Journal of Statistical Planning and Inference 34 (1993) 333-348
North-Holland

On monotone empirical Bayes estimators of a binomial parameter

TaChen Liang

Department of Mathematics, Wayne State University, Detroit, MI 48202, USA

Wen-Tao Huang

Institute of Statistical Sciences, Academia Sinica, Nankang, Taipei, Taiwan, ROC

Received 27 March 1989; revised manuscript received 30 March 1992
Recommended by S. Panchapakesan

Abstract: Two monotone empirical Bayes estimators for the binomial parameters are considered: One using the monotonizing method of Van Houwelingen (*Statistica Neerlandica* **31**, 1977), and the other using the isotonic regression method. The corresponding asymptotic optimality is investigated. It is proved that for each of them, the associated rate of convergence is of order n^{-1} where n is the number of past observations at hand. Improved empirical Bayes estimators are obtained by Rao–Blackwellizing the two monotone empirical Bayes estimators. The small sample performance of the proposed empirical Bayes estimators as well as some other known empirical Bayes estimators is investigated using Monte Carlo simulation. The performance of the proposed empirical Bayes estimators is much better than that of the others, especially when n is small.

AMS Subject Classification: Primary 62C12, Secondary 62F15.

Key words and phrases: Empirical Bayes; asymptotically optimal; isotonic regression; rate of convergence; Rao–Blackwellization.

1. Introduction

Consider a sequence of M Bernoulli trials. Let p denote the probability of success for each trial and Y stand for the number of successes among the total M trials. Suppose that the parameter p is a realization of a random variable P having a prior distribution G. Under the squared error loss, given $Y=y$, the Bayes estimator of p is the posterior mean of P denoted by

$$\varphi_{GM}(y) = E[P \mid Y=y] = \frac{\omega_{GM}(y)}{f_{GM}(y)}, \tag{1.1}$$

Correspondence to: Prof. TaChen Liang, Dept. of Mathematics, Wayne State University, Detroit MI 48202, USA.

where

$$f_{GM}(y) = \int_0^1 f_M(y \mid p)\,\mathrm{d}G(p),$$

$$\omega_{GM}(y) = \int_0^1 pf_M(y \mid p)\,\mathrm{d}G(p),$$

and

$$f_M(y \mid p) = \binom{M}{y} p^y (1-p)^{M-y}.$$

The minimum Bayes risk is given by

$$R_M(G) \equiv R_M(G, \varphi_{GM}) = E[(\varphi_{GM}(Y) - P)^2], \tag{1.2}$$

where the expectation E is taken with respect to both Y and P.

When the prior distribution G is unknown, many authors treated this estimation problem via the empirical Bayes approach of Robbins (1956, 1964, 1983). For details, the reader is referred to Berry and Christensen (1979), Griffin and Krutchkoff (1971), Gupta and Liang (1988), Gutman (1982), Liang (1989), Martz and Lian (1974) and Vardeman (1978), among others. As pointed out by Robbins (1956, 1964), this estimation problem has the interesting feature that estimators which are asymptotically optimal in the usual empirical Bayes sense do not exist. This is due to the fact that the function $\omega_{GM}(y)$ can not be consistently estimated when the prior distribution G is completely unknown. To remedy this deficiency, Robbins (1956) suggested observing one more Bernoulli trial at each stage, and proposed an estimator which is asymptotically optimal in a modified sense (i.e., asymptotically optimal with respect to $R_M(G)$ instead of $R_{M+1}(G)$). Vardeman (1978) has studied two variants of the Robbins' estimator and has shown that the Bayes risks (under squared error loss) of these modified estimators converge to the minimum Bayes risk $R_M(G)$ at least with rate of order $n^{-1/2}$. Later Liang (1989) improved the Vardeman result by providing a lower bound and an upper bound for the rate of convergence. These bounds are of order n^{-1} and $n^{-1}\log n$, respectively. Gupta and Liang (1988) have also investigated two monotone empirical Bayes estimators, which are based on Vardeman's (1978) and Liang's (1989) original empirical Bayes estimators, and proved that for each of them, the corresponding rate of convergence is of order n^{-1}. Though this rate is the best rate one can obtain for this estimation problem, one is also interested in the small sample performance of the empirical Bayes estimators.

In this paper, we deal with this estimation problem through the nonparametric empirical Bayes approach. Two monotone empirical Bayes estimators are constructed: one using the monotonizing method of Van Houwelingen (1977), and the other using the isotonic regression method. The corresponding asymptotic optimality is investigated. For each of them it is found that the associated rate of convergence is of order n^{-1}. Improved empirical Bayes estimators are obtained by using the Rao–Blackwellizing method. These improved empirical Bayes estimators are better than the original one in the sense that the improved empirical Bayes estimators have smaller Bayes risks. Monte Carlo simulation is used to study the small sample performance of the proposed empirical Bayes estimators and some other known empirical Bayes estimators. The simulation results indicate that the proposed empirical Bayes estimators perform much better than the other known empirical Bayes estimators, especially when the number of past observations n, is small.

2. Monotone empirical Bayes estimators

In this section, a modified empirical Bayes framework of Robbins (1956, 1964) is adopted. Following his suggestion, an experimenter observes one more Bernoulli trial at each stage. For each $i=1,2,\ldots$, let $X_i=(X_{i,1},\ldots,X_{i,M+1})$ denote $M+1$ independent Bernoulli variables with probability of success p_i at stage i. The parameter p_i is a realization of a random variable P_i. It is assumed that the random variables $P_1, P_2, \ldots$ are independently distributed with a common unknown prior distribution G. Under this model, $X_1, X_2, \ldots,$ are i.i.d. Suppose that now we are at stage $n+1$. Thus, we have n past observations $X_1,\ldots,X_n$ and the present observation X_{n+1}. Our goal is to construct empirical Bayes estimator for the current random parameter P_{n+1}.

Before we go further to construct the empirical Bayes estimators for the estimation problem under study, we first look at certain properties associated with this estimation problem.

Note that the class of binomial probability function $\{f_M(y \mid p) \mid 0<p<1\}$ has monotone likelihood ratio in y and therefore, the Bayes estimator $\varphi_{GM}(y)$ is an increasing function of y. Recall that under the squared error loss, all the monotone estimators form an essentially complete class, see Berger (1985). Hence, monotonicity is a desirable property for an empirical Bayes estimator. In the literature, Van Houwelingen (1977) has studied a method to monotonize empirical Bayes estimators for the discrete exponential family. A resulting monotonized empirical Bayes estimator has been shown to be as good as the original one in terms of Bayes risks. However, the performance of the monotonized empirical Bayes estimator is strongly dependent on that of the original empirical Bayes estimator. Hence, to apply his monotonizing method, it is important to find a 'good' initial empirical Bayes estimator.

