

11-9-2012

# Bayesian Estimation of Small Proportions Using Binomial Group Test

Shihua Luo

*Florida International University*, shihualuo2000@yahoo.com

**DOI:** 10.25148/etd.FI12112706

Follow this and additional works at: <http://digitalcommons.fiu.edu/etd>

---

## Recommended Citation

Luo, Shihua, "Bayesian Estimation of Small Proportions Using Binomial Group Test" (2012). *FIU Electronic Theses and Dissertations*. 744.

<http://digitalcommons.fiu.edu/etd/744>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

BAYESIAN ESTIMATION OF SMALL PROPORTIONS USING BINOMIAL  
GROUP TEST

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

in

STATISTICS

by

Shihua Luo

2012

To: Dean Kenneth Furton  
College of Arts and Sciences

This thesis, written by Shihua Luo, and entitled Bayesian Estimation of Small Proportions Using Binomial Group Test, having been approved in respect to style and intellectual content, is referred to you for judgement.

We have read this thesis and recommend that it be approved.

---

Florence George

---

Kai Huang, Co-Major Professor

---

Jie Mi, Co-Major Professor

Date of Defense: November 9, 2012

The thesis of Shihua Luo is approved.

---

Dean Kenneth Furton  
College of Arts and Sciences

---

Dean Lakshmi N. Reddi  
University Graduate School

Florida International University, 2012

## ACKNOWLEDGMENTS

First of all, I would like to express my sincere thanks to my major professor, Dr. Jie Mi for his patient guidance, enthusiasm, encouragement and friendship throughout this whole study. I couldn't finish my research without his great support. I would like to express deepest appreciation to my co-major professor, Dr. Kai Huang, for his technical support and patient guidance in Latex and Matlab.

I would also like to thank the member of my committee, Dr. Florence George for her time, valuable advice and great encouragement. In addition, I would like to thank the Department of Mathematics and Statistics, all the professors who supported and encouraged me throughout my life in FIU.

I wish to thank aunt Miao, Dr. Jie Mi's wife. She gave me substantial big help in my life. I also thank my classmates, friends. I am so happy I have them around me.

Finally, I wish to thank my family for their strong support and their absolute confidence in me.

ABSTRACT OF THE THESIS

BAYSIAN ESTIMATION OF SMALL PROPORTIONS USING BINOMIAL  
GROUP TEST

by

Shihua Luo

Florida International University, 2012

Miami, Florida

Professor Jie Mi, Co-Major Professor

Professor Kai Huang, Co-Major Professor

Group testing has long been considered as a safe and sensible relative to one-at-a-time testing in applications where the prevalence rate  $p$  is small. In this thesis, we applied Bayes approach to estimate  $p$  using Beta-type prior distribution. First, we showed two Bayes estimators of  $p$  from prior on  $p$  derived from two different loss functions. Second, we presented two more Bayes estimators of  $p$  from prior on  $\pi$  according to two loss functions. We also displayed credible and HPD interval for  $p$ . In addition, we did intensive numerical studies. All results showed that the Bayes estimator was preferred over the usual maximum likelihood estimator (MLE) for small  $p$ . We also presented the optimal  $\beta$  for different  $p, m$ , and  $k$ .

*Keywords:* Group Test, Bayes Estimator, Beta Distribution, MLE, MSE.

## TABLE OF CONTENTS

CHAPTER	PAGE
1 Introduction .....	1
2 Bayes Inferences from Prior on $p$ .....	7
2.1 Bayes Point Estimators of $p$ .....	7
2.2 Approximation to Bayes Estimators of $p$ .....	11
2.3 Comparison in the Special Case of $Y=0$ .....	13
2.4 Credible and HPD Interval for $p$ .....	15
3 Bayes Estimators of $p$ from Prior on $\pi$ .....	17
4 Numerical Studies .....	21
5 Conclusion .....	42
REFERENCES .....	43
APPENDIX .....	45

## LIST OF TABLES

TABLE .....	PAGE
1 Optimal $\beta$ For $k = 5, \alpha = 1, N = 10000$ .....	22
2 Optimal $\beta$ For $k = 10, \alpha = 1, N = 10000$ .....	23
3 Optimal $\beta$ For $k = 15, \alpha = 1, N = 10000$ .....	24
4 <i>MSE</i> For $p = 0.005, k = 5, m = 5, \alpha = 1, N = 10000$ .....	25
5 <i>MSE</i> For $p = 0.01, k = 5, m = 5, \alpha = 1, N = 10000$ .....	25
6 <i>MSE</i> For $p = 0.005, k = 5, m = 15, \alpha = 1, N = 10000$ .....	26
7 <i>MSE</i> For $p = 0.01, k = 5, m = 15, \alpha = 1, N = 10000$ .....	26
8 <i>MSE</i> For $p = 0.005, 0.006, 0.007, 0.008, k = 5, \alpha = 1, N = 10000$ .....	27
9 <i>MSE</i> For $p = 0.009, 0.01, 0.02, 0.03, k = 5, \alpha = 1, N = 10000$ .....	28
10 <i>MSE</i> For $p = 0.04, 0.05, 0.06, 0.07, k = 5, \alpha = 1, N = 10000$ .....	29
11 <i>MSE</i> For $p = 0.08, 0.09, 0.1, k = 5, \alpha = 1, N = 10000$ .....	30
12 <i>MSE</i> For $p = 0.005, 0.006, 0.007, 0.008, k = 10, \alpha = 1, N = 10000$ .....	31
13 <i>MSE</i> For $p = 0.009, 0.01, 0.02, 0.03, k = 10, \alpha = 1, N = 10000$ .....	32
14 <i>MSE</i> For $p = 0.04, 0.05, 0.06, 0.07, k = 10, \alpha = 1, N = 10000$ .....	33
15 <i>MSE</i> For $p = 0.08, 0.09, 0.1, k = 10, \alpha = 1, N = 10000$ .....	34
16 <i>MSE</i> For $p = 0.005, 0.006, 0.007, 0.008, k = 15, \alpha = 1, N = 10000$ .....	35
17 <i>MSE</i> For $p = 0.009, 0.01, 0.02, 0.03, k = 15, \alpha = 1, N = 10000$ .....	36
18 <i>MSE</i> For $p = 0.04, 0.05, 0.06, 0.07, k = 15, \alpha = 1, N = 10000$ .....	37
19 <i>MSE</i> For $p = 0.08, 0.09, 0.1, k = 15, \alpha = 1, N = 10000$ .....	38

## LIST OF FIGURES

FIGURE .....	PAGE
1 Optimal $\beta$ for $k = 5, m = 10$ .....	39
2 Optimal $\beta$ for $k = 5, m = 20$ .....	39
3 Optimal $\beta$ for $k = 5, m = 10$ or $20$ .....	40
4 Optimal $\beta$ for $k = 10, m = 10$ .....	40
5 Optimal $\beta$ for $k = 5$ or $10, m = 10$ .....	41



## 1 Introduction

Almost daily, the news media report the results of some poll. Pollsters often give some data, for example, the percentage of people in favor of the current president, the rate of people who smoke cigarettes, the death rate of a certain disease, the percent of people who have hepatitis C virus or a rare disease, the rate of contaminated fish in a river, the proportion of defective products, et cetera. All of these issues have a common property—the ratio of a part to a whole. How can we estimate this proportion  $p$ ? In estimating the proportion of subjects that have a certain characteristic in a population in which each individual can be classified into two categories, say with or without certain characteristic, one logical method of estimating  $p$  for the population is to use the proportion of 'successes' meaning with that characteristic in the sample. For example, how would we estimate the true fraction of all U.S. citizens who trust the president? We can estimate the  $p$  by calculating  $\hat{p}=x/n$ : where  $x$  is the number of people in the sample who trust the president and  $n$  is the number of people sampled. Experimenters used this method which is called the maximum likelihood estimator (MLE) for many years. But this method has some disadvantages when  $p$  is very small. For instance, if researchers want to estimate the prevalence of human immunodeficiency virus (HIV), the ratio  $\hat{p}$  is extremely small and so the estimation is not stable in using the MLE method. Then we need to test every sample unit to get the estimate of ratio  $p$ . In addition, the examination of the individual members of a large sample is an expensive and tedious work. To overcome these shortcomings, Dorfman (1943) first proposed the method of group testing and applied it to testing

blood samples for the syphilis antigen.

The idea of group testing is as follows. Assuming that a random sample of size  $n$  is drawn from a target population in which some units have certain characteristic that we are interested in, then we divide this sample into  $m$  groups with  $k$  units in each group. It is assumed that if at least one unit in a group has that characteristic (positive), then this group will be positive; on the other hand, if none of the units in a group is positive, then this group will be negative. The key assumption is valid for many situations in biological, environmental and medical studies. Let  $p$  denote the probability of a unit being positive in a certain population, then the probability of a group being positive, denoted as  $\pi$ , can be expressed in terms of  $p$ ,  $\pi = 1 - (1 - p)^k$ . Instead of directly estimating  $p$  from examining each unit in a sample, group testing will estimate  $\pi$  and then  $p$ .

