n^*(x)}{h_n^*(x)} & \text{if } h_n^*(x) \neq 0, \\ \frac{1}{2} & \text{if } h_n^*(x) = 0. \end{cases}$$

Since $\varphi_n(x)$ may not be a nondecreasing function of x for $\frac{N}{2} \le x \le N$, we consider the isotonic regression $\{\varphi_n^*(x)\}_{x=\frac{N}{2}}^{N}$ of $\{\varphi_n(x)\}_{x=\frac{N}{2}}^{N}$ with equal weights. Also, for $0 \le x \le \frac{N}{2} - 1$, define $\varphi_n^*(x) = 1 - \varphi_n^*(N-x)$. Now one can see that $\varphi_n^*(x)$ is nondecreasing in x for $x = 0, 1, \ldots, N$. We propose using $\varphi_n^*(x)$ as an estimator of $\varphi_G(x), x = 0, 1, \ldots, N$.

Remark 2.1. By the nondecreasing property of $w_n^*(x)$ on $x, x = \frac{N}{2}, \ldots, N, \varphi_n(x) \ge \frac{1}{2}$ for all $x \ge \frac{N}{2}$ and hence, $\varphi_n^*(x) \ge \frac{1}{2}$ for $x \ge \frac{N}{2}$. Also, $h_n^*(x) = 0$ iff $w_n^*(x) = 0$ iff $\tilde{w}_n(y) = 0$ for all $\frac{N}{2} \le y \le x$ iff $\tilde{h}_n(y) = 0$ for all $\frac{N}{2} \le y \le x$ iff $h_n(y) = 0$ for all $y = N - x, \ldots, x$, where $x \ge \frac{N}{2}$.

3. ASYMPTOTIC OPTIMALITY

Let $\psi_n(y)$ denote an empirical Bayes estimator based on the current observation y and the past data $\underset{\sim}{Y}_n = (Y_1, \ldots, Y_n)$. Let $r(G, \psi_n)$ denote the conditional Bayes risk (conditional on $\underset{\sim}{Y}_n$) of the estimator ψ_n and $Er(G, \psi_n)$ the associated overall Bayes risk where the expectation E is taken with respect to $\underset{\sim}{Y}_n$. Since $r(G)$ is the minimum Bayes risk, $r(G, \psi_n) - r(G) \ge 0$ and therefore $Er(G, \psi_n) - r(G) \ge 0$. The nonnegative difference $Er(G, \psi_n) - r(G)$ is often used as a measure of the optimality of the empirical Bayes estimator ψ_n.

Definition 3.1. A sequence of empirical Bayes estimators $\{\psi_n\}_{n=1}^{\infty}$ is said to be asymptotically optimal in E at least of order β_n relative to the prior distribution G if $Er(G, \psi_n) - r(G) \le O(\beta_n)$ where $\{\beta_n\}_{n=1}^{\infty}$ is a sequence of positive numbers such that $\lim\limits_{n \to \infty} \beta_n = 0$.

The usefulness of empirical Bayes estimators in applications clearly depend on the convergence rates for which the risks of the successive estimators approach the minimum

Bayes risk. In the following, the performance of the proposed empirical Bayes estimators $\{\varphi_n^*\}$ is evaluated on basis of the rates of convergence of the nonnegative difference $Er(G,\varphi_n^*) - r(G)$. In the following, all the computations are made under Assumption A.

Lemma 3.1. For each $y = \frac{N}{2}+1,\ldots,N$, and for $0 < t < \min(1-\varphi_G(y), \varphi_G(y)-\frac{1}{2})$, if $h_n^*(y) > 0$ and $|\varphi_n^*(y) - \varphi_G(y)| > t$, then

$$(N+2)\sum_{x=\frac{N}{2}}^{N} |h_n(x) - h(x)|^2 > [h(N/2)t]^2 .$$

Proof: $|\varphi_n^*(y) - \varphi_G(y)| > t \Rightarrow \varphi_n^*(y) - \varphi_G(y) > t$ or $\varphi_n^*(y) - \varphi_G(y) < -t$. By the definition of $\varphi_n^*(y), w_n^*(x), \tilde{w}_n(x), \tilde{h}_n(x)$, Lemma 2.1 and Theorem 2.1 of Barlow, et. al. (1972), we have:

$$\begin{aligned}
& \varphi_n^*(y) - \varphi_G(y) > t \\
\Rightarrow\ & \varphi_n(x) - \varphi_G(y) > t \quad \text{for some } N/2+1 \le x \le y. \\
\Rightarrow\ & w_n^*(x)[1-\varphi_G(y)-t] - w_n^*(x-1)[\varphi_G(y)+t] > 0 \text{ for some } N/2+1 \le x \le y \\
\Rightarrow\ & [w_n^*(x) - w(x)][1-\varphi_G(y)-t] - [w_n^*(x-1) - w(x-1)][\varphi_G(y)+t] > h(N/2)t \\
& \qquad \text{for some } N/2+1 \le x \le y \\
\Rightarrow\ & [w_n^*(x) - w(x)] > h(N/2)t \text{ or } [w_n^*(x-1) - w(x-1)] < -h(N/2)t \\
& \qquad \text{for some } N/2+1 \le x \le y \\
\Rightarrow\ & \sup_{\frac{N}{2} \le x \le N} |w_n^*(x) - w(x)| > h(N/2)t \\
\Rightarrow\ & \sum_{x=\frac{N}{2}}^{N} |w_n^*(x) - w(x)|^2 > [h(N/2)t]^2 \\
\Rightarrow\ & \sum_{x=\frac{N}{2}}^{N} |\tilde{w}_n(x) - w(x)|^2 > [h(N/2)t]^2,
\end{aligned}$$

since $\sum_{x=\frac{N}{2}}^{N} [w_n^*(x) - w(x)]^2 \le \sum_{x=\frac{N}{2}}^{N} (\tilde{w}_n(x) - w(x))^2$, see Theorem 2.1 of Barlow, et. al. (1972).

Now, by the definition of $\tilde{w}_n(x)$, we have, for each $x = \frac{N}{2}, \ldots, N$,

$$\begin{aligned}
&[\tilde{w}_n(x) - w(x)]^2 \\
&= \left[\sum_{i=0}^{x-\frac{N}{2}} [\tilde{h}_n(x-i) - h(x-i)](-1)^i - (-1)^{x-\frac{N}{2}} [\tilde{h}_n(N/2) - h(N/2)]/2 \right]^2 \\
&\leq 2 \sum_{x=\frac{N}{2}}^{N} [\tilde{h}_n(x) - h(x)]^2 \\
&\leq 2 \sum_{x=\frac{N}{2}}^{N} [h_n(x) - h(x)]^2
\end{aligned}$$

where the last inequality is again from Theorem 2.1 of Barlow, et. al. (1972).

Based on the above discussions, we conclude that

$$\varphi_n^*(y) - \varphi_G(y) > t \Rightarrow (N+2) \sum_{x=\frac{N}{2}}^{N} [h_n(x) - h(x)]^2 > [h(N/2)t]^2 . \tag{3.1}$$

Analogous to the preceding discussion, we can obtain:

$$\varphi_n^*(y) - \varphi_G(y) < -t \text{ and } h_n^*(y) > 0 \Rightarrow (N+2) \sum_{z=\frac{N}{2}}^{N} [h_n(z) - h(z)]^2 > [h(N/2)t]^2 . \tag{3.2}$$

Therefore, (3.1) and (3.2) together lead to the result of the lemma. □

Remark 3.1. Let $\mathcal{Y}_n$ be the sample space generated by $\underset{\sim}{Y}_n$. Then for each $y = \frac{N}{2}+1, \ldots, N$, as $t > 1 - \varphi_G(y)$, $\{\underset{\sim}{Y}_n \in \mathcal{Y}_n | \varphi_n^*(y) - \varphi_G(y) > t\} = \phi$; also, as $t > \varphi_G(y) - \frac{1}{2}$, $\{\underset{\sim}{Y}_n \in \mathcal{Y}_n | \varphi_n^*(y) - \varphi_G(y) < -t\} = \phi$.