For each $i=1,2,\ldots,n+1$, let $Y_{i,m}=\sum_{j=1,j\neq m}^{M+1} X_{ij}$, $m=1,\ldots,M+1$. Then, conditional on $P_i=p_i$, $Y_{i,m}$ and $X_{i,m}$ are independent, and are, respectively, $B(M,p_i)$ and $B(1,p_i)$.

For each $y=0,1,\ldots,M$, let

$$f_{Mn}(y)=\frac{1}{n(M+1)}\sum_{i=1}^{n}\sum_{j=1}^{M+1} I_{\{y\}}(Y_{i,j}), \tag{2.1}$$

$$\omega_{Mn}(y)=\frac{1}{n(M+1)}\sum_{i=1}^{n}\sum_{j=1}^{M+1} X_{i,j}I_{\{y\}}(Y_{i,j}),$$

where $I_A(\cdot)$ denotes the indicator function of the set A.

Note that $E[f_{Mn}(y)]=f_{GM}(y)$ and $E[\omega_{Mn}(y)]=\omega_{GM}(y)$. Since $X_1,\ldots,X_n$ are i.i.d., by the strong law of large numbers, $f_{Mn}(y)$ and $\omega_{Mn}(y)$ are consistent estimators of $f_{GM}(y)$ and $\omega_{GM}(y)$, respectively, for each $y=0,1,\ldots,M$. To estimate $\varphi_{GM}(y)$ defined in (1.1), it is thus reasonable to use the consistent estimator $\varphi_{Mn}(y)$ defined by

$$\varphi_{Mn}(y)=\frac{\omega_{Mn}(y)}{f_{Mn}(y)}, \tag{2.2}$$

where $0/0\equiv 0$. Note that $0\leqslant\varphi_{Mn}(y)\leqslant 1$, $y=0,1,\ldots,M$.

However, the function $\varphi_{Mn}(y)$ may not possess the monotonicity property. Based on the previous reasoning, it is important to consider monotone empirical Bayes estimators. Two monotone empirical Bayes estimators are constructed based on the initial empirical Bayes estimator φ_{Mn}. One is the monotone empirical Bayes estimator which is obtained by monotonizing the empirical Bayes estimator φ_{Mn} by the monotonizing method of Van Houwelingen (1977). We denote this monotonized version of φ_{Mn} by φ^*_{Mn}. Since there is no closed form for φ^*_{Mn}, the reader is referred to Van Houwelingen (1977) for details. We propose using $\varphi^*_{Mn}(Y_{n+1,M+1})$ to estimate P_{n+1}. This monotone empirical Bayes estimator $\varphi^*_{Mn}(Y_{n+1,M+1})$ has the following nice property:

$$E[(\varphi^*_{Mn}(Y_{n+1,M+1})-P_{n+1})^2]\leqslant E[(\varphi_{Mn}(Y_{n+1,M+1})-P_{n+1})^2]. \tag{2.3}$$

For details, see Van Houwelingen (1977).

The other monotone empirical Bayes estimator is constructed using the isotonic regression method. For each $y=0,1,\ldots,M$, define

$$\tilde{\varphi}_{Mn}(y)=\max_{0\leqslant s\leqslant y}\ \min_{s\leqslant t\leqslant M}\left\{\sum_{x=s}^{t}\omega_{Mn}(x)\Big/\sum_{x=s}^{t}f_{Mn}(x)\right\}. \tag{2.4}$$

It is easily seen that $\tilde{\varphi}_{Mn}(y)$ is nondecreasing in y. We propose using $\tilde{\varphi}_{Mn}(y)$ to estimate $\varphi_{GM}(y)$, and suggest estimating P_{n+1} by $\tilde{\varphi}_{Mn}(Y_{n+1,M+1})$.

Remark 2.1. (a) Note that $E[P_{n+1}]=\sum_{y=0}^{M}\varphi_{GM}(y)f_{GM}(y)$. The expectation of the random parameter P_{n+1} is thus a weighted sum of the posterior means $\varphi_{GM}(y)$ with weights $f_{GM}(y)$, $y=0,1,\ldots,M$. Hence, $\tilde{\varphi}_{Mn}(y)$, $y=0,\ldots,M$, can be viewed as the isotonic regression estimators of $\varphi_{GM}(y)$, $y=0,1,\ldots,M$, based on the naive estimators $\varphi_{Mn}(y)$, $y=0,1,\ldots,M$, with random weight functions $f_{Mn}(y)$, $y=0,1,\ldots,M$.

(b) In the literature, Robbins (1956), Vardeman (1978) and Liang (1989) have proposed some nonparametric empirical Bayes estimators for $\varphi_{GM}(y)$. These estimators are consistent in the sense that they converge to $\varphi_{GM}(y)$ in probability. However, none of them possesses the monotonicity property. Gupta and Liang (1988) have constructed two monotone empirical Bayes estimators which are based on Vardeman's and Liang's original estimators using the isotonic regression method with equal weights.

(c) All the empirical Bayes estimators, proposed by those authors just mentioned in (b), use only the information contained in $(Y_{i,M+1}, X_{i,M+1})$, $i=1,\dots,n$, to estimate $f_{GM}(y)$ and $\omega_{GM}(y)$. However, the proposed empirical Bayes estimators φ^*_{Mn} and $\tilde{\varphi}_{Mn}$ use all the information of $(Y_{i,m}, X_{i,m})$, $m=1,\dots,M+1$, $i=1,\dots,n$.

3. Asymptotic optimality

For an empirical Bayes estimator $\psi_{Mn}(Y_{n+1,M+1})$ of P_{n+1}, let $R_M(G,\psi_{Mn})$ denote its associated conditional Bayes risk (conditional on the past observations $X_1,\dots,X_n$) and $ER_M(G,\psi_{Mn})$ the associated overall Bayes risk, where the expectation E is taken with respect to $(X_1,\dots,X_n)$. Since $R_M(G)$ is the minimum Bayes risk, $R_M(G,\psi_{Mn})-R_M(G)\geqslant 0$ for all $(X_1,\dots,X_n)$ and for all n, therefore $ER_M(G,\psi_{Mn})-R_M(G)\geqslant 0$ for all n. The nonnegative difference $ER_M(G,\psi_{Mn})-R_M(G)$ is often used as a measure of optimality of the empirical Bayes estimator $\psi_{Mn}(Y_{n+1,M+1})$.