Instead of conducting  $N$  chemical analyses are required to check out all members of a population of size  $N$ , Dorfman (1943) applied group testing for checking the syphilis antigen. Suppose that after the blood serum is drawn for each individual, all the blood samples are pooled in groups of five, and the group rather than the individual serum are subjected to chemical analysis. If none of the five sera contains syphilis antigen, the result of this group will be negative. If the result of a group shows positive, it means that this group includes at least one individual of syphilis antigen. Then the individuals making up the pool must be retested to decide which of the members are infected. He also discussed related costs for selected prevalence rates and optimum group sizes. He suggested that the prevalence rate be sufficiently

small to make worthwhile economics possible. Afterwards, many scientists used group testing in their study fields. Gibbs and Gower (1960) applied this method to estimate the frequency of success of attempts to transmit a virus disease from one plant to another. Chiang and Reeves (1962) used this approach to estimate the infection rates of mosquitoes. Bhattacharyya et al.(1979) employed group testing to give point estimates and confidence intervals for infection rates. Worlund and Taylor (1983) applied the procedure to evaluate the disease incidence in a large population of fish. Swallow (1985) used this method to estimate infection rates and probabilities of pathogen transmission by a single vector. To evaluate human immunodeficiency virus (HIV) seroprevalence in population surveys, pooled sera was used by Kline et al. (1989). Rodoni et al. (1994) applied a sequential batch testing procedure combined with Enzyme-Linked Immunosorbent Assay (ELISA) to estimate levels of virus incidence in Victoria cut-flower sim carnation. Hepworth (1996) constructed confidence intervals for proportion of infected unites in a population involving unequaled sized group. Gastwirth (2000) proposed that mutations in individual patients could be tested more fruitful by being checked in pools. Xie et al. (2001) used group testing to develops models and estimation procedures in order to obtain quantitative information from data in the process to discovery and development of a new drug. Katholi and Unnasch (2006) discussed the suitable sampling protocol about important experimental parameters for deciding infection rates. Because of its wide applications, group testing is also known as pooled testing, or composite sampling, in different fields. In recent years, more and more scholars recognized the advantage of the group testing

experimental design. It is well known that this method is economical, time saving, and efficient particularly when it is applied to estimating small proportion of a disease or some other rare attributes.

The traditional way to estimate  $\pi$  or  $p$  is using the method of maximum likelihood (MLE). The definition of MLE is as follows. If  $x_1, x_2, \dots, x_n$  are independent and identically distributed observation from a population with *pdf* or *pmf*  $f(\mathbf{x}; \theta)$ , where  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  is the fixed sample point, and  $\theta$  is the unknown parameter of interest, the likelihood function is defined as the product of all the *pdfs* or *pmfs*  $f(x_i; \theta)$ . For the sample point  $\mathbf{x}$ , let  $\theta(\mathbf{x})$  be a parameter value at which likelihood function attains its maximum as a function of  $\theta$ , with  $\mathbf{x}$  held fixed. Then  $\theta(\mathbf{x})$  is the *MLE* of  $\theta$ .

An alternative way to estimate  $\pi$  or  $p$  is applying the Bayes method. The Bayesian approach is totally different from the approach used in MLE analyse. In classical methods the parameter to be estimated is thought to be an unknown, but fixed quantity. The Bayesian approach treats the unknown parameter as a random variable. In many problems the investigator has some prior information about the unknown parameter. For instance, in sampling inspection, the quality engineer has some idea about the true defective rate, the unknown parameter, of a production process from past experience. The Bayesian approach assumes that this prior knowledge can be summarized in the form of a probability distribution on unknown parameter, called the prior distribution. This is a subjective distribution, based on the experimenter's belief, and is formulated before the data are seen. A sample is then taken

from a population indexed by the unknown parameter  $\theta$  and the prior distribution is updated with this sample information. The updated prior distribution is named the posterior distribution. The posterior distribution is now used to make inference about  $\theta$ . To use the Bayes method, we also need a loss function. The most common loss function is the square loss function. If we use this square loss function, a natural estimator for  $\theta$  is the mean of the posterior distribution. If necessary, of course, we can use alternative loss function, but the square loss function is the most popular one.

Usually, the posterior distribution will become very complex with each added measurement. Whereas, if we choose a conjugate prior, then the posterior distribution will be easier to obtain. A conjugate prior is defined as a prior distribution belonging to some parametric family, for which the resulted posterior distribution also belongs to the same family. This is an important property. Such prior distributions have been identified for some standard distributions  $f(x|\theta)$ . When we want Bayes estimator for proportion  $\theta = p$ , we assume the prior distribution on parameter  $\theta$  is Beta  $(\alpha, \beta)$ . The posterior distribution is also a Beta distribution function which makes the estimation of parameter  $p$  much more convenient.

Some researchers prefer to use classical Bayesian approaches because the Bayesian method gave a better point estimator of  $p$ , especially when  $p$  is very small. Gastwirth et al. (1991) suggested classical Bayesian approaches in group testing. Chaubey and Li (1995) researched the difference between *MLE* and Bayes method for estimation of binomial probability with group testing. They observed that the Bayes methodology

gave an alternative choices to the experimenter with possible reduction in cost as well as error. Chick (1996) has used this method to do his studies.

In the current study, we want to do further study on estimating  $p$  by using Bayes methods. We will confine our inference about  $p$  to the range  $(0, 0.1)$ . We will try two Bayesian estimators corresponding to two different loss functions with assumption that  $p$  has Beta prior distribution. In addition, we will try to obtain a credible and HPD interval for  $p$ . Assuming Beta prior on the probability  $\pi$  of a group being positive, we will investigate two Bayes estimators corresponding to two loss functions. The criterion for comparing the accurate of estimators is its *MSE* (mean squared error). The performance of the proposed estimators will be studied based on heavy Monte Carlo simulation. The proposed Bayes estimators will also be compared with other estimator existed in the literature.

## 2 Bayes Inferences from Prior on $p$

In this section, we will consider Bayes inferences about  $p$ . It is assumed that  $p$  has Beta prior distribution  $Beta(\alpha, \beta)$  with density function

$$f_P(p) = \frac{1}{B(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1}, \quad 0 < p < 1, \quad \alpha > 0, \quad \beta > 0.$$

**2.1 Bayes Point Estimators of  $p$**  Suppose that a random sample of size  $n = mk$  is obtained from a population of screened subjects, where  $m, k$  are positive integers. We assume that  $p$  is the probability that a randomly selected individual's test is positive. In this research our concern is the inference about small  $p$ , say  $0 < p \leq 0.1$ . The  $n$  subjects are classified into  $m$  groups each of which includes  $k$  subjects. Throughout this research it is assumed that any one of the  $m$  groups is tested as positive if and only if it includes at least one positive subject.

Let  $Y$  be the number of positive groups. It is known that  $Y \sim B(m, \pi)$  where  $\pi = 1 - (1-p)^k$ , i.e.,  $P(Y = y) = \binom{m}{y} \pi^y (1-\pi)^{m-y}$ .

Clearly, the joint distribution of  $(Y, p)$  is

$$\begin{aligned} f_{Y,P}(y, p) &= \frac{1}{B(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1} \binom{m}{y} \pi^y (1-\pi)^{m-y} \\ &= \frac{\binom{m}{y}}{B(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1} (1 - (1-p)^k)^y (1-p)^{k(m-y)} \end{aligned} \quad (2.1)$$

Thus the marginal distribution of  $Y$  is given as

$$f_Y(y) = \int_0^1 f_{Y,P}(y, p) dp$$

$$\begin{aligned}
&= \frac{\binom{m}{y}}{B(\alpha, \beta)} \int_0^1 p^{\alpha-1} (1-p)^{k(m-y)+\beta-1} (1-(1-p)^k)^y dp \\
&= \frac{\binom{m}{y}}{B(\alpha, \beta)} \int_0^1 p^{\alpha-1} (1-p)^{k(m-y)+\beta-1} \sum_{j=0}^y \binom{y}{j} (-1)^j (1-p)^{kj} dp \\
&= \binom{m}{y} (B(\alpha, \beta))^{-1} \sum_{j=0}^y (-1)^j \binom{y}{j} \int_0^1 p^{\alpha-1} (1-p)^{k(m+j-y)+\beta-1} dp \\
&= \binom{m}{y} (B(\alpha, \beta))^{-1} \sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, k(m+j-y) + \beta) \tag{2.2}
\end{aligned}$$

Therefore, the posterior distribution of  $p$  is

$$\begin{aligned}
f_{P|Y}(p|y) &= \frac{f_{Y,P}(y, p)}{f_Y(y)} \\
&= \frac{\binom{m}{y} (B(\alpha, \beta))^{-1} p^{\alpha-1} (1-p)^{k(m-y)+\beta-1} (1-(1-p)^k)^y}{\binom{m}{y} (B(\alpha, \beta))^{-1} \sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, k(m+j-y) + \beta)} \\
&= \frac{p^{\alpha-1} (1-p)^{k(m-y)+\beta-1} (1-(1-p)^k)^y}{\sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, k(m+j-y) + \beta)} \tag{2.3}
\end{aligned}$$

The Bayes estimator of  $p$  for the case of square loss function  $L_1(p, a) = (p-a)^2$  is derived by Chaubey and Li (1995) as

$$\hat{p}_{B1} = \frac{\sum_{j=0}^y \binom{y}{j} (-1)^j B(\alpha + 1, km + kj - ky + \beta)}{\sum_{j=0}^y \binom{y}{j} (-1)^j B(\alpha, km + kj - ky + \beta)}. \tag{2.4}$$

Specifically, if  $y = 0$ , then

$$\hat{p}_{B1} = \frac{\alpha}{n + \alpha + \beta}.$$

Now consider loss function  $L_2(p, a) = p^{-1}(p-a)^2$  and will derive the corresponding Bayes estimator  $\hat{p}_{B2}$ .



Note that if loss function has the form  $L(p, a) = w(p)(p - a)^2$ , then the corresponding Bayes estimator is given by

$$\hat{P}_{BL} = \frac{\int_0^1 w(p) p f_{P|Y}(p|y) dp}{\int_0^1 w(p) f_{P|Y}(p|y) dp} = \frac{E_{P|Y}(w(p)p|y)}{E_{P|Y}(w(p)|y)} \quad (2.5)$$

According to Equation (2.5) we have to evaluate  $\int_0^1 w(p) p f_{P|Y}(p|y) dp$  and  $\int_0^1 w(p) f_{P|Y}(p|y) dp$ .