Lemma 3.2. For each $y = \frac{N}{2} + 1, \ldots, N$ and $t > 0$,

$$P\{|\varphi_n^*(y) - \varphi_G(y)| > t \text{ and } h_n^*(y) > 0\} \leq \sum_{x=\frac{N}{2}}^{N} 2e^{-\frac{4nh^2(\frac{N}{2})\binom{N}{x}^2 t^2}{(N+2)^2}} .$$

Proof: By Remark 3.1, $P\{|\varphi_n^*(y) - \varphi_G(y)| > t, h_n^*(y) > 0\} = 0$ if $t \geq \max(1 - \varphi_G(y),$

$\varphi_G(y) - \frac{1}{2})$. Thus, as $0 < t < \max(1 - \varphi_G(y), \varphi_G(y) - \frac{1}{2})$, from Lemma 3.1,

$$\begin{aligned}
& P\{|\varphi_n^*(y) - \varphi_G(y)| > t, h_n^*(y) > 0\} \\
\leq\ & P\left\{ \sum_{z=\frac{N}{2}}^{N} [h_n(z) - h(z)]^2 > \frac{[th\left(\frac{N}{2}\right)]^2}{N+2} \right\} \\
\leq\ & \sum_{z=\frac{N}{2}}^{N} P\left\{ |h_n(z) - h(z)| > \frac{\sqrt{2}th\left(\frac{N}{2}\right)}{N+2} \right\} \\
=\ & \sum_{z=\frac{N}{2}}^{N} P\left\{ |f_n(z) - f_G(z)| > \frac{\sqrt{2}th\left(\frac{N}{2}\right)\binom{N}{z}}{N+2} \right\} \\
\leq\ & \sum_{z=\frac{N}{2}}^{N} 2e^{-\frac{4nh^2\left(\frac{N}{2}\right)\binom{N}{z}^2 t^2}{(N+2)^2}}
\end{aligned}$$

where the last inequality is obtained from Theorem 1 of Hoeffding (1963). □

The following theorem is our main result.

Theorem 3.1. Let $\{\varphi_n^*\}_{n=1}^{\infty}$ be the sequence of empirical Bayes estimators constructed in Section 2. Then, under Assumption A,

$$Er(G, \varphi_n^*) - r(G) \leq O(n^{-1}).$$

Proof: First, we consider the case where $G(0) < \frac{1}{2}$. Straightforward computation leads to the following.

$$\begin{aligned}
& 0 \leq Er(G, \varphi_n^*) - r(G) \\
& = \sum_{y=0}^{N} E[(\varphi_n^*(y) - \varphi_G(y))^2] f_G(y) \\
& = 2 \sum_{y=\frac{N}{2}+1}^{N} E[(\varphi_n^*(y) - \varphi_G(y))^2] f_G(y).
\end{aligned} \tag{3.3}$$

For each $y = \frac{N}{2} + 1, \ldots, N$,

$$\begin{aligned}
& E[(\varphi_n^*(y) - \varphi_G(y))^2] \\
= & \int_0^{\max(1-\varphi_G(y), \varphi_G(y)-\frac{1}{2})} 2tP\{|\varphi_n^*(y) - \varphi_G(y)| > t, h_n^*(y) > 0\} dt \\
& + (\varphi_G(y) - 1/2)^2 P\{h_n^*(y) = 0\}.
\end{aligned}$$

Now, from Remark 2.1,

$$
\begin{aligned}
P\{h_n^*(y)=0\} &= P\{f_n(x)=0 \text{ for all } x=N-y,\dots,y\} \\
&= [1-F_G(y)+F_G(N-y-1)]^n \\
&= \exp(-nln(1-F_G(y)+F_G(N-y-1))^{-1}) \\
&\le O(n^{-1}).
\end{aligned} \tag{3.4}
$$

where $F_G(\cdot)$ is the marginal distribution function of Y. Also, from Lemma 3.2, and the fact that $\max(1-\varphi_G(y),\varphi_G(y)-\frac{1}{2})<\frac{1}{2}$ for $y\ge\frac{N}{2}+1$, we have