Definition 3.1. A sequence of empirical Bayes estimators $\{\psi_{Mn}\}$ is said to be asymptotically optimal in E of order β_n relative to $R_M(G)$ if $ER_M(G,\psi_{Mn})-R_M(G)=O(\beta_n)$, where $\{\beta_n\}$ is a sequence of positive numbers such that $\lim_{n\to\infty}\beta_n=0$.

The usefulness of empirical Bayes estimators in practical applications clearly depends on the convergence rates at which the risks of the successive estimators approach the minimum Bayes risk. In the following, we evaluate the performance of the two sequences of empirical Bayes estimators $\{\tilde{\varphi}_{Mn}\}$ and $\{\varphi^*_{Mn}\}$ on basis of the rates of convergence.

For the empirical Bayes estimator $\tilde{\varphi}_{Mn}(Y_{n+1,M+1})$, straightforward computation leads to the following:

$$
\begin{aligned}
0&\leqslant ER_M(G,\tilde{\varphi}_{Mn})-R_M(G)\\
&=E[(\tilde{\varphi}_{Mn}(Y_{n+1,M+1})-\varphi_{GM}(Y_{n+1,M+1}))^2]\\
&=\sum_{y=0}^{M} E[(\tilde{\varphi}_{Mn}(y)-\varphi_{GM}(y))^2]f_{GM}(y)\\
&=\sum_{y\in A} E[(\tilde{\varphi}_{Mn}(y)-\varphi_{GM}(y))^2]f_{GM}(y),
\end{aligned}
\tag{3.1}
$$

where $A=\{y\mid f_{GM}(y)>0,\ y=0,1,\dots,M\}$. The case $A=\{0,1,\dots,M\}$ will be considered first, followed by the case where $A\neq\{0,1,\dots,M\}$. First, as $A=\{0,1,\dots,M\}$, we let $c_1=\min_{0\leqslant y\leqslant M} f_{GM}(y)$. Note that $c_1>0$.

Lemma 3.1. *$P\{f_{Mn}(x)=0\}\leqslant\exp\{-2nc_1^2\}$ for all $x=0,1,\dots,M$.*

Proof.

$$\begin{aligned}P\{f_{Mn}(x)=0\} &= P\{f_{Mn}(x)-f_{GM}(x)=-f_{GM}(x)\}\\ &\leqslant P\{f_{Mn}(x)-f_{GM}(x)\leqslant -c_1\}\\ &\leqslant \exp\{-2nc_1^2\},\end{aligned}$$

where the second inequality follows from Theorem 2 of Hoeffding (1963). □

Lemma 3.2. (a) *For $0\leqslant y\leqslant x\leqslant M$ and for each $t\in(0,\varphi_{GM}(y))$,*

$$P\{\varphi_{Mn}(x)-\varphi_{GM}(y)\leqslant -t,\ f_{Mn}(x)>0\}\leqslant \exp\{-2nt^2c_1^2\}.$$

(b) *For $0\leqslant x\leqslant y\leqslant M$ and for each $t\in(0,1-\varphi_{GM}(y))$,*

$$P\{\varphi_{Mn}(x)-\varphi_{GM}(y)>t\}\leqslant \exp\{-2nt^2c_1^2\}.$$

Proof. We prove part (a) only. By the definition of $\varphi_{Mn}(x)$, following direct computation, we obtain

$$\begin{aligned}&P\{\varphi_{Mn}(x)-\varphi_{GM}(y)>-t,\ f_{Mn}(x)>0\}\\ &\quad\leqslant P\{\omega_{Mn}(x)-f_{Mn}(x)[\varphi_{GM}(y)-t]<0\}\\ &\quad\leqslant P\{\omega_{Mn}(x)-f_{Mn}(x)[\varphi_{GM}(y)-t]-\omega_{GM}(x)\\ &\qquad\quad +f_{GM}(x)[\varphi_{GM}(y)-t]<-tf_{GM}(x)\}\\ &\quad= P\left\{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{M+1}[X_{ij}-\varphi_{GM}(y)+t]I_{\{x\}}(Y_{i,j})/(M+1)<-tf_{GM}(x)\right\}\\ &\quad\leqslant \exp\{-2nt^2c_1^2\},\end{aligned}$$

where the last inequality follows from Theorem 2 of Hoeffding (1963) and the definition of c_1. □

Lemma 3.3. (a) *For $0\leqslant y\leqslant x\leqslant M$,*

$$\int_0^{\varphi_{GM}(y)} tP\{\varphi_{Mn}(x)-\varphi_{GM}(y)<-t\}\,\mathrm{d}t=\mathrm{O}(n^{-1}).$$

(b) *For $0\leqslant x\leqslant y\leqslant M$,*

$$\int_0^{1-\varphi_{GM}(y)} tP\{\varphi_{Mn}(x)-\varphi_{GM}(y)>t\}\,\mathrm{d}t=\mathrm{O}(n^{-1}).$$

Proof. (a) Note that for $0\leqslant y\leqslant x\leqslant M$ and for each $t\in(0,\varphi_{GM}(y))$,

$$\begin{aligned}&P\{\varphi_{Mn}(x)-\varphi_{GM}(y)<-t\}\\ &\quad\leqslant P\{f_{Mn}(x)=0\}+P\{\varphi_{Mn}(x)-\varphi_{GM}(y)<-t,\ f_{Mn}(x)>0\}.\end{aligned}$$

Then, by Lemmas 3.1 and 3.2 (a), after taking integration over the domain of t, the result follows immediately.