For  $w(p) = p^{-1}$  in loss function  $L_2(p, a)$  we have

$$\int_0^1 w(p) p f_{P|Y}(p|y) dp = \int_0^1 p^{-1} p f_{P|Y}(p|y) dp = 1$$

and

$$\begin{aligned} \int_0^1 w(p) f_{P|Y}(p|y) dp &= \int_0^1 p^{-1} \frac{p^{\alpha-1} (1-p)^{k(m-y)+\beta-1} (1-(1-p)^k)^y}{\sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, k(m+j-y) + \beta)} dp \\ &= \frac{\int_0^1 p^{\alpha-2} (1-p)^{k(m-y)+\beta-1} (1-(1-p)^k)^y dp}{\sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, k(m+j-y) + \beta)} \end{aligned} \quad (2.6)$$

The value of the integral of the numerator in (2.6) is

$$\begin{aligned} &\int_0^1 p^{\alpha-2} (1-p)^{k(m-y)+\beta-1} (1-(1-p)^k)^y dp \\ &= \int_0^1 (1-u)^{\alpha-2} u^{k(m-y)+\beta-1} (1-u^k)^y du \quad (y > 0) \\ &= \sum_{i=0}^{k-1} \int_0^1 u^{k(m-y)+i+\beta-1} (1-u^k)^{y-1} (1-u)^{\alpha-1} du \end{aligned} \quad (2.7)$$

$$\begin{aligned}
&= \sum_{i=0}^{k-1} \int_0^1 u^{k(m-y)+i+\beta-1} \left( \sum_{j=0}^{y-1} \binom{y-1}{j} (-1)^j u^{kj} \right) (1-u)^{\alpha-1} du \\
&= \sum_{i=0}^{k-1} \sum_{j=0}^{y-1} (-1)^j \binom{y-1}{j} \int_0^1 u^{k(m-y)+i+\beta-1+kj} (1-u)^{\alpha-1} du \\
&= \sum_{i=0}^{k-1} \sum_{j=0}^{y-1} (-1)^j \binom{y-1}{j} B(km+kj+i+\beta-ky, \alpha)
\end{aligned}$$

Therefore, if  $y > 0$  we have

$$\hat{p}_{B2} = \frac{\sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, km+kj+\beta-ky)}{\sum_{i=0}^{k-1} \sum_{j=i}^{y-1} (-1)^j \binom{y-1}{j} B(km+kj+i+\beta-ky, \alpha)} \quad (2.8)$$

In the case  $y = 0$ , we let  $\alpha > 1$  and notice that equation (2.7) becomes

$$\int_0^1 p^{\alpha-2} (1-p)^{km+\beta-1} dp = B(\alpha-1, km+\beta), \quad (\alpha > 1).$$

Summarizing the above, we obtain

**Theorem 2.1.** *Let  $p$  have prior distribution  $B(\alpha, \beta)$ . For the loss function  $L_2(p, a) = p^{-1}(p-a)^2$  the Bayes estimator of  $p$ , denoted as  $\hat{p}_{B2}$ , is given as*

$$\hat{p}_{B2} = \begin{cases} \frac{\sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, km+kj+\beta-ky)}{\sum_{i=0}^{k-1} \sum_{j=i}^{y-1} (-1)^j \binom{y-1}{j} B(km+kj+\beta+i-ky, \alpha)} & \text{if } y > 0 \\ \frac{B(\alpha, km+\beta)}{B(\alpha-1, km+\beta)} & \text{if } y = 0 \text{ and } \alpha > 1. \end{cases}$$

## 2.2 Approximation to Bayes Estimators of $\mathbf{p}$

It is easy to show that for sufficiently large  $N$  it holds that

$$\frac{\Gamma(N+a)}{\Gamma(N+b)} \approx N^{a-b} \quad (2.9)$$

As it was mentioned in the previous section the group test is applied for the case of large sample size  $n = mk$ . Thus we could simplify the expression of  $\hat{p}_{B1}$  and  $\hat{p}_{B2}$

First we examine the expression (2.4) for  $\hat{p}_{B1}$ . We have

$$B(\alpha, km + kj - ky + \beta) = \frac{\Gamma(\alpha)\Gamma(km + kj - ky + \beta)}{\Gamma(km + kj - ky + \beta + \alpha)} \approx \Gamma(\alpha)(km + kj - ky + \beta)^{-\alpha}$$

since  $y$  is usually a small number and  $n = km$  is large. Similarly we have

$$B(\alpha + 1, km + kj - ky + \beta) \approx \Gamma(\alpha + 1)(km + kj - ky + \beta)^{-\alpha-1}$$

Hence we obtain

$$\begin{aligned} \hat{p}_{B1} &= \frac{\sum_{j=0}^y \binom{y}{j} (-1)^j B(\alpha + 1, km + kj - ky + \beta)}{\sum_{j=0}^y \binom{y}{j} (-1)^j B(\alpha, km + kj - ky + \beta)} \\ &\approx \frac{\Gamma(\alpha + 1) \sum_{j=0}^y \binom{y}{j} (-1)^j (km + kj - ky + \beta)^{-\alpha-1}}{\Gamma(\alpha) \sum_{j=0}^y \binom{y}{j} (-1)^j (km + kj - ky + \beta)^{-\alpha}} \end{aligned}$$

$$\begin{aligned}
& \alpha \sum_{j=0}^y (-1)^j \binom{y}{j} (km + kj - ky + \beta)^{-\alpha-1} \\
&= \frac{\alpha \sum_{j=0}^y (-1)^j \binom{y}{j} (km + kj - ky + \beta)^{-\alpha-1}}{\sum_{j=0}^y (-1)^j \binom{y}{j} (km + kj - ky + \beta)^{-\alpha}} \equiv \hat{p}_{B'1}
\end{aligned} \tag{2.10}$$

where  $\hat{p}_{B'1}$  is an approximation to the Bayes estimator  $\hat{p}_{B1}$ .

As for  $\hat{p}_{B2}$  notice that if  $y \geq 1$ , then

$$\begin{aligned}
B(km + kj + i + \beta - ky, \alpha) &= \frac{\Gamma(\alpha)\Gamma(km + kj + i + \beta - ky)}{\Gamma(km + kj + i + \beta - ky + \alpha)} \\
&\approx \Gamma(\alpha)(km + kj + i + \beta - ky)^{-\alpha}
\end{aligned}$$

Therefore, if  $y \geq 1$  it holds from Theorem 2.1 that

$$\begin{aligned}
\hat{p}_{B2} &\approx \frac{\sum_{j=0}^y (-1)^j \binom{y}{j} \Gamma(\alpha)(km + kj - ky + \beta)^{-\alpha}}{\sum_{i=0}^{k-1} \sum_{j=0}^{y-1} (-1)^j \binom{y-1}{j} \Gamma(\alpha)(km + kj + i + \beta - ky)^{-\alpha}} \\
&= \frac{\sum_{j=0}^y (-1)^j \binom{y}{j} (km + kj - ky + \beta)^{-\alpha}}{\sum_{i=0}^{k-1} \sum_{j=0}^{y-1} (-1)^j \binom{y-1}{j} (km + kj + i + \beta - ky)^{-\alpha}} \equiv \hat{p}_{B'2}
\end{aligned} \tag{2.11}$$

On the other hand if  $y = 0$  and  $\alpha > 1$ , then we have

$$\begin{aligned}
\hat{p}_{B2} &= \frac{B(\alpha, km + \beta)}{B(\alpha - 1, km + \beta)} = \frac{\Gamma(\alpha)\Gamma(km + \beta)}{\Gamma(km + \alpha + \beta)} \frac{\Gamma(km + \alpha + \beta - 1)}{\Gamma(\alpha - 1)\Gamma(km + \beta)} \\
&= \frac{(\alpha - 1)\Gamma(km + \alpha + \beta - 1)}{\Gamma(km + \alpha + \beta)} = \frac{\alpha - 1}{km + \alpha + \beta - 1} = \frac{\alpha - 1}{n + \alpha + \beta - 1} \quad \text{if } \alpha > 1.
\end{aligned} \tag{2.12}$$

From (2.11) and (2.12) we see that the Bayes estimator  $\hat{p}_{B2}$  can be approximated

by

$$\hat{p}_{B'2} = \begin{cases} \frac{\sum_{j=0}^y (-1)^j \binom{y}{j} (km+kj-ky+\beta)^{-\alpha}}{\sum_{i=0}^{k-1} \sum_{j=0}^{y-1} (-1)^j \binom{y-1}{j} (km+kj+i+\beta-ky)^{-\alpha}}, & \text{if } y \geq 1; \\ \frac{\alpha-1}{n+\alpha+\beta-1}, & \text{if } y = 0 \text{ and } \alpha > 1. \end{cases} \quad (2.13)$$

### 2.3 Comparison in the special case of $Y=0$

Note the fact that the event ( $Y = 0$ ) is the same as event ( $X = 0$ ) where  $X$  is the number of positive subjects when subjects are tested individually. So it is interesting to compare the Bayes estimators based on  $X = 0$  and  $Y = 0$ . Clearly  $X \sim B(n, p) = B(mk, p)$ . Assuming that the proportion  $p$  has  $Beta(\alpha, \beta)$  as its prior distribution, it is well known that the posterior distribution of  $p$  given  $X = x$  is  $Beta(x + \alpha, n - x + \beta)$ .

