Journal of the KSQC
Vol. 19, No. 2, Dec.1991

1991년 12월 韓國品質管理學會誌 제19권 제2호

EMPIRICAL BAYES ESTIMATION OF BINOMIAL PARAMETER WITH AN APPLICATION TO THE LOT ACCEPTANCE SAMPLING PROBLEM**

INHA JUNG* SEUNG-HO LEE*

ABSTRACT

Empircal Bayes estimation of the binomial parameter θ is considered when the sample size in each component problem is random. The result is applied to the lot acceptanec sampling problem in the presence of various kinds of costs including cost for sampling.

1. Introduction

In the usual empirical Bayes problem introduced by Robbins(1951, 1956, 1964) solutions are obtained by utilizing the fact that the observations associated with each repetition of the component problem are iid. Most of the empirical Bayes works following Robbins have concerned with iid components with a few exceptions. O'Bryan(1972, 1976) introduced the nonparametric empirical Bayes decision problem with non-iid components by allowing unequal nonrandom sample

* Departmemt of Mathematics, Ajou University, Suwon 441-749, Korea.

** This research was supported by the Korean Ministry of Education.

size in the component problem. Gilliland and Karunamuni(1988) consider the possibility of varying stochastic sample sizes. Karunamuni(1985, 1988) studies an empirical Bayes problem with a sequential component with a linear loss and multiple decision loss structure.

We consider an empirical Bayes decision problem in which data ccumulated over past component problems are allowed to be used in selection both the sample size and the decision rule to be used in the current component problem.

The component problem that will be considered here, which will be described in detail in *section 2*, is the squared error loss estimation of θ in Binomial $(m, \theta)$when there is a constant cost $c>0$ per observation. Here $m>0$ is given positive integer and the prior distribution G for θ is assumed to be in the parametric family of conjugate priors

$$g = \{\text{Beta}(\alpha, \beta) \mid \alpha>0, \beta>0\}.$$

Let D_n denote the set of all decision rules based on n observations. Minimum Bayes risk $R_n(G)$ is attained by a Bayes rule $d_G \in D_n$. Let

$$r_n(G) = R_n + cn$$

the minimun Bates risk plus cost for n observations.

As functions of n $=0,1,2,\cdots\cdots$,$R_n(G)$ is nonincreasing and cn is strictly increasing in n. Therefore there is an integer-valued minimizer $n^*(G)$ of $r_n(G)$ over $n=0,1,2,\cdots$, i.e.

$$r(G) = r_{n^*(G)}(G) = \min\{ r_n(G) \mid n=0,1,2,\cdots \}.$$

We call the integer minimizer $n^*(G)$ an potimal sample size with respect to G.

By using a Bayes rule $d_G \in D_{n^*(G)}$ based on the optimal sample size $n^*(G)$, the best possible risk goal is achieved.

At the $(i+1)$th component problem with $G \in g$ unknown, G is estimated using the data from the previous i component problems. With respect to this estimated prior, optimal sample size N_{i+1} and a Bayes rule $d_{i+1} \in D_{N_{i+1}}$ for the current $(i+1)th$ component are obtained.

Let

$$\underline{N} = (N_1, N_2, \cdots)$$
$$\underline{d} = (d_1, d_2, \cdots)$$

, where N_l, d_l are nonrandom choices.

We will be concerned with the risk behavior of the empirical Bayes procedure$(\underline{N}, \underline{d})$.

The risk for the decision about θ_i is

$$Er_{N_i}(G, d_i) = ER_{N_i}(G, d_i) + cEN_i$$

where E denotes the expectation with respect to the observations from the earlier component problems.

Definition1.1 If the empirical Bayes procedure $(\underline{N}, \underline{d})$ possesses the property :

$$\lim_{i} E r_{N_i}(G,d_i) = r(G)$$

for all $G \in g$, we say it is asymptotically optimal (a.o.).

This means that in the limit, the empirical Bayes procedure has the best possible risk behavior.

In *section 3,* we develop an a.o. empirical Bayes procedure $(\underline{N},\underline{d})$ for the component problem described in *section 2.*

The following remark shows that asymptotic optimality implies the convergence of the sample size N_i to the set of optimal sample sizes with respect to each $G \in g$.

Remark1.1 Let $S(G)$ denote the set of optimal sizes with respect to G.

(a) If $(\underline{N},\underline{d})$ *is a. o. at G,* then $P(N_i \in S(G)) \rightarrow 1$ as $i \rightarrow \infty$.