(b) Similarly, by Lemma 3.2 (b) and taking integration, then the result follows. □

Theorem 3.1. *Let $\{\tilde{\varphi}_{Mn}\}$ be the sequence of the empirical Bayes estimators defined in (2.4). Then, $ER_M(G,\tilde{\varphi}_{Mn}) - R_M(G) = O(n^{-1})$.*

Proof. First, we consider the case where $A = \{0, 1, \dots, M\}$. By the definition of $\tilde{\varphi}_{Mn}(y)$, straightforward computation leads to the following: For each $y=0, 1, \dots, M$,

$$\begin{aligned}
&E[(\tilde{\varphi}_{Mn}(y) - \varphi_{GM}(y))^2] \\
&\quad = \int_0^{\varphi_{GM}(y)} 2tP\{\tilde{\varphi}_{Mn}(y) - \varphi_{GM}(y) < -t\}\, \mathrm{d}t \\
&\qquad + \int_0^{1-\varphi_{GM}(y)} 2tP\{\tilde{\varphi}_{Mn}(y) - \varphi_{GM}(y) > t\}\, \mathrm{d}t \\
&\quad \leqslant \sum_{x=y}^{M} \int_0^{\varphi_{GM}(y)} 2tP\{\varphi_{Mn}(x) - \varphi_{GM}(y) < -t\}\, \mathrm{d}t \\
&\qquad + \sum_{x=0}^{y} \int_0^{1-\varphi_{GM}(y)} 2tP\{\tilde{\varphi}_{Mn}(x) - \varphi_{GM}(y) > t\}\, \mathrm{d}t.
\end{aligned}$$

Then, by the finiteness of M, Lemma 3.3 and (3.1), the result follows directly.

Next, we consider the case where $A \neq \{0, 1, \dots, M\}$. Note that by the definition of $f_{GM}(y)$, that $f_{GM}(y) > 0$ for some $y \in \{1, \dots, M-1\}$ will imply that $f_{GM}(y) > 0$ for all $y=0, 1, \dots, M$. Also, that $f_{GM}(y)=0$ for some $y=\{1, \dots, M-1\}$ implies that $f_{GM}(y)=0$ for all $y \in \{1, \dots, M-1\}$. Thus, $A \neq \{0, 1, \dots, M\}$ iff $A \subset \{0, M\}$. This situation occurs iff the support of the random variable P_{n+1} is contained in $\{0, 1\}$. Let g denote the probability function of such a prior distribution G with support contained in $\{0, 1\}$. Then, $0 \leqslant g(0), g(1) \leqslant 1$ and $g(0) + g(1) = 1$. Therefore, we can obtain that $P\{X_{i,1} = \cdots = X_{i,M+1}\} = 1$ for all $i = 1, \dots, n+1$. In particular, $g(x) = 1$ iff $P\{X_{i,1} = \cdots = X_{i,M+1} = x\} = 1$ for each $x=0, 1$. Under such situation, by the definition of $\tilde{\varphi}_{Mn}(y)$ and (3.1), a straightforward computation shows that $ER_M(G,\tilde{\varphi}_{Mn}) - R_M(G) = 0$. Hence, the theorem holds true.

When both $g(0)$ and $g(1)$ are positive, then $f_{GM}(0) > 0$ and $f_{GM}(M) > 0$. One can also obtain results like Lemmas 3.1-3.3, but with $c_1 = \min(f_{GM}(0), f_{GM}(M)) > 0$. The remaining proof is similar to the case where $A = \{0, 1, \dots, M\}$. We omit the details here. □

For the sequence of empirical Bayes estimators $\{\varphi_{Mn}^*\}$, we also have the following theorem.

Theorem 3.2. *Let $\{\varphi_{Mn}^*\}$ be the sequence of empirical Bayes estimators obtained* *from* $\{\varphi_{Mn}\}$ *by using the monotonizing method of Van Houwelingen* (1977). *Then,* $ER_M(G, \varphi^*_{Mn}) - R_M(G) = O(n^{-1})$.

Proof. By Lemmas 3.1-3.3 and following argument analogous to the proof of Theorem 3.1, one can prove that, for the sequence of the empirical Bayes estimators $\{\varphi_{Mn}\}$,

$$ER_M(G, \varphi_{Mn}) - R_M(G) = O(n^{-1}). \tag{3.2}$$

Therefore, the result of the theorem follows immediately from (2.3) and (3.2). □

4. Rao-Blackwellization of empirical Bayes estimators

The asymptotic optimality of the empirical Bayes estimators proved in Section 3 is a modified empirical Bayes optimality, i.e., optimality with respect to $R_M(G)$ instead of $R_{M+1}(G)$. As mentioned earlier, for this estimation problem, estimators which are asymptotically optimal in the usual empirical Bayes sense do not exist and $\tilde{\varphi}_{Mn}$ and φ^*_{Mn} are not asymptotically optimal with respect to $R_{M+1}(G)$.

Despite this undesirable fact, we still can improve the performance of the empirical Bayes estimators by Rao-Blackwellizing the empirical Bayes estimators $\tilde{\varphi}_{Mn}$ and φ^*_{Mn}.

In general, let $\psi_{Mn}(Y_{n+1,M+1})$ be an empirical Bayes estimator of P_{n+1}. Let $S_{n+1} = Y_{n+1,M+1} + X_{n+1,M+1}$, the number of successes among the total $M+1$ Bernoulli trials taken at stage $n+1$. Let

$$\psi_{M+1,n}(S_{n+1}) = E[\psi_{Mn}(Y_{n+1,M+1}) \mid S_{n+1}], \tag{4.1}$$

where the condition expectation is taken with respect to $Y_{n+1,M+1}$ conditional on S_{n+1}. Straightforward computation yields

$$\psi_{M+1,n}(S_{n+1}) = \frac{M+1-S_{n+1}}{M+1}\psi_{Mn}(S_{n+1}) + \frac{S_{n+1}}{M+1}\psi_{Mn}(S_{n+1}-1). \tag{4.2}$$

where $\psi_{Mn}(M+1) \equiv 1$ and $\psi_{Mn}(-1) \equiv 0$.

We have the following result.

Theorem 4.1. *Let* $\psi_{Mn}(Y_{n+1,M+1})$ *be an empirical Bayes estimator of* P_{n+1} *and let* $\psi_{M+1,n}(S_{n+1})$ *be defined in* (4.1). *Then*

$$E[(\psi_{M+1,n}(S_{n+1}) - P_{n+1})^2] \leqslant E[(\psi_{Mn}(Y_{n+1,M+1}) - P_{n+1})^2]$$

for all n.