Suppose that square loss function  $L_1(p, a) = (p - a)^2$  is used, then the Bayes estimator  $\hat{p}_{B1}^*$  based on  $X = 0$  is

$$\hat{p}_{B1}^* = E_{P|X}(p|x = 0) = \frac{\alpha}{n + \alpha + \beta}$$

In the mean time the Bayes estimator  $\hat{p}_{B1}$  based on  $Y = 0$  is exactly the same as  $\hat{p}_{B1}^*$ .(bottom of p8)

Now consider the loss function  $L_2(p, a) = p^{-1}(p - a)^2$ . For this loss function the Bayes estimator  $\hat{p}_{B2}^*$  based on  $X$  can be obtained as follows. We have

$$\int_0^1 w(p) p f_{P|X}(p|x) dp = \int_0^1 f_{P|X}(p|x) dp = 1$$

and

$$\begin{aligned}\int_0^1 w(p)f_{P|X}(p|x)dp &= \int_0^1 p^{-1}(B(x+\alpha, n-x+\beta))^{-1}p^{x+\alpha-1}(1-p)^{n-x+\beta-1}dp \\ &= (B(x+\alpha, n-x+\beta))^{-1} \int_0^1 p^{x+\alpha-2}(1-p)^{n-x+\beta-1}dp\end{aligned}$$

Consider two cases. If  $x \geq 1$  which is equivalent to  $y \geq 1$ . In this case

$$\begin{aligned}\int_0^1 w(p)f_{P|X}(p|x)dp &= \left( \frac{B(x+\alpha, n-x+\beta)}{B(x+\alpha-1, n-x+\beta)} \right)^{-1} \\ &= \left( \frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(n+\alpha+\beta)} \frac{\Gamma(n+\alpha+\beta-1)}{\Gamma(x+\alpha-1)\Gamma(n-x+\beta)} \right)^{-1} \\ &= \left( \frac{x+\alpha-1}{n+\alpha+\beta-1} \right)^{-1}\end{aligned}$$

Thus

$$\hat{p}_{B2}^* = \frac{x+\alpha-1}{n+\alpha+\beta-1}$$

However, if  $x = 0$ , then it must be assumed that  $\alpha > 1$  and so

$$\begin{aligned}\int_0^1 w(p)f_{P|X}(p|0)dp &= (B(\alpha, n+\beta))^{-1} \int_0^1 p^{\alpha-2}(1-p)^{n+\beta-1}dp \\ &= \frac{B(\alpha-1, n+\beta)}{B(\alpha, n+\beta)} \\ &= \frac{\Gamma(\alpha-1)\Gamma(n+\beta)}{\Gamma(n+\alpha+\beta-1)} \frac{\Gamma(n+\alpha+\beta)}{\Gamma(\alpha)\Gamma(n+\beta)} \\ &= \frac{n+\alpha+\beta-1}{\alpha-1}\end{aligned}$$

Consequently, it follows that

$$\hat{p}_{B2}^* = \frac{\alpha - 1}{n + \alpha + \beta - 1}, \quad \text{if } x = 0 \text{ and } \alpha > 1.$$

Therefore

$$\hat{p}_{B2}^* = \begin{cases} \frac{x+\alpha-1}{n+\alpha+\beta-1} & \text{if } x \geq 1; \\ \frac{\alpha-1}{n+\alpha+\beta-1} & \text{if } x = 0 \text{ and } \alpha > 1. \end{cases}$$

From this and equation (2.13), i.e.,

$$\hat{p}_{B'2} = \frac{\alpha - 1}{n + \alpha + \beta - 1} \quad \text{if } y = 0 \text{ and } \alpha > 1.$$

we see that in this case  $\hat{p}_{B'2}$  and  $\hat{p}_{B2}^*$  are exactly the same when there is no positive subject in the entire sample.

## 2.4 Credible and HPD Interval for $p$

Suppose that  $p$  has prior distribution  $Beta(\alpha, \beta)$ . From equation (2.3) it is known that the posterior distribution of  $p$  given  $Y = y$  is

$$f_{P|Y}(p|y) = \frac{p^{\alpha-1}(1-p)^{k(m-y)+\beta-1}(1-(1-p)^k)^y}{\sum_{j=0}^y \binom{y}{j} (-1)^j B(\alpha, k(m+j-y) + \beta)}$$

Then for any set  $A \subset (0, 1)$ , the credibal probability of  $A$  is

$$P(p \in A) = \int_A f_{P|Y}(p|y) dp$$

and  $A$  is a credible set for  $p$ . By numerical computation it is not difficult to obtain a  $1 - \gamma$  credible interval for  $p$ .

In a special case when  $\alpha = 1$ , i.e.,  $p$  has  $Beta(1, \beta)$  as its prior distribution, by equation (2.2) the marginal distribution of  $Y$  has density

$$\begin{aligned} f_Y(y) &= \int_0^1 \beta(1-p)^{\beta-1} \binom{m}{y} (1 - (1-p)^k)^y (1-p)^{k(m-y)} dp \\ &= \frac{\beta}{k} \binom{m}{y} B\left(m - y + \frac{\beta}{k}, y + 1\right) \end{aligned} \quad (2.14)$$

and consequently the posterior distribution of  $p$  given  $Y = y$  has density

$$\begin{aligned} f_{P|Y}(p|y) &= \frac{\beta \binom{m}{y} (1-p)^{k(m-y)+\beta-1} (1 - (1-p)^k)^y}{\beta k^{-1} \binom{m}{y} B\left(m - y + \frac{\beta}{k}, y + 1\right)} \\ &= \frac{k(1-p)^{k(m-y)+\beta-1} (1 - (1-p)^k)^y}{B\left(m - y + \frac{\beta}{k}, y + 1\right)} \end{aligned} \quad (2.15)$$

We further let  $\beta \geq 1$ . The assumption is reasonable because we are dealing with small  $p$  and the majority of the probability distribution  $Beta(1, \beta)$  is close to zero if  $\beta$  is relatively large. With the assumption it holds that  $m - y + (\beta - 1)/k \geq 0$ . It can be shown easily that the posterior  $f_{P|Y}(p|y)$  is unimodal. Therefore, the HPD interval for  $p$  can be obtained by

$$\{p : f_{P|Y}(p|y) \geq a\} \quad \text{where} \quad \int_{\{p: f_{P|Y}(p|y) \geq a\}} f_{P|Y}(p|y) dp = 1 - \gamma$$

for any given  $\gamma \in (0, 1)$ .

For the case  $0 < \beta < 1$ , the posterior density  $f_{P|Y}(p|y)$  strictly decreases in  $p \in (0, 1)$ , so the HPD interval for  $p$  can be obtained as



$$\{p : f_{P|Y}(p|y) \geq a\} \quad \text{where} \quad \int_0^a f_{P|Y}(p|y) dp = 1 - \gamma.$$

### 3 Bayes Estimators of $p$ from Prior on $\pi$

In the present section we will assume  $\pi \equiv 1 - (1 - p)^k$  has a  $Beta(\alpha, \beta)$  prior distribution. On the basis of the observation  $Y = y$  we first calculate the Bayes estimator  $\hat{\pi}_B$ , then derive the Bayes estimator of  $p$ . Following Chaubey and Li (1995) this procedure is called the *indirect Bayes procedure*, the produced estimators is called the *indirect Bayes estimators* and denoted as  $\hat{p}_{IB1}$  or  $\hat{p}_{IB2}$  depending on the loss function  $L_1(p, a) = (p - a)^2$  and  $L_2(p, a) = p^{-1}(p - a)^2$ .

Obviously the posterior distribution of  $\pi$  given  $Y = y$  is  $Beta(y + \alpha, m - y + \beta)$ , i.e.,

$$f_{\pi|Y}(\pi|y) = (B(y + \alpha, m - y + \beta))^{-1} \pi^{y+\alpha-1} (1 - \pi)^{m-y+\beta-1}$$

From  $\pi = 1 - (1 - p)^k = \pi(p)$  we see that  $p = p(\pi) = 1 - (1 - \pi)^{1/k}$ .

Consider loss function  $L(p, a) = w(p)(p - a)^2$ . It is easy to see that the corresponding indirect Bayes estimator  $\hat{p}_{IB}$  is given as

$$\hat{p}_{IB} = \frac{E(pw(p)|Y = y)}{E(w(p)|Y = y)}$$

For the case  $L_1(p, a) = (p - a)^2$  the indirect Bayes estimator  $\hat{p}_{IB1}$  is given in Chaubey and Li (1995) as

$$\hat{p}_{IB1} = 1 - \frac{\Gamma(m + \alpha + \beta)\Gamma(m - y + \beta + 1/k)}{\Gamma(m - y + \beta)\Gamma(m + \alpha + \beta + 1/k)} \quad (3.1)$$

which can be further approximated as

$$\hat{p}_{IB'1} = 1 - \left( \frac{m + \beta - y}{m + \alpha + \beta} \right)^{1/k} \quad (3.2)$$

Now let us consider loss function  $L_2(p, a) = p^{-1}(p - a)^2$ . Clearly

$$E(pw(p)|Y = y) = \int_0^1 pp^{-1}f_{\pi|Y}(\pi|y)d\pi = \int_0^1 f_{\pi|Y}(\pi|y)d\pi = 1$$

and

$$\begin{aligned} E(w(p)|Y = y) &= \int_0^1 p^{-1}f_{\pi}(\pi|y)d\pi \\ &= \int_0^1 p^{-1}(B(y + \alpha, m - y + \beta))^{-1}\pi^{y+\alpha-1}(1 - \pi)^{m-y+\beta-1}d\pi \\ &= \frac{\int_0^1 (1 - (1 - \pi)^{1/k})^{-1}\pi^{y+\alpha-1}(1 - \pi)^{m-y+\beta-1}d\pi}{B(y + \alpha, m - y + \beta)} \end{aligned} \quad (3.3)$$

To evaluate the integral on the right side of (3.3) we let  $u = (1 - \pi)^{1/k}$  and

have

$$\begin{aligned} &\int_0^1 (1 - (1 - \pi)^{1/k})^{-1}\pi^{y+\alpha-1}(1 - \pi)^{m-y+\beta-1}d\pi \\ &= \int_0^1 (1 - u)^{-1}(1 - u^k)^{y+\alpha-1}u^{k(m-y+\beta-1)}ku^{k-1}du \end{aligned}$$

$$\begin{aligned}
&= k \sum_{i=0}^{k-1} \int_0^1 u^{k(m-y+\beta)+i-1} (1-u^k)^{y+\alpha-2} du \\
&= \sum_{i=0}^{k-1} \int_0^1 u^{k(m-y+\beta)+i-k} (1-u^k)^{y+\alpha-2} k u^{k-1} du \\
&= \sum_{i=0}^{k-1} \int_0^1 v^{m-y+\beta+i/k-1} (1-v)^{y+\alpha-2} dv \\
&= \sum_{i=0}^{k-1} B\left(m-y+\beta+\frac{i}{k}, y+\alpha-1\right),
\end{aligned}$$

here it is required that  $\alpha > 1$  if  $y = 0$ .