(b) If $r_{N_i}((G, d_i) \rightarrow r(G)$ a.s., then $p(N_i \in S(G)$ for all but a finite number of i) = 1.

In *section 4,* we consider an application to the lot acceptance sampling problem by introducing various costs including cost for sampling.

2. The Component Problem

Let $X_1, X_2, \cdots$ be i.i.d. Binomial(m,θ), where m is a given positive integer and the parameter θ have prior distribution G in the beta family $g = \{Beta\ (\alpha, \beta) \mid \alpha > 0, \beta > 0\}$.

Estimation of θ is considerde for squared-error loss. Let $c>0$ be a constant cost per observbation. Let $d \in D_n$ be a decision rule based on the observation $\underline{X}^n = (X_1, \cdots, X_n)$.

The decision loss plus cost for observation is given by$[\theta - d(\underline{X}^n)]^2 + cn$.

The marginal distribution of X_i is Beta-Binomial. We let ξ and η denote the first two momemts of $G = \mathrm{Beta}(\alpha,\beta)$, that is,

$$\xi = E_G\theta = \frac{\alpha}{\alpha+\beta}$$

$$F_G\theta^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \quad (2.1$$

and note that $0 < \xi^2 < \eta < 1$ since $\alpha>0$, $\beta>0$. Also

$$E(X_i) = m\xi$$

$$E(X_i^2) = (\xi - \eta) + m^2\eta, \quad (2.2)$$

and from (2.2) it follows that

$$\alpha=\frac{\xi(\xi-\eta)}{\eta-\xi^2}$$
$$\beta=\frac{(1-\xi)(\xi-\eta)}{\eta-\xi^2} \qquad (2.3)$$

In the empirical Bayes application,(2.2) and (2.3) will useful in the construction of consistent estimators for α and β. We will use the method of momemts to obtain estimates of ξ and η and will use (2.3) to obtain estimates for the parameters α and β.

A Bayes rule exists and is given by the posterior mean of θ, given $\underline{X}^m$. The posterior distribution of θ given $\underline{X}^m$, is Beta $(\alpha+n\overline{X}_n, \beta+mn-n\overline{X})$, where $\overline{X}_n$ denotes the average of $X_1, \cdots, X_n$.

Hence, a Bayes rule $d_G \in D_n$ is

$$d_G(\underline{X}^n)=\alpha+n\overline{X}_n$$
$$=\frac{\alpha}{\alpha+\beta+mn}+\frac{n}{\alpha+\beta+mn}\overline{X}_n. \qquad (2.4)$$

if $G=\text{Beta}(\alpha,\beta)$.

Remark2.1 For $G=\text{Beta}(\alpha,\beta)$ and $G'=\text{Beta}(\alpha', \beta')$,

$$R_n(G,d_{G'})=\frac{1}{[(\alpha'+\beta'+mn)^2}\{[\alpha'+\beta')^2-mn]\eta-[2\alpha'(\alpha'\ \beta')-mn]\xi+(\alpha')^2\}, \qquad (2.5)$$

$$|\,Rm(G,d_G)-Rm(G',d_G)\,|\leq 2\,|\,\xi-\xi'|+|\,\eta-\eta'|, \qquad (2.6)$$

and

$$Rn(G)=\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)(\alpha+\beta+mn)} \qquad (2.7)$$

From(2.7) the minimum Bayes risk including cost for observations is

$$r_n(G)=\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)}(\alpha+\beta+mn)^{-1}+cn. \qquad (2.8)$$

We seek the optimal sample size $n^*(G)=n^*(\alpha,\beta)$. $r_n(G)$ is a continuous and convex function of real n. Consider the equation

$$0=\frac{d}{dn}r_n(G)=-\frac{m\alpha\beta}{(\alpha-\beta)(\alpha+\beta+1)}(\alpha+\beta+mn)^{-2}+c.$$

Its larger solution is

$$\nu=\left[\frac{m}{c}\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)}\right]^{1/2} \qquad (2.9)$$

and an optimal fixed sample size $n^*=n^*(\alpha,\beta)$ is given by

$$n^*(G)=\begin{cases} 1, & if\ \nu<1 \\ \nu, & if\ \nu\in\{1,2,3,\cdots\} \\ [\nu] \text{ or } [\nu]+1 \text{ depending on which} & \\ \text{integer minimizes } r_n(G), & \text{otherwise.} \end{cases} \quad (2.10)$$

Here [] dedotes the greatest integer function and take $n^*(G)=[\nu]$ if both $[\nu]$ and $[\nu]+1$ minimize $r_n(G)$.