Proof. For each fixed p, $0<p<1$, conditional on $(X_1, \ldots, X_n)$, by Rao-Blackwell theorem, we can obtain

$$E_{S_{n+1}}[(\psi_{M+1,n}(S_{n+1})-p)^2 \mid X_1,\dots,X_n]$$
$$\leqslant E_{Y_{n+1,M+1}}[(\psi_{Mn}(Y_{n+1,M+1})-p)^2 \mid X_1,\dots,X_n], \tag{4.3}$$

where $E_{S_{n+1}}$ is computed with respect to a $B(M+1,p)$ distribution while $E_{Y_{n+1,M+1}}$ is computed with respect to a $B(M,p)$ distribution. Since (4.3) holds for all $p\in(0,1)$, $X_1,\dots,X_n$ and n, we have.

$$\begin{aligned} ER_{M+1}(G,\psi_{M+1,n}) &\equiv E[(\psi_{M+1,n}(S_{n+1})-P_{n+1})^2] \\ &\leqslant E[(\psi_{Mn}(Y_{n+1,M+1})-P_{n+1})^2] \\ &\equiv ER_M(G,\psi_{Mn}) \qquad \square \end{aligned} \tag{4.4}$$

We denote the Rao-Blackwellized version of the empirical Bayes estimators φ^*_{Mn} and $\tilde\varphi_{Mn}$ by $\varphi^*_{M+1,n}$ and $\tilde\varphi_{M+1,n}$, respectively. That is,

$$\varphi^*_{M+1,n}(S_{n+1})=\frac{M+1-S_{n+1}}{M+1}\varphi^*_{Mn}(S_{n+1})+\frac{S_{n+1}}{M+1}\varphi^*_{Mn}(S_{n+1}-1), \tag{4.5}$$

$$\tilde\varphi_{M+1,n}(S_{n+1})=\frac{M+1-S_{n+1}}{M+1}\tilde\varphi_{Mn}(S_{n+1})+\frac{S_{n+1}}{M+1}\tilde\varphi_{Mn}(S_{n+1}-1).$$

Since both φ^*_{Mn} and $\tilde\varphi_{Mn}$ possess the monotonicity property, it is easy to see both $\varphi^*_{M+1,n}$ and $\tilde\varphi_{M+1,n}$ possess the monotonicity property.

5. Simulation comparisons

In this section, we compare the small sample performance of the proposed empirical Bayes estimators φ^*_{Mn}, $\tilde\varphi_{Mn}$, $\varphi^*_{M+1,n}$ and $\tilde\varphi_{M+1,n}$ with several known nonparametric empirical Bayes estimators via Monte Carlo simulation study. We let $X_{i,M+1}$ and $Y_{i,M+1}$ $i=1,2,\dots$ be defined as that in Section 2.

For each $y=0,1,\dots,M$, let

$$f_{MnR}(y)=\frac{1}{n+1}\sum_{i=1}^{n}I_{\{y\}}(Y_{i,M+1})+\frac{1}{n+1},$$

$$\omega_{MnR}(y)=\frac{y+1}{(n+1)(M+1)}\sum_{i=1}^{n}I_{\{y\}}(Y_{i,M+1}+X_{i,M+1}).$$

Robbins (1956) suggested estimating the binomial parameter P_{n+1} by $\varphi_{MnR}(Y_{n+1,M+1})$, where

$$\varphi_{MnR}(Y_{n+1,M+1})=\frac{\omega_{MnR}(Y_{n+1,M+1})}{f_{MnR}(Y_{n+1,M+1})}. \tag{5.1}$$

Note that $P\{\varphi_{MnR}(Y_{n+1,M+1})>1\}>0$. However, it is known that $0\leqslant\varphi_{GM}(Y_{n+1,M+1})\leqslant 1$. Thus, Vardeman (1978) considered a variant of (5.1) and suggested estimating P_{n+1} by $\varphi_{MnV}(Y_{n+1,M+1})$ where

$$\varphi_{MnV}(Y_{n+1,M+1}) = \min\{\varphi_{MnR}(Y_{n+1,M+1}), 1\}. \tag{5.2}$$

One can see that $\varphi_{MnV}(Y_{n+1,M+1})$ is superior to $\varphi_{MnR}(Y_{n+1,M+1})$ in the sense that $R_M(G,\varphi_{MnV}) \leqslant R_M(G,\varphi_{MnR})$ for any prior distribution G and any past observations $X_1,\ldots,X_n$.

Liang (1989) proposed using $\varphi_{MnL}(Y_{n+1,M+1})$ to estimate P_{n+1}, where

$$\varphi_{MnL}(y_{n+1,M+1}) = \frac{\omega_{MnL}(Y_{n+1,M+1})}{f_{MnR}(Y_{n+1,M+1})}, \tag{5.3}$$

with

$$\omega_{MnL}(y) = \frac{1}{n+1}\sum_{i=1}^{n} X_{i,M+1} I_{\{y\}}(Y_{i,M+1}) + \frac{1}{n+1} X_{n+1,M+1},$$

$$y = 0, 1, \ldots, M.$$

Gupta and Liang (1988) proposed a monotone empirical Bayes estimator, say $\varphi_{MnGL}(Y_{n+1,M+1})$ using the isotonic regression method with equal weights. The estimator φ_{MnGL} is constructed as follows. Let, for each $y=0,1,\ldots,M$,

$$f_n(y) = \frac{1}{n}\sum_{i=1}^{n} I_{\{y\}}(Y_{i,M+1}) + \frac{1}{n},$$

$$\omega_n(y) = \frac{1}{n}\sum_{i=1}^{n} X_{i,M+1} I_{\{y\}}(Y_{i,M+1}) + \frac{1}{n},$$

and

$$\varphi_n(y) = \frac{\omega_n(y)}{f_n(y)}.$$

Then, $\varphi_{MnGL}(y)$, $y=0,1,\ldots,M$, is defined as follows:

$$\varphi_{MnGL}(y) = \max_{0\leqslant s\leqslant y}\ \min_{s\leqslant t\leqslant M} \left\{ \sum_{y=s}^{t} \varphi_n(y)/(t-s+1) \right\}. \tag{5.4}$$

Since φ_{MnV} and φ_{MnL} are not monotone, the monotonizing method of Van Houwelingen (1977) is applied to them and the resulting monotone empirical Bayes estimators are denoted by φ^*_{MnV} and φ^*_{MnL}, respectively. Finally, we let $\varphi^*_{M+1,nV}$, $\varphi^*_{M+1,nL}$ and $\varphi_{M+1,nGL}$ denote the Rao–Blackwellized version of the empirical Bayes estimators φ^*_{MnV}, φ^*_{MnL} and φ_{MnGL}, respectively.