Therefore, the indirect Bayes estimator  $\hat{p}_{IB2}$  based on observation  $Y = y$  is

$$\hat{p}_{IB2} = \frac{B(y+\alpha, m-y+\beta)}{\sum_{i=0}^{k-1} B\left(m-y+\beta+\frac{i}{k}, y+\alpha-1\right)} \quad (3.4)$$

where  $\alpha > 1$  if  $y = 0$ .

We can rewrite  $\hat{p}_{IB2}$  as

$$\begin{aligned}
\hat{p}_{IB2} &= \frac{\Gamma(y+\alpha)\Gamma(m-y+\beta)/\Gamma(m+\alpha+\beta)}{\sum_{i=0}^{k-1} \Gamma\left(m-y+\beta+\frac{i}{k}\right)\Gamma(y+\alpha-1)/\Gamma\left(m+\alpha+\beta+\frac{i}{k}-1\right)} \\
&= (y+\alpha-1) \frac{\Gamma(m-y+\beta)/\Gamma(m+\alpha+\beta)}{\sum_{i=0}^{k-1} \Gamma\left(m-y+\beta+\frac{i}{k}\right)/\Gamma\left(m+\alpha+\beta+\frac{i}{k}-1\right)} \\
&= (y+\alpha-1) \sum_{i=0}^{k-1} \frac{\Gamma(m-y+\beta)}{\Gamma\left(m-y+\beta+\frac{i}{k}\right)} \cdot \frac{\Gamma\left(m+\alpha+\beta+\frac{i}{k}-1\right)}{\Gamma(m+\alpha+\beta)} \quad (3.5)
\end{aligned}$$

From (3.5) we see that  $\hat{p}_{IB2}$  can be approximated by

$$\hat{p}_{IB'2} = \frac{y+\alpha-1}{\sum_{i=0}^{k-1} (m-y+\beta)^{i/k} (m+\alpha+\beta)^{-i/k+1}}$$

$$= \frac{y + \alpha - 1}{m + \alpha + \beta} \frac{1}{\sum_{i=0}^{k-1} \left( \frac{m-y+\beta}{m+\alpha+\beta} \right)^{i/k}} \quad (3.6)$$

here  $\alpha > 1$  if  $y = 0$ .

## 4 Numerical Studies

A MATLAB simulation is performed in order to analyze the behaviors of different estimators under several combinations. I chose  $m = 5, 10, 15, 20, 25$  and  $k = 5, 10, 15$ . Fifteen cases of  $p$  between 0.005 and 0.1 are considered. First, I simulate a sample that is from an independent, identically Bernoulli distribution and divided the observations into  $m$  groups with  $k$  units in each group. I use this ‘new’ grouped sample to calculate the value of  $MLE$ . Second, I use the same sample to calculate the values of Bayes estimator and approximate Bayes estimator under different  $\beta$ s. The process is repeated 10000 times. Then I will obtain the  $MSE$  of these estimators.

Using the  $MSE$  as a criterion. Table 1-3 display the optimal  $\beta$ s which yield the smallest  $MSE$  of Bayes estimator associated with the square loss function. From these table, we can see that the optimal  $\beta$  decreases with the increase of  $p$ ; the optimal  $\beta$  increases with the increase of  $m$  and  $k$ . Figures 1-5 also indicate these trends.

Tables 4-7 show how I got the optimal  $\beta$ s. Tables 8-19 compare the  $MSE$  of Bayes estimator under the optimal  $\beta$ s with the approximate Bayes estimator as well as  $MLE$ . From these tables, we can draw a conclusion that the Bayes estimator is the best one as far as  $MSE$  is concerned.  $MLE$  is not a good estimator when  $p$  is small. Moreover, the approximate Bayes estimator is also much better than the  $MLE$ .

Table 1:  $k = 5, \alpha = 1, N = 10000$

$k = 5$					
p	m				
	5	10	15	20	25
0.005	222	237	252	269	274
0.006	187	205	218	229	239
0.007	163	180	192	201	209
0.008	144	163	168	181	187
0.009	130	146	157	166	171
0.01	121	132	143	152	155
0.02	66	76	79	83	85
0.03	47	54	57	58	59
0.04	37	42	44	45	47
0.05	31	34	36	37	38
0.06	26	29	31	32	32
0.07	23	26	27	27	29
0.08	21	23	24	25	25
0.09	19	21	22	22	22
0.1	18	19	20	21	21

This is the optimal  $\beta$  based on the smallest  $MSE$  of Bayes estimator by using  $k = 5$  with combinations of different  $m$  and  $p$ .

Table 2:  $k = 10, \alpha = 1, N = 10000$

$k = 10$					
p	m				
	5	10	15	20	25
0.005	237	266	292	303	314
0.006	207	233	247	264	269
0.007	183	199	216	229	236
0.008	163	183	193	202	213
0.009	147	166	176	189	192
0.01	135	151	163	167	176
0.02	77	86	91	94	97
0.03	55	62	64	66	67
0.04	43	48	51	54	57
0.05	37	40	41	44	47
0.06	32	35	37	37	37
0.07	27	32	33	34	36
0.08	25	28	30	30	32
0.09	22	25	27	27	28
0.1	20	24	25	26	27

This is the optimal  $\beta$  based on the smallest  $MSE$  of Bayes estimator by using  $k = 10$  with combinations of different  $m$  and  $p$ .

Table 3:  $k = 15, \alpha = 1, N = 10000$

$k = 15$					
p	m				
	5	10	15	20	25
0.005	259	290	313	321	327
0.006	221	255	268	273	286
0.007	194	221	233	241	251
0.008	175	198	208	218	223
0.009	162	179	191	199	203
0.01	147	163	173	181	186
0.02	84	93	99	103	107
0.03	61	67	71	72	74
0.04	48	54	56	58	63
0.05	40	45	48	49	54
0.06	35	40	42	44	48
0.07	31	35	37	39	39
0.08	27	32	34	36	37
0.09	24	29	32	33	33
0.1	21	26	29	31	33

This is the optimal  $\beta$  based on the smallest  $MSE$  of Bayes estimator by using  $k = 15$  with combinations of different  $m$  and  $p$ .



Table 4:  $k = 5, m = 5, p = 0.005, \alpha = 1, N = 10000$

$\beta$	215	216	217	218	219
$MSE(\hat{p}_{B1})$	2.2647E-06	2.2594E-06	2.2549E-06	2.2513E-06	2.2485E-06
$MSE(\hat{p}_{B'1})$	2.2895E-06	2.2830E-06	2.2773E-06	2.2725E-06	2.2686E-06
$\beta$	220	221	222	223	224
$MSE(\hat{p}_{B1})$	2.2465E-06	2.2453E-06	2.2448E-06	2.2451E-06	2.2461E-06
$MSE(\hat{p}_{B'1})$	2.2655E-06	2.2632E-06	2.2617E-06	2.2610E-06	2.2610E-06
$\beta$	225	226	227	228	229
$MSE(\hat{p}_{B1})$	2.2479E-06	2.2503E-06	2.2534E-06	2.2572E-06	2.2616E-06
$MSE(\hat{p}_{B'1})$	2.2617E-06	2.2632E-06	2.2654E-06	2.2682E-06	2.2717E-06

Table 5:  $k = 5, m = 5, p = 0.01, \alpha = 1, N = 10000$

$\beta$	114	115	116	117	118
$MSE(\hat{p}_{B1})$	1.3915E-05	1.3859E-05	1.3813E-05	1.3776E-05	1.3749E-05
$MSE(\hat{p}_{B'1})$	1.4185E-05	1.4111E-05	1.4048E-05	1.3996E-05	1.3954E-05
$\beta$	119	120	121	122	123
$MSE(\hat{p}_{B1})$	1.3732E-05	1.3723E-05	1.3722E-05	1.3730E-05	1.3745E-05
$MSE(\hat{p}_{B'1})$	1.3921E-05	1.3898E-05	1.3883E-05	1.3878E-05	1.3880E-05
$\beta$	124	125	126	127	128
$MSE(\hat{p}_{B1})$	1.3768E-05	1.3798E-05	1.3834E-05	1.3877E-05	1.3927E-05
$MSE(\hat{p}_{B'1})$	1.3891E-05	1.3909E-05	1.3934E-05	1.3966E-05	1.4005E-05

The comparison of  $MSE$  of Bayes estimator and approximation to the Bayes estimator under different  $\beta$ s.