$$
\begin{aligned}
&\int_0^{\max(1-\varphi_G(y),\varphi_G(y)-\frac{1}{2})} 2tP\{|\varphi_n^*(y)-\varphi_G(y)|>t, h_n^*(y)>0\}dt. \\
\le &\int_0^{\frac{1}{2}} 4t \sum_{x=\frac{N}{2}}^{N} e^{-\frac{4nh^2(\frac{N}{2})\binom{N}{x}^2 t^2}{(N+2)^2}} dt \\
\le &\ \frac{1}{n}\sum_{x=\frac{N}{2}}^{N} \frac{(N+2)^2}{2h^2\left(\frac{N}{2}\right)\binom{N}{x}^2} \\
= &\ O(n^{-1}).
\end{aligned} \tag{3.5}
$$

From (3.4) and (3.5), we conclude that for each $y=\frac{N}{2}+1,\dots,N$,

$$
E[(\varphi_n^*(y)-\varphi_G(y))^2]\le O(n^{-1}). \tag{3.6}
$$

Since N is finite and fixed, (3.6) and (3.3) together complete the proof of the theorem.

Next, we consider the case where $G(0)=\frac{1}{2}$. Under this case, by the symmetric property of G, one can show that $f_G(y)=0$ for all $y=1,\dots,N-1$. Thus, (3.3) can be reduced to

$$
\begin{aligned}
0 &\le Er(G,\varphi_n^*)-r(G) \\
&= 2E[(\varphi_n^*(N)-\varphi_G(N))^2]f_G(N),
\end{aligned} \tag{3.7}
$$

where $\varphi_G(N)=1$, which can be verified by straight computation.

Also,

$$P\{Y_j \in \{0, N\} \text{ for all } j = 1, \ldots, n\} = 1. \tag{3.8}$$

Thus, with probability one, we have: $h_n(N) = h_n(0) = 1$ and $h_n(y) = 0, 1 \leq y \leq N - 1$. Following the way to construct the isotonic estimators $\varphi_n^*(y), y = 0, 1, \ldots, N$, one can see that $\varphi_n^*(N) = 1$. This fact and (3.7) together imply $Er(G, \varphi_n^*) - r(G) = 0$, which concludes the theorem. $\square$

ACKNOWLEDGEMENT

This research was supported in part by the Office of Naval Research Contract N00014–88–K–0170 at Purdue University. The author is grateful to a referee for helpful comments and suggestions to improve the presentation of this article.

BIBLIOGRAPHY

Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. (1972). *Statistical Inference Under Order Restrictions.* Wiley, New York.

Gupta, S.S. and Liang, T. (1986). Empirical Bayes rules for selecting good binomial populations. *Adaptive Statistical Procedures and Related Topics* (Ed. J. Van Ryzin), IMS Lecture Notes–Monograph Series, *Vol. 8*, 110–128.

Gupta, S.S. and Liang, T. (1989). Selecting the best binomial population: parametric empirical Bayes approach. *J. Statist. Plann. Inference*, **23**, 21–31.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, **58**, 13–30.

Liang, T. (1984). Some contributions to empirical Bayes, sequential and locally optimal subset selection rules. Ph.D. Dissertation, Department of Statistics, Purdue University, West Lafayette, Indiana.

Liang, T. and Huang, W.-T. (1988). On a monotone empirical Bayes estimator of binomial parameter. Submitted for publication.

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

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

Stijnen, Th. (1982). A monotone empirical Bayes estimator and test for the one-parameter continuous exponential family based on spacings. *Scand. J. Statist.*, **9**, 153–158.

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

Van Houwelingen, J.C. and Stijnen, Th. (1983). Monotone empirical Bayes estimators for the continuous one-parameter exponential family. *Statist. Neerl.*, **37**, 29–43.

Vardeman, S.B. (1978). Bounds on the empirical Bayes and compound risks of truncated versions of Robbins's estimator of a binomial parameter. *J. Statist. Plann. Inference*, **2**, 245–252.