In the following, we compare the performance among the empirical Bayes estimators φ^*_{Mn}, $\tilde{\varphi}_{Mn}$, φ_{MnV}, φ^*_{MnV}, φ_{MnL}, φ^*_{MnL}, φ_{MnGL} and the associated Rao–Blackwellized versions. We let the prior distribution G be a member in the family of beta distributions with parameters (α,β). The simulation scheme used in this paper is described as follows:

(1) At stage $i, i=1,2,\ldots$, generate random value p_i according to the prior distribution G. Then, generate $M+1$ Bernoulli random variables $X_{i,1},\ldots,X_{i,M+1}$ from a $B(1,p_i)$ distribution.

(2) For each n, based on the data $\{X_{i,j}, Y_{i,j}, j=1,\ldots,M+1, i=1,\ldots,n+1\}$, con-

struct the empirical Bayes estimators φ^*_{Mn}, $\tilde{\varphi}_{Mn}$, φ_{MnV}, φ^*_{MnV}, φ_{MnL}, φ^*_{MnL} and φ_{MnGL}. Then, compute the corresponding conditional Bayes risk $R_M(G, \psi_{Mn})$ where ψ_{Mn} denote the related empirical Bayes estimator.

(3) For each of the empirical Bayes estimators, derive the corresponding Rao–Blackwellized version and compute the associated conditional Bayes risk $R_{M+1}(G, \psi_{M+1,n})$ where $\psi_{M+1,n}$ is the Rao–Blackwellized version of ψ_{Mn}.

(4) The process is repeated 500 times. For each n, the average, denoted by $\hat{E}R_M(G, \psi_{Mn})$ (or $\hat{E}R_{M+1}(G, \psi_{M+1,n})$ for the Rao–Blackwellized empirical Bayes estimator $\psi_{M+1,n}$) based on the 500 conditional Bayes risks, is used as an estimator of the corresponding overall expected Bayes risk $ER_M(G, \psi_{Mn})$ ($ER_{M+1}(G, \psi_{M+1,n})$).

The simulation study has been carried out for several values of the parameters (α, β) and M. Since the simulation comparison indicates similar result, we only report the results of the cases where $(\alpha, \beta) = (0.5, 0.5)$, $(5, 5)$ and $M = 15$. In the tables, the numbers in the parentheses are the estimated standard errors of the corresponding estimators $\hat{E}R_M(G, \psi_{Mn})(\hat{E}R_{M+1}(G, \psi_{M+1,n}))$.

The simulation results indicate that the performance of the empirical Bayes estimators $\tilde{\varphi}_{Mn}$ and φ^*_{Mn} is much better than that of φ_{MnV}, φ^*_{MnV}, φ_{MnL}, φ^*_{MnL} and φ_{MnGL}, especially when the number of the past observations n is very small. Similar results also hold for the Rao–Blackwellized version $\tilde{\varphi}_{M+1,n}$ and $\varphi^*_{M+1,n}$ compared with the others. The effect may be due to the use of all the information $(X_{i,j}, Y_{i,j})$, $j = 1, \ldots, M+1$, $i = 1, \ldots, n$. If we compare the performance of the empirical Bayes estimator $\varphi_{MnV}(\varphi_{MnL})$ with the corresponding monotonized version $\varphi^*_{MnV}(\varphi^*_{MnL})$, we can see that the small-sample performance of $\varphi^*_{MnV}(\varphi^*_{MnL})$ is much better than that of $\varphi_{MnL}(\varphi_{MnV})$. This fact may also indicate the importance of the monotonicity property for empirical Bayes estimators for the estimation problem under study. Finally, we compare the performance of $\tilde{\varphi}_{Mn}(\tilde{\varphi}_{M+1,n})$ and $\varphi^*_{Mn}(\varphi^*_{M+1,n})$. From Tables 1 and 2, it can be seen that the performance of $\varphi^*_{Mn}(\varphi^*_{M+1,n})$ is always better than that of $\tilde{\varphi}_{Mn}(\tilde{\varphi}_{M+1,n})$. This may indicate the superiority of the Van Houwelingen's monotonizing method to the isotonic monotonizing method.

Acknowledgement

The authors are grateful to a referee for very helpful comments that led to an improved article.

Table 1(a)

Estimated Bayes risks for estimators with no Rao-Blackwellization. $M=15$, $(\alpha,\beta)=(0.5,0.5)$, $R_M(G)=7.813\times10^{-3}$, $R_{M+1}(G)=7.353\times10^{-3}$

n	$\hat{E}R_M(G,\tilde{\varphi}_{Mn})$	$\hat{E}R_M(G,\varphi^*_{Mn})$	$\hat{E}R_M(G,\varphi_{MnV})$	$\hat{E}R_M(G,\varphi^*_{MnV})$	$\hat{E}R_M(G,\varphi_{MnL})$	$\hat{E}R_M(G,\varphi^*_{MnL})$	$\hat{E}R_M(G,\varphi_{MnGL})$
20	14.199×10^{-3}	10.361×10^{-3}	134.981×10^{-3}	29.466×10^{-3}	151.539×10^{-3}	45.830×10^{-3}	108.677×10^{-3}
	(0.254×10^{-3})	(0.147×10^{-3})	(1.756×10^{-3})	(0.599×10^{-3})	(1.956×10^{-3})	(0.930×10^{-3})	(1.442×10^{-3})
40	11.711×10^{-3}	9.121×10^{-3}	74.789×10^{-3}	17.068×10^{-3}	86.925×10^{-3}	25.403×10^{-3}	55.108×10^{-3}
	(0.067×10^{-3})	(0.036×10^{-3})	(1.042×10^{-3})	(0.223×10^{-3})	(1.184×10^{-3})	(0.405×10^{-3})	(0.898×10^{-3})
60	10.984×10^{-3}	8.996×10^{-3}	52.976×10^{-3}	13.476×10^{-3}	61.149×10^{-3}	18.245×10^{-3}	36.883×10^{-3}
	(0.058×10^{-3})	(0.028×10^{-3})	(0.748×10^{-3})	(0.122×10^{-3})	(0.913×10^{-3})	(0.274×10^{-3})	(0.619×10^{-3})
80	10.546×10^{-3}	8.889×10^{-3}	40.545×10^{-3}	11.829×10^{-3}	47.037×10^{-3}	15.382×10^{-3}	27.842×10^{-3}
	(0.047×10^{-3})	(0.022×10^{-3})	(0.569×10^{-3})	(0.086×10^{-3})	(0.730×10^{-3})	(0.200×10^{-3})	(0.469×10^{-3})
100	10.239×10^{-3}	8.861×10^{-3}	33.560×10^{-3}	10.916×10^{-3}	38.431×10^{-3}	13.639×10^{-3}	22.769×10^{-3}
	(0.040×10^{-3})	(0.022×10^{-3})	(0.469×10^{-3})	(0.066×10^{-3})	(0.538×10^{-3})	(0.169×10^{-3})	(0.352×10^{-3})
120	10.018×10^{-3}	8.846×10^{-3}	29.160×10^{-3}	10.366×10^{-3}	33.207×10^{-3}	12.609×10^{-3}	19.690×10^{-3}
	(0.036×10^{-3})	(0.020×10^{-3})	(0.386×10^{-3})	(0.366×10^{-3})	(0.386×10^{-3})	(0.136×10^{-3})	(0.282×10^{-3})
160	9.705×10^{-3}	8.758×10^{-3}	23.995×10^{-3}	9.698×10^{-3}	26.308×10^{-3}	11.652×10^{-3}	16.198×10^{-3}
	(0.032×10^{-3})	(0.018×10^{-3})	(0.297×10^{-3})	(0.035×10^{-3})	(0.346×10^{-3})	(0.108×10^{-3})	(0.190×10^{-3})
200	9.473×10^{-3}	8.678×10^{-3}	21.218×10^{-3}	9.393×10^{-3}	22.176×10^{-3}	10.958×10^{-3}	14.436×10^{-3}
	(0.027×10^{-3})	(0.017×10^{-3})	(0.278×10^{-3})	(0.028×10^{-3})	(0.261×10^{-3})	(0.080×10^{-3})	(0.146×10^{-3})