Table 6:  $k = 5, m = 15, p = 0.005, \alpha = 1, N = 10000$

$\beta$	245	246	247	248	249
$MSE(\hat{p}_{B1})$	4.0451E-06	4.0426E-06	4.0404E-06	4.0388E-06	4.0375E-06
$MSE(\hat{p}_{B'1})$	4.0718E-06	4.0684E-06	4.0655E-06	4.0630E-06	4.0610E-06
$\beta$	250	251	252	253	254
$MSE(\hat{p}_{B1})$	4.0366E-06	4.0362E-06	4.0361E-06	4.0365E-06	4.0372E-06
$MSE(\hat{p}_{B'1})$	4.0594E-06	4.0582E-06	4.0574E-06	4.0570E-06	4.0570E-06
$\beta$	255	256	257	258	259
$MSE(\hat{p}_{B1})$	4.0383E-06	4.0398E-06	4.0416E-06	4.0438E-06	4.0463E-06
$MSE(\hat{p}_{B'1})$	4.0575E-06	4.0583E-06	4.0594E-06	4.0610E-06	4.0629E-06

Table 7:  $k = 5, m = 15, p = 0.01, \alpha = 1, N = 10000$

$\beta$	136	137	138	139	140
$MSE(\hat{p}_{B1})$	1.9446E-05	1.9422E-05	1.9401E-05	1.9385E-05	1.9372E-05
$MSE(\hat{p}_{B'1})$	1.9675E-05	1.9641E-05	1.9612E-05	1.9587E-05	1.9566E-05
$\beta$	141	142	143	144	145
$MSE(\hat{p}_{B1})$	1.9364E-05	1.9359E-05	1.9357E-05	1.9359E-05	1.9364E-05
$MSE(\hat{p}_{B'1})$	1.9548E-05	1.9535E-05	1.9525E-05	1.9519E-05	1.9517E-05
$\beta$	146	147	148	149	150
$MSE(\hat{p}_{B1})$	1.9373E-05	1.9385E-05	1.9399E-05	1.9417E-05	1.9438E-05
$MSE(\hat{p}_{B'1})$	1.9518E-05	1.9523E-05	1.9531E-05	1.9541E-05	1.9556E-05

The comparison of  $MSE$  of Bayes estimator and approximation to the Bayes estimator under different  $\beta$ s.

Table 8:  $k = 5$ ,  $p = 0.005, 0.006, 0.007, 0.008$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.005$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	2.2448E-06	3.3817E-06	4.0361E-06	4.6045E-06	4.7323E-06
$MSE(\hat{p}_{B'1})$	2.2617E-06	3.4029E-06	4.0574E-06	4.6270E-06	4.7537E-06
$MSE(\hat{p}_{MLE})$	2.4104E-04	1.0604E-04	6.8361E-05	5.3956E-05	4.0660E-05
$p = 0.006$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	3.6664E-06	5.3663E-06	6.2593E-06	6.8083E-06	6.9873E-06
$MSE(\hat{p}_{B'1})$	3.6992E-06	5.4038E-06	6.2974E-06	6.8459E-06	7.0234E-06
$MSE(\hat{p}_{MLE})$	2.9072E-04	1.3120E-04	8.4518E-05	6.3058E-05	4.9948E-05
$p = 0.007$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	5.4291E-06	7.9326E-06	9.1134E-06	9.3897E-06	9.6176E-06
$MSE(\hat{p}_{B'1})$	5.4817E-06	7.9946E-06	9.1770E-06	9.4481E-06	9.6735E-06
$MSE(\hat{p}_{MLE})$	3.3223E-04	1.5615E-04	1.0109E-04	7.2744E-05	5.7832E-05
$p = 0.008$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	7.5515E-06	1.0967E-05	1.1873E-05	1.2491E-05	1.2826E-05
$MSE(\hat{p}_{B'1})$	7.6302E-06	1.1058E-05	1.1961E-05	1.2576E-05	1.2907E-05
$MSE(\hat{p}_{MLE})$	3.6853E-04	1.8285E-04	1.0860E-04	8.4066E-05	6.6827E-05

Table-8 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 5$  with different combinations of  $p = 0.005, 0.006, 0.007, 0.008$  and  $m = 5, 10, 15, 20, 25$ .

Table 9:  $k = 5$ ,  $p = 0.009, 0.01, 0.02, 0.03$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.009$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.0420E-05	1.3972E-05	1.5494E-05	1.6192E-05	1.5946E-05
$MSE(\hat{p}_{B'1})$	1.0544E-05	1.4098E-05	1.5617E-05	1.6313E-05	1.6062E-05
$MSE(\hat{p}_{MLE})$	4.3057E-04	1.9751E-04	1.2918E-04	9.7241E-05	7.5947E-05
$p = 0.01$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.3722E-05	1.7754E-05	1.9357E-05	2.0338E-05	1.9712E-05
$MSE(\hat{p}_{B'1})$	1.3883E-05	1.7927E-05	1.9525E-05	2.0507E-05	1.9866E-05
$MSE(\hat{p}_{MLE})$	4.9391E-04	2.1441E-04	1.4237E-04	1.0983E-04	8.3501E-05
$p = 0.02$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	7.1288E-05	7.9758E-05	7.4927E-05	7.1622E-05	6.6506E-05
$MSE(\hat{p}_{B'1})$	7.2702E-05	8.1032E-05	7.6108E-05	7.2738E-05	6.7577E-05
$MSE(\hat{p}_{MLE})$	9.9164E-04	4.6150E-04	2.8147E-04	2.1224E-04	1.6707E-04
$p = 0.03$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.7145E-04	1.6946E-04	1.5602E-04	1.3832E-04	1.2629E-04
$MSE(\hat{p}_{B'1})$	1.7615E-04	1.7343E-04	1.5965E-04	1.4169E-04	1.2943E-04
$MSE(\hat{p}_{MLE})$	1.5344E-03	6.9309E-04	4.4324E-04	3.1779E-04	2.5145E-04

Table-9 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 5$  with different combinations of  $p = 0.009, 0.01, 0.02, 0.03$  and  $m = 5, 10, 15, 20, 25$ .

Table 10:  $k = 5$ ,  $p = 0.04, 0.05, 0.06, 0.07$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.04$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	3.1116E-04	2.8077E-04	2.4544E-04	2.1612E-04	1.9608E-04
$MSE(\hat{p}_{B'1})$	3.2268E-04	2.9046E-04	2.5333E-04	2.2362E-04	2.0262E-04
$MSE(\hat{p}_{MLE})$	2.4326E-03	9.4918E-04	5.8556E-04	4.3233E-04	3.4986E-04
$p = 0.05$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	4.7085E-04	4.1457E-04	3.5361E-04	2.9887E-04	2.6260E-04
$MSE(\hat{p}_{B'1})$	4.9176E-04	4.3175E-04	3.7010E-04	3.1185E-04	2.7459E-04
$MSE(\hat{p}_{MLE})$	3.4564E-03	1.1715E-03	7.6870E-04	5.4140E-04	4.3144E-04
$p = 0.06$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	6.6019E-04	5.4481E-04	4.6408E-04	3.9214E-04	3.4498E-04
$MSE(\hat{p}_{B'1})$	6.9704E-04	5.7466E-04	4.8941E-04	4.1474E-04	3.6812E-04
$MSE(\hat{p}_{MLE})$	4.4409E-03	1.4282E-03	9.3129E-04	6.8010E-04	5.4684E-04
$p = 0.07$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	8.6912E-04	6.9911E-04	5.8303E-04	4.8215E-04	4.1385E-04
$MSE(\hat{p}_{B'1})$	9.2762E-04	7.4170E-04	6.2164E-04	5.1935E-04	4.4459E-04
$MSE(\hat{p}_{MLE})$	6.2587E-03	1.7286E-03	1.1043E-03	7.9370E-04	6.3559E-04

Table-10 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 5$  with different combinations of  $p = 0.04, 0.05, 0.06, 0.07$  and  $m = 5, 10, 15, 20, 25$ .

Table 11:  $k = 5, p = 0.08, 0.09, 0.1, \alpha = 1, N = 10000$

$p = 0.08$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.0716E-03	8.7340E-04	6.9105E-04	5.7976E-04	4.9486E-04
$MSE(\hat{p}_{B'1})$	1.1431E-03	9.4426E-04	7.4734E-04	6.3048E-04	5.4688E-04
$MSE(\hat{p}_{MLE})$	7.8350E-03	2.1054E-03	1.2674E-03	9.3950E-04	7.2941E-04
$p = 0.09$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.3188E-03	1.0258E-03	8.1458E-04	7.0277E-04	5.7674E-04
$MSE(\hat{p}_{B'1})$	1.4227E-03	1.1176E-03	8.9544E-04	7.8028E-04	6.4741E-04
$MSE(\hat{p}_{MLE})$	1.1193E-02	2.3781E-03	1.4831E-03	1.1029E-03	8.3620E-04
$p = 0.1$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.6198E-03	1.2040E-03	9.5955E-04	8.0905E-04	6.7701E-04
$MSE(\hat{p}_{B'1})$	1.7472E-03	1.3411E-03	1.0694E-03	9.0490E-04	7.7305E-04
$MSE(\hat{p}_{MLE})$	1.6287E-02	2.9874E-03	1.6952E-03	1.2618E-03	9.5998E-04

Table-11 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 5$  with different combinations of  $p = 0.08, 0.09, 0.1$  and  $m = 5, 10, 15, 20, 25$ .