If α and β were known constants, we can use $d_G\in D_{n^*(G)}$ to achieve minimum Bayes risk, i.e.,

$$r(G)=\min\{r_n(G)\mid n=1,2,\cdots\}.$$

In the next section we show how (α,β) is estimated in the empirical Bayes problem with this component and establish the asymptotic optimality for the resulting procedure.

3. The Empirical Bayes Decision Procedure

Consider the binomial component problem of the last section. Let $\hat{\alpha}_0$, $\hat{\beta}_0$ be initial nonrandom estimates of α, β and let $N_1 = n^*(\hat{\alpha}_0, \hat{\beta}_0)$ be the sample size chosen for the first component. (See(2.20) for the definition of the optimal fixed size function n^*.) Recall that $\underline{X}^1=(X_{11},X_{12},\cdots,X_{1N_1})$ denotes the vector of observations from the first component.

We will define a sequence of estimates $\hat{\alpha}_i$, $\hat{\beta}_i$ based on $(\underline{X}^1, \underline{X}^2, \cdots, \underline{X}^i)$. *Then for component* $i+1$, the empirical Bayes sample size is $N_{i+1}=n^*(\hat{\alpha}_i, \hat{\beta}_i)$ and the empirical Bayes estimator of θ_{i+1} is

$$d_{i+1}(\underline{X}^{i+1})=\frac{\hat{\alpha}_i+N_{i+1}Y_{i+1}}{\hat{\alpha}_i+\hat{\beta}_i+mN_{i+1}},\qquad i=0,1,\cdots \quad (3.1)$$

, where

$$Y_i=\frac{1}{N_i}\sum_{j=1}^{N_i}X_{ij},\qquad i=1,2,\cdots \quad (3.2)$$

We will give estimates based on the method of momemts and will find it useful to consider

$$Z_i=\frac{1}{N_i}X_{ij}^2,\qquad i=1,2,\cdots \quad (3.3)$$

and denote average of Y_j, Z_j $j=1,2,\cdots,i$ *as* $\underline{Y}_i$, $\underline{Z}_i$, $i=1,2,\cdots$.

Let F_0 be the trivial σ-field and let $F_j=\sigma(\underline{X}^1,\underline{X}^2,\cdots,\underline{X}^j)$, $j=1,2,\cdots$. The sample size N_j is F_{j-1} measurable, $j=1,2,\cdots$ and we see that

$$E(Y_j\mid F_{j-1})=m\xi,\qquad j=1,2,\cdots$$
$$E(Z_j\mid F_{j-1})=m(\xi-\eta)+m^2\eta,\quad j=1,2,\cdots \quad (3.4)$$

follow from (2.2).

Since $Y_j \leqq m$ and $Z_j \leqq m^2$, $j=1,2,\cdots$ the strong law for centerings at conditional expectation (see Hall and Heyde (1980), Theorem 2.19)) implies

$$\overline{Y}_i - \frac{1}{i}\sum_{j=1}^{i} E(Y_j \mid F_{j-1}) \rightarrow 0 \quad \text{a.s}$$
$$\overline{Z}_i - \frac{1}{i}\sum_{j=1}^{i} E(Z_j \mid F_{j-1}) \rightarrow 0. \quad \text{a.s.} \tag{3.5}$$

From (2.14) and (2.15) we have

$$\overline{Y}_i \rightarrow m\xi \quad \text{a.s.}$$
$$\overline{Z}_i \rightarrow m(\xi - \eta) \quad \text{a.s.} \tag{3.6}$$

Lemma 3.1 Let $m \geq 2$. The estimators defined for $i=1,2,\cdots$ by

$$\hat{\xi}_i \equiv \frac{\overline{X}_i}{m}$$
$$\hat{\eta}_i \equiv \frac{\overline{Z}_i - \overline{Y}_i}{m(m-1)} \tag{3.7}$$

$$\hat{\alpha}_i \equiv \left[\frac{\hat{\xi}_i(\hat{\xi}_i - \hat{\eta}_i)}{\hat{\eta}_i - \hat{\xi}_i^2} \right]^+$$
$$\hat{\beta}_i \equiv \left[\frac{(1-\hat{\xi}_i)(\hat{\xi}_i - \hat{\eta}_i)}{\hat{\eta}_i - \hat{\xi}_i^2} \right]^+ \tag{3.8}$$

are a.s,. consistent. (In (3.8) take ratios 0/0 to be 0)

Proof. The a.s. convergence of the estimates (3.7) follows from (3.6). The a.s. convergence of the estimates (3.8) then follows from (2.3).