Table 1(b)
Estimated Bayes risks for estimators with no Rao–Blackwellization. $M=15$, $(\alpha,\beta)=(5,5)$, $R_M(G)=9.091\times10^{-3}$, $R_{M+1}(G)=8.741\times10^{-3}$

n	$\hat{E}R_M(G,\tilde{\varphi}_{Mn})$	$\hat{E}R_M(G,\varphi^*_{Mn})$	$\hat{E}R_M(G,\varphi_{MnV})$	$\hat{E}R_M(G,\varphi^*_{MnV})$	$\hat{E}R_M(G,\varphi_{MnL})$	$\hat{E}R_M(G,\varphi^*_{MnL})$	$\hat{E}R_M(G,\varphi_{MnGL})$
20	16.254×10^{-3}	13.529×10^{-3}	110.037×10^{-3}	20.378×10^{-3}	131.002×10^{-3}	31.682×10^{-3}	97.054×10^{-3}
	(0.140×10^{-3})	(0.132×10^{-3})	(1.196×10^{-3})	(0.200×10^{-3})	(1.453×10^{-3})	(0.815×10^{-3})	(1.441×10^{-3})
40	13.347×10^{-3}	11.740×10^{-3}	71.683×10^{-3}	14.731×10^{-3}	84.387×10^{-3}	21.020×10^{-3}	58.626×10^{-3}
	(0.073×10^{-3})	(0.051×10^{-3})	(0.916×10^{-3})	(0.805×10^{-3})	(0.139×10^{-3})	(0.415×10^{-3})	(1.058×10^{-3})
60	12.627×10^{-3}	11.239×10^{-3}	57.123×10^{-3}	13.181×10^{-3}	63.782×10^{-3}	17.723×10^{-3}	42.533×10^{-3}
	(0.060×10^{-3})	(0.043×10^{-3})	(0.823×10^{-3})	(0.052×10^{-3})	(0.811×10^{-3})	(0.259×10^{-3})	(0.730×10^{-3})
80	12.012×10^{-3}	10.848×10^{-3}	47.832×10^{-3}	12.355×10^{-3}	51.529×10^{-3}	15.845×10^{-3}	33.556×10^{-3}
	(0.049×10^{-3})	(0.033×10^{-3})	(0.676×10^{-3})	(0.043×10^{-3})	(0.687×10^{-3})	(0.193×10^{-3})	(0.562×10^{-3})
100	11.588×10^{-3}	10.616×10^{-3}	41.010×10^{-3}	11.721×10^{-3}	44.257×10^{-3}	14.674×10^{-3}	29.017×10^{-3}
	(0.040×10^{-3})	(0.027×10^{-3})	(0.618×10^{-3})	(0.040×10^{-3})	(0.577×10^{-3})	(0.170×10^{-3})	(0.453×10^{-3})
120	11.277×10^{-3}	10.474×10^{-3}	36.149×10^{-3}	11.437×10^{-3}	37.913×10^{-3}	13.834×10^{-3}	25.951×10^{-3}
	(0.036×10^{-3})	(0.025×10^{-3})	(0.512×10^{-3})	(0.030×10^{-3})	(0.448×10^{-3})	(0.136×10^{-3})	(0.438×10^{-3})
160	10.974×10^{-3}	10.263×10^{-3}	29.697×10^{-3}	10.959×10^{-3}	31.080×10^{-3}	13.111×10^{-3}	21.900×10^{-3}
	(0.030×10^{-3})	(0.021×10^{-3})	(0.392×10^{-3})	(0.026×10^{-3})	(0.341×10^{-3})	(0.115×10^{-3})	(0.309×10^{-3})
200	10.732×10^{-3}	10.296×10^{-3}	26.170×10^{-3}	10.653×10^{-3}	26.623×10^{-3}	12.389×10^{-3}	19.256×10^{-3}
	(0.029×10^{-3})	(0.019×10^{-3})	(0.336×10^{-3})	(0.022×10^{-3})	(0.274×10^{-3})	(0.086×10^{-3})	(0.255×10^{-3})