Table 12:  $k = 10$ ,  $p = 0.005, 0.006, 0.007, 0.008$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.005$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	3.3306E-06	4.4508E-06	4.9904E-06	4.9980E-06	4.9499E-06
$MSE(\hat{p}_{B'1})$	3.3508E-06	4.4718E-06	5.0114E-06	5.0174E-06	4.9690E-06
$MSE(\hat{p}_{MLE})$	1.1841E-04	5.4817E-05	3.7781E-05	2.6902E-05	2.1330E-05
$p = 0.006$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	5.4620E-06	6.9512E-06	7.0983E-06	7.1791E-06	6.8084E-06
$MSE(\hat{p}_{B'1})$	5.4980E-06	6.9887E-06	7.1327E-06	7.2117E-06	6.8490E-06
$MSE(\hat{p}_{MLE})$	1.5569E-04	7.0589E-05	4.3417E-05	3.3162E-05	2.5273E-05
$p = 0.007$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	7.9479E-06	9.2850E-06	9.6347E-06	9.6161E-06	9.2208E-06
$MSE(\hat{p}_{B'1})$	8.0065E-06	9.3409E-06	9.6871E-06	9.6664E-06	9.2686E-06
$MSE(\hat{p}_{MLE})$	1.8707E-04	7.5878E-05	5.0300E-05	3.8194E-05	2.9938E-05
$p = 0.008$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.0768E-05	1.2447E-05	1.2669E-05	1.2093E-05	1.1611E-05
$MSE(\hat{p}_{B'1})$	1.0852E-05	1.2528E-05	1.2747E-05	1.2167E-05	1.1690E-05
$MSE(\hat{p}_{MLE})$	2.0946E-04	9.1635E-05	5.8055E-05	4.2742E-05	3.4843E-05

Table-12 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 10$  with different combinations of  $p = 0.005, 0.006, 0.007, 0.008$  and  $m = 5, 10, 15, 20, 25$ .

Table 13:  $k = 10$ ,  $p = 0.009, 0.01, 0.02, 0.03$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.009$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.4095E-05	1.5807E-05	1.5575E-05	1.5423E-05	1.4294E-05
$MSE(\hat{p}_{B'1})$	1.4221E-05	1.5920E-05	1.5681E-05	1.5524E-05	1.4404E-05
$MSE(\hat{p}_{MLE})$	3.3372E-04	1.0322E-04	6.5449E-05	5.1053E-05	3.9094E-05
$p = 0.01$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.8210E-05	1.9936E-05	1.9093E-05	1.7952E-05	1.6800E-05
$MSE(\hat{p}_{B'1})$	1.8379E-05	2.0097E-05	1.9231E-05	1.8085E-05	1.6930E-05
$MSE(\hat{p}_{MLE})$	2.6769E-04	1.1665E-04	7.4129E-05	5.3540E-05	4.3466E-05
$p = 0.02$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	7.9248E-05	7.2587E-05	6.4383E-05	5.5948E-05	4.9063E-05
$MSE(\hat{p}_{B'1})$	8.0445E-05	7.3586E-05	6.5370E-05	5.6769E-05	4.9932E-05
$MSE(\hat{p}_{MLE})$	7.7790E-04	2.4565E-04	1.6019E-04	1.1424E-04	8.9265E-05
$p = 0.03$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.6815E-04	1.4399E-04	1.2083E-04	1.0304E-04	8.8492E-05
$MSE(\hat{p}_{B'1})$	1.7165E-04	1.4721E-04	1.2349E-04	1.0547E-04	9.0659E-05
$MSE(\hat{p}_{MLE})$	2.7501E-03	4.0664E-04	2.4659E-04	1.8104E-04	1.4063E-04

Table-13 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 10$  with different combinations of  $p = 0.009, 0.01, 0.02, 0.03$  and  $m = 5, 10, 15, 20, 25$ .



Table 14:  $k = 10$ ,  $p = 0.04, 0.05, 0.06, 0.07$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.04$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	2.7947E-04	2.1812E-04	1.8658E-04	1.5466E-04	1.3503E-04
$MSE(\hat{p}_{B'1})$	2.8766E-04	2.2464E-04	1.9258E-04	1.5984E-04	1.3835E-04
$MSE(\hat{p}_{MLE})$	4.7970E-03	5.4403E-04	3.5728E-04	2.6227E-04	2.0367E-04
$p = 0.05$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	4.2111E-04	3.2064E-04	2.5194E-04	2.0914E-04	1.8079E-04
$MSE(\hat{p}_{B'1})$	4.3432E-04	3.3344E-04	2.6276E-04	2.1774E-04	1.8758E-04
$MSE(\hat{p}_{MLE})$	1.2563E-02	8.4855E-04	4.5156E-04	3.3252E-04	2.6620E-04
$p = 0.06$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	5.5055E-04	4.2260E-04	3.3451E-04	2.7243E-04	2.2721E-04
$MSE(\hat{p}_{B'1})$	5.7192E-04	4.4230E-04	3.5185E-04	2.8719E-04	2.4088E-04
$MSE(\hat{p}_{MLE})$	2.1022E-02	1.4193E-03	6.0974E-04	4.2062E-04	3.2314E-04
$p = 0.07$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	7.1297E-04	5.2898E-04	4.1495E-04	3.4058E-04	2.9036E-04
$MSE(\hat{p}_{B'1})$	7.5187E-04	5.5571E-04	4.3954E-04	3.6249E-04	3.1361E-04
$MSE(\hat{p}_{MLE})$	3.2272E-02	2.3502E-03	7.5571E-04	7.5571E-04	4.1002E-04

Table-14 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 10$  with different combinations of  $p = 0.04, 0.05, 0.06, 0.07$  and  $m = 5, 10, 15, 20, 25$ .

Table 15:  $k = 10$ ,  $p = 0.08, 0.09, 0.1$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.08$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	8.8583E-04	6.5048E-04	5.0621E-04	4.0891E-04	3.5726E-04
$MSE(\hat{p}_{B'1})$	9.3492E-04	6.8854E-04	5.4078E-04	4.4174E-04	3.8616E-04
$MSE(\hat{p}_{MLE})$	5.4293E-02	3.8652E-03	1.0961E-03	6.4116E-04	5.2369E-04
$p = 0.09$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.0383E-03	7.6432E-04	6.0528E-04	4.8103E-04	4.1071E-04
$MSE(\hat{p}_{B'1})$	1.1162E-03	8.2394E-04	6.5740E-04	5.2627E-04	4.5140E-04
$MSE(\hat{p}_{MLE})$	7.2972E-02	6.2664E-03	1.4612E-03	7.5971E-04	5.9529E-04
$p = 0.1$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.2246E-03	9.2362E-04	6.8347E-04	5.7343E-04	4.7303E-04
$MSE(\hat{p}_{B'1})$	1.3279E-03	9.9429E-04	7.4465E-04	6.3427E-04	5.2576E-04
$MSE(\hat{p}_{MLE})$	9.6444E-02	1.3330E-02	1.6181E-03	1.3371E-03	7.1557E-04

Table-15 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 10$  with different combinations of  $p = 0.08, 0.09, 0.1$  and  $m = 5, 10, 15, 20, 25$ .

Table 16:  $k = 15$ ,  $p = 0.005, 0.006, 0.007, 0.008$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.005$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	4.1599E-06	4.8808E-06	5.0095E-06	4.8111E-06	4.5628E-06
$MSE(\hat{p}_{B'1})$	4.1809E-06	4.9008E-06	5.0285E-06	4.8287E-06	4.5792E-06
$MSE(\hat{p}_{MLE})$	8.9782E-05	3.8597E-05	2.5179E-05	1.7880E-05	1.3810E-05
$p = 0.006$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	6.3672E-06	7.2169E-06	6.9268E-06	6.6213E-06	6.2627E-06
$MSE(\hat{p}_{B'1})$	6.4031E-06	7.2514E-06	6.9575E-06	6.6499E-06	6.2910E-06
$MSE(\hat{p}_{MLE})$	1.0678E-04	4.9047E-05	2.9716E-05	2.1206E-05	1.7196E-05
$p = 0.007$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	8.8637E-06	9.7900E-06	9.3750E-06	8.6453E-06	8.3146E-06
$MSE(\hat{p}_{B'1})$	8.9197E-06	9.8425E-06	9.4233E-06	8.6896E-06	8.3581E-06
$MSE(\hat{p}_{MLE})$	1.2354E-04	5.5867E-05	3.4883E-05	2.5128E-05	2.0663E-05
$p = 0.008$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.2017E-05	1.2810E-05	1.1758E-05	1.0961E-05	1.0185E-05
$MSE(\hat{p}_{B'1})$	1.2099E-05	1.2887E-05	1.1826E-05	1.1024E-05	1.0246E-05
$MSE(\hat{p}_{MLE})$	2.4358E-04	6.4797E-05	3.9662E-05	2.9437E-05	2.3370E-05

Table-16 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 15$  with different combinations of  $p = 0.005, 0.006, 0.007, 0.008$  and  $m = 5, 10, 15, 20, 25$ .

Table 17:  $k = 15$ ,  $p = 0.009, 0.01, 0.02, 0.03$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.009$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.5806E-05	1.5904E-05	1.4711E-05	1.3719E-05	1.2185E-05
$MSE(\hat{p}_{B'1})$	1.5921E-05	1.6011E-05	1.4807E-05	1.3810E-05	1.2276E-05
$MSE(\hat{p}_{MLE})$	2.7447E-04	7.2743E-05	4.6269E-05	3.4602E-05	2.6411E-05
$p = 0.01$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.9290E-05	1.9094E-05	1.7777E-05	1.5895E-05	1.4252E-05
$MSE(\hat{p}_{B'1})$	1.9441E-05	1.9229E-05	1.7902E-05	1.6012E-05	1.4381E-05
$MSE(\hat{p}_{MLE})$	1.9090E-04	7.9142E-05	5.1203E-05	3.7888E-05	2.9314E-05
$p = 0.02$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	7.7196E-05	6.3557E-05	5.3951E-05	4.5169E-05	4.0730E-05
$MSE(\hat{p}_{B'1})$	7.8202E-05	6.4408E-05	5.4716E-05	4.5838E-05	4.1305E-05
$MSE(\hat{p}_{MLE})$	2.2374E-03	1.7653E-04	1.1397E-04	8.2246E-05	6.5161E-05
$p = 0.03$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	1.5784E-04	1.2243E-04	9.6288E-05	8.0800E-05	7.0920E-05
$MSE(\hat{p}_{B'1})$	1.6083E-04	1.2488E-04	9.8370E-05	8.2735E-05	7.2528E-05
$MSE(\hat{p}_{MLE})$	7.8835E-03	2.9993E-04	1.8318E-04	1.3259E-04	1.0567E-04

Table-17 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 15$  with different combinations of  $p = 0.009, 0.01, 0.02, 0.03$  and  $m = 5, 10, 15, 20, 25$ .