Theorem 3.1 Let $m \geq 2$. The empirical Bayes procedure $(\underline{N}, \underline{d})$ defined above is asymptotically optimal at each $G = (\alpha, \beta)$.

Proof. By (2.6),

$$0 \leq r_{N_i+1}(G, d_{i+1}) - r(G) \leq 4 \mid \hat{\xi}_i - \xi \mid + 2 \mid \hat{\eta}_i - \eta \mid . \tag{3.9}$$

Since $\mid \hat{\xi}_i - \xi \mid \leq 1$ and $\mid \hat{\eta}_i - \eta \mid$ for all i, the DCT and (3.9) imply that $r_{N_i+1}(G, d_{i+1}) \rightarrow r(G)$.

Remark 3.2 In the component problem under consideration the marginal distribution of a single observation is Beta-Binomial with parameters m, α, β. If $m=1$, this is Binomial $(1, \alpha/\alpha+\beta)$ and the pair (α, β) is not identified. Our method of estimation in the empirical Bayes version requires that $m \geq 2$. This assumption can be removed if we require that the $N_i \geq 2$ and use estimators based on pooled data. Requiring $N_i \geq 2$ i.o. would suffice but details of these variations will not be presented.

4. Application to Lot Acceptance Sampling Problem.

The lot of K units from a certain production process is shipped or not shipped according to the number of defective units found by inspection in the sample of size $n(\leq K)$ from the lot. Let X and Y denote the number of defectives in the sample and in the remaining lot after sampling. Any defective unit found in the sample by inspection is reworked at a cost d. Cost of sampling is cn. Besides these costs, there are also costs a for each bad item sent and cost b for each good item not sent. Then the loss from not shipping the lot is

$$cn+dX+aY \tag{4.1}$$

and the loss from shipping the lot is

$$cn+dX+b(K-n-Y). \tag{4.2}$$

Here the random variable X is the sum of n independent Bernoulli random variables with parameter θ and θ is a random variable with prior distribution G on $\Theta = [0,1]$. Let $\chi^{(n)}=\{0,1,\cdots,n\}$, *the sample space of* X. The action space is $A = \{0,1\}$ where 0 denotes the action "Ship the lot" and 1 denots the action "Do not ship the lot". Let D_n denote the set of all nonrandomized decision rules δ, which are A-valued functions on $\chi^{(n)}$.

Any non-randomized decision rule $\delta \in D_n$ is an indicator of a set $A \subset \chi^{(n)}$ i. e. for each $x \in \chi^{(n)}$

$$\delta(x) = I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

It follows from (4.1) and (4.2) that for each observed (X, Y), the loss incurred from using the rule $\delta = I_A \in D_n$ is

$$L((X,Y), \delta(X)) = cn+dX+aY(1-I_A(X))+b(K-n-Y)I_A(X). \tag{4.4}$$

[2] Hall, P. and Heyde, C.C.(1980). "*Martingle limit theory and its application.*" Academic press. New York.

[3] Karunamuni, R.(1985). "*Empirical Bayes with sequential components.*" Ph. D. Thesis. Dept. of Statistics & Probability, Michigan State University.

[4] O'Bryan, T.(1972). "*Empirical Bayes results in the case of non-identical components.*" Ph. D. thesis. RM−306, Statistics and Probability, Michigan State University.

[5] O'Bryan, T. and Susarla, V.(1976). "*An empirical Bayes two−cation problem with non-identical components for a translated exponential distribution.*" Comm. Statist. 4(8) 767−775.

[6] Robbins, H.(1951). "*Asymptotically subminimax solutions of compound statistical decision problems.*" Proc.2nd Berkeley Symp. Math. Statist. Prob. 131−148, University of California Press.

[7] Robbins, H.(1956). "*The empirical Bayes approach to statistics.*" Proc. 3rd Berkeley Symp. Math. Statist. Prob. 1 157−163, University of California Press.

[8] Robbins, H.(1964). "*The empirical Bayes approach to statistical decision problems.*" Ann. Math. Statist. 351−20.