Table 2(a)
Estimated Bayes risks for Rao-Blackwellized estimators. $M=15$, $(\alpha,\beta)=(0.5,0.5)$, $R_M(G)=7.813\times10^{-3}$, $R_{M+1}(G)=7.353\times10^{-3}$

n	$\hat{E}R_{M+1}(G,\tilde{\varphi}_{M+1,n})$	$\hat{E}R_{M+1}(G,\varphi^*_{M+1,n})$	$\hat{E}R_{M+1}(G,\varphi^*_{M+1,nV})$	$\hat{E}R_{M+1}(G,\varphi^*_{M+1,nL})$	$\hat{E}R_{M+1}(G,\varphi_{M+1,nGL})$
20	12.321×10^{-3}	9.670×10^{-3}	28.832×10^{-3}	44.960×10^{-3}	107.368×10^{-3}
	(0.239×10^{-3})	(0.137×10^{-3})	(0.598×10^{-3})	(0.917×10^{-3})	(1.444×10^{-3})
40	10.086×10^{-3}	8.469×10^{-3}	16.485×10^{-3}	24.625×10^{-3}	53.654×10^{-3}
	(0.055×10^{-3})	(0.032×10^{-3})	(0.221×10^{-3})	(0.442×10^{-3})	(0.895×10^{-3})
60	9.511×10^{-3}	8.337×10^{-3}	12.895×10^{-3}	17.498×10^{-3}	35.403×10^{-3}
	(0.045×10^{-3})	(0.025×10^{-3})	(0.120×10^{-3})	(0.267×10^{-3})	(0.614×10^{-3})
80	9.160×10^{-3}	8.224×10^{-3}	11.240×10^{-3}	14.655×10^{-3}	26.412×10^{-3}
	(0.035×10^{-3})	(0.019×10^{-3})	(0.085×10^{-3})	(0.196×10^{-3})	(0.467×10^{-3})
100	8.924×10^{-3}	8.187×10^{-3}	10.323×10^{-3}	12.927×10^{-3}	21.410×10^{-3}
	(0.030×10^{-3})	(0.019×10^{-3})	(0.065×10^{-3})	(0.165×10^{-3})	(0.349×10^{-3})
120	8.759×10^{-3}	8.158×10^{-3}	9.702×10^{-3}	11.889×10^{-3}	18.332×10^{-3}
	(0.027×10^{-3})	(0.019×10^{-3})	(0.052×10^{-3})	(0.133×10^{-3})	(0.279×10^{-3})
160	8.533×10^{-3}	8.069×10^{-3}	9.080×10^{-3}	10.939×10^{-3}	14.917×10^{-3}
	(0.022×10^{-3})	(0.015×10^{-3})	(0.033×10^{-3})	(0.105×10^{-3})	(0.187×10^{-3})
200	8.359×10^{-3}	7.996×10^{-3}	8.773×10^{-3}	10.241×10^{-3}	13.205×10^{-3}
	(0.019×10^{-3})	(0.013×10^{-3})	(0.026×10^{-3})	(0.078×10^{-3})	(0.143×10^{-3})

Table 2(b)
Estimated Bayes risks for Rao-Blackwellized estimators. $M=15$, $(\alpha,\beta)=(5,5)$, $R_M(G)=9.091\times10^{-3}$, $R_{M+1}(G)=8.741\times10^{-3}$

n	$\hat{E}R_{M+1}(G,\tilde{\varphi}_{M+1,n})$	$\hat{E}R_{M+1}(G,\varphi^*_{M+1,n})$	$\hat{E}R_{M+1}(G,\varphi^*_{M+1,nI'})$	$\hat{E}R_{M+1}(G,\varphi^*_{M+1,nL})$	$\hat{E}R_{M+1}(G,\varphi_{M+1,nGL})$
20	13.873×10^{-3}	12.416×10^{-3}	19.137×10^{-3}	30.180×10^{-3}	96.618×10^{-3}
	(0.111×10^{-3})	(0.117×10^{-3})	(0.195×10^{-3})	(0.792×10^{-3})	(1.450×10^{-3})
40	11.576×10^{-3}	10.939×10^{-3}	13.956×10^{-3}	19.881×10^{-3}	57.717×10^{-3}
	(0.054×10^{-3})	(0.045×10^{-3})	(0.079×10^{-3})	(0.398×10^{-3})	(1.063×10^{-3})
60	11.054×10^{-3}	10.480×10^{-3}	12.438×10^{-3}	16.671×10^{-3}	41.572×10^{-3}
	(0.043×10^{-3})	(0.038×10^{-3})	(0.050×10^{-3})	(0.247×10^{-3})	(0.737×10^{-3})
80	10.598×10^{-3}	10.137×10^{-1}	11.629×10^{-3}	14.907×10^{-3}	32.670×10^{-3}
	(0.035×10^{-3})	(0.029×10^{-3})	(0.041×10^{-3})	(0.183×10^{-3})	(0.569×10^{-3})
100	10.279×10^{-3}	9.923×10^{-3}	11.033×10^{-3}	13.298×10^{-3}	28.201×10^{-3}
	(0.028×10^{-3})	(0.023×10^{-3})	(0.032×10^{-3})	(0.161×10^{-3})	(0.458×10^{-3})
120	10.082×10^{-3}	9.798×10^{-3}	10.753×10^{-3}	12.991×10^{-3}	25.106×10^{-3}
	(0.025×10^{-3})	(0.022×10^{-3})	(0.029×10^{-3})	(0.128×10^{-3})	(0.444×10^{-3})
160	9.879×10^{-3}	9.617×10^{-3}	10.293×10^{-3}	12.309×10^{-3}	21.900×10^{-3}
	(0.022×10^{-3})	(0.018×10^{-3})	(0.023×10^{-3})	(0.108×10^{-3})	(0.309×10^{-3})
200	9.715×10^{-3}	9.470×10^{-3}	10.012×10^{-3}	11.634×10^{-3}	18.485×10^{-3}
	(0.019×10^{-3})	(0.015×10^{-3})	(0.020×10^{-3})	(0.081×10^{-3})	(0.261×10^{-3})

References

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis.* Second Edition, Springer-Verlag, New York.

Berry, D.A. and R. Christensen (1979). Empirical Bayes estimation of a binomial parameter via mixtures of Dirichlet processes. *Ann. Statist.* **7**, 558-567.

Griffin, B.S. and R.G. Krutchkoff (1971). Optimal linear estimators: an empirical Bayes version with application to the binomial distribution. *Biometrika* **58**, 195-201.

Gupta, S.S. and T. Liang (1988). Empirical Bayes rules for selecting the best binomial population. *Statistical Decision Theory and Related Topics-IV* (Eds. S.S. Gupta and J.O. Berger), Springer-Verlag, Berlin, Vol. I, 213-224.

Gutman, S. (1982). Minimax linear empirical Bayes estimation of binomial parameter. *Commun. Statist.—Theor. Meth.* **11** (18) 2075-2082.

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

Liang, T. (1989). Convergence rates for empirical Bayes estimation of binomial parameter. *J. Statist. Plann. Inference* **21**, 117-124.

Martz, H.F. and M.G. Lian (1974). Empirical Bayes estimation of the binomial parameter. *Biometrika* **61**, 517-523.

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

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

Robbins, H. (1983). Some thoughts on empirical Bayes estimation. *Ann. Statist.* **11**, 713-723.

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

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.