Table 18:  $k = 15$ ,  $p = 0.04, 0.05, 0.06, 0.07$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.04$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	2.4810E-04	1.9384E-04	1.4871E-04	1.2263E-04	1.0360E-04
$MSE(\hat{p}_{B'1})$	2.5412E-04	1.9927E-04	1.5293E-04	1.2636E-04	1.0591E-04
$MSE(\hat{p}_{MLE})$	1.9655E-02	1.1963E-03	2.6913E-04	1.9428E-04	1.5192E-04
$p = 0.05$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	3.6837E-04	2.6225E-04	2.0940E-04	1.7137E-04	1.4769E-04
$MSE(\hat{p}_{B'1})$	3.7958E-04	2.7037E-04	2.1701E-04	1.7790E-04	1.5514E-04
$MSE(\hat{p}_{MLE})$	4.2438E-02	1.8604E-03	5.6624E-04	2.7102E-04	2.1758E-04
$p = 0.06$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	4.8791E-04	3.5900E-04	2.7970E-04	2.2146E-04	1.9689E-04
$MSE(\hat{p}_{B'1})$	5.0210E-04	3.7193E-04	2.9144E-04	2.3149E-04	2.0192E-04
$MSE(\hat{p}_{MLE})$	7.2819E-02	7.0501E-03	1.0527E-03	3.6908E-04	2.9080E-04
$p = 0.07$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	6.0603E-04	4.3622E-04	3.4544E-04	2.8133E-04	2.3842E-04
$MSE(\hat{p}_{B'1})$	6.2674E-04	4.5475E-04	3.6176E-04	2.9559E-04	2.5317E-04
$MSE(\hat{p}_{MLE})$	1.1716E-01	1.4181E-02	2.4547E-03	6.4547E-04	4.4632E-04

Table-18 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 15$  with different combinations of  $p = 0.04, 0.05, 0.06, 0.07$  and  $m = 5, 10, 15, 20, 25$ .

Table 19:  $k = 15$ ,  $p = 0.08, 0.09, 0.1$ ,  $\alpha = 1$ ,  $N = 10000$

$p = 0.08$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	7.3871E-04	5.4315E-04	4.1452E-04	3.5102E-04	2.9617E-04
$MSE(\hat{p}_{B'1})$	7.7028E-04	5.6469E-04	4.3557E-04	3.6997E-04	3.1453E-04
$MSE(\hat{p}_{MLE})$	1.5989E-01	2.7586E-02	5.0902E-03	1.6245E-03	1.0531E-03
$p = 0.09$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	8.6129E-04	6.5064E-04	5.1550E-04	4.3114E-04	3.6849E-04
$MSE(\hat{p}_{B'1})$	9.0280E-04	6.8725E-04	5.4165E-04	4.6025E-04	3.9929E-04
$MSE(\hat{p}_{MLE})$	2.0509E-01	5.3086E-02	1.3889E-02	4.9522E-03	1.1073E-03
$p = 0.1$					
$m$	5	10	15	20	25
$MSE(\hat{p}_{B1})$	9.9767E-04	7.7144E-04	6.1556E-04	5.1293E-04	4.3627E-04
$MSE(\hat{p}_{B'1})$	1.0679E-03	8.1737E-04	6.5614E-04	5.4559E-04	4.6183E-04
$MSE(\hat{p}_{MLE})$	2.5111E-01	7.7235E-02	2.6940E-02	8.4818E-03	3.4771E-03

Table-19 compares the  $MSE$  of Bayes estimator and approximation to the Bayes estimator with the  $MSE$  of the  $MLE$  by using  $k = 15$  with different combinations of  $p = 0.08, 0.09, 0.1$  and  $m = 5, 10, 15, 20, 25$ .

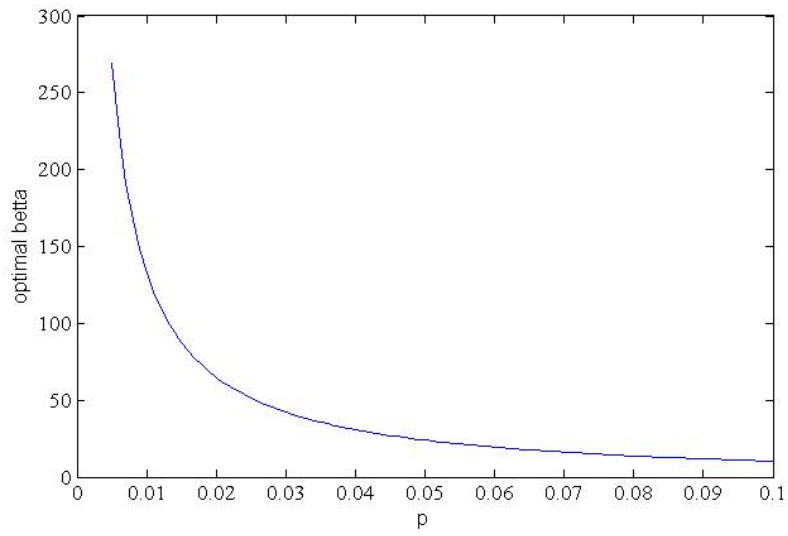


Figure 1:  $K = 5, m = 10$

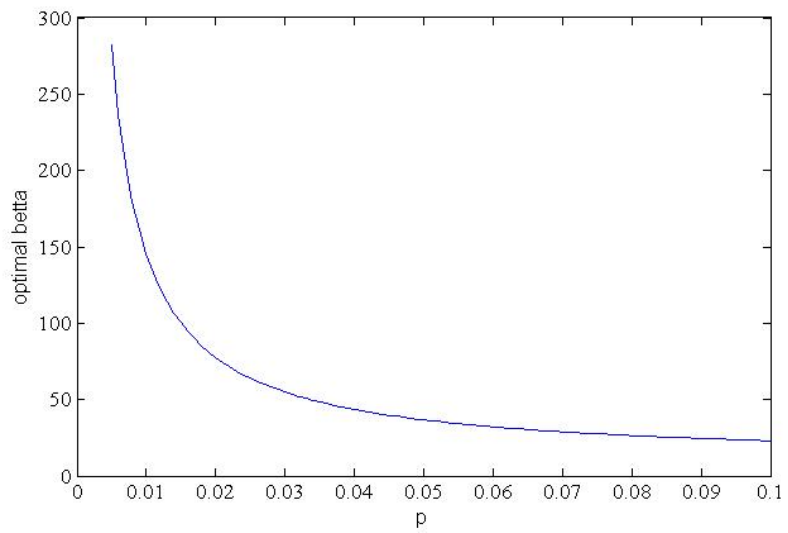


Figure 2:  $K = 5, m = 20$

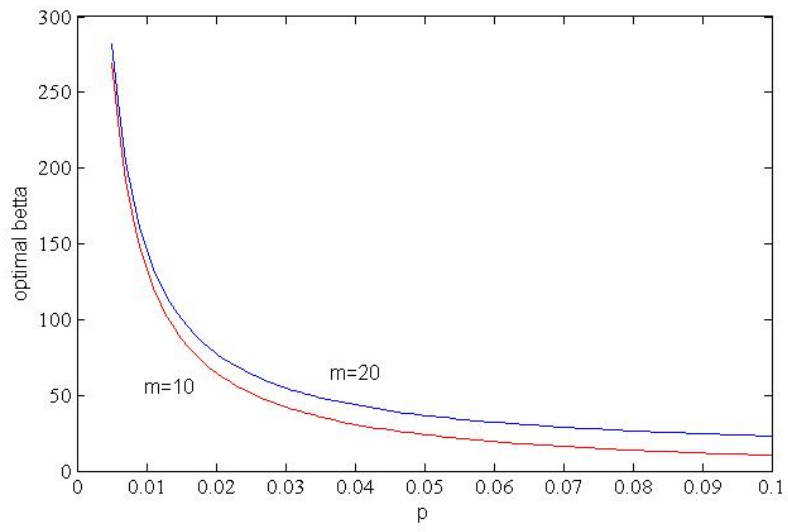


Figure 3:  $K = 5$ ,  $m = 10$  or  $20$

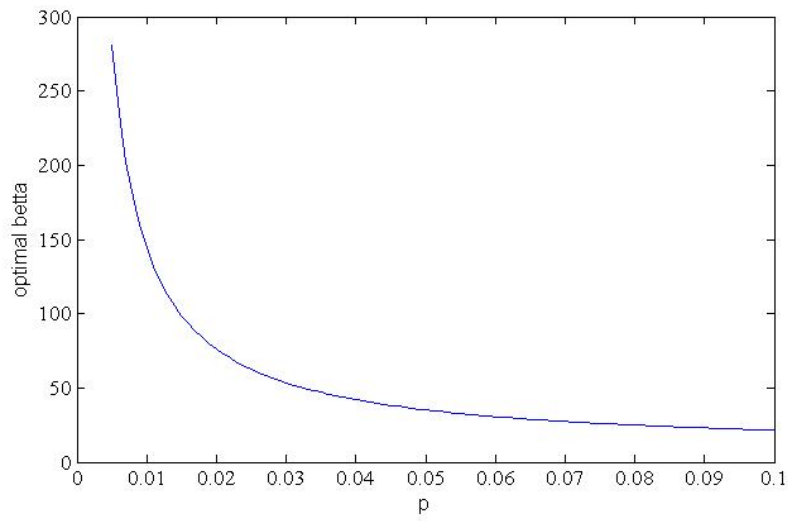


Figure 4:  $K = 10$ ,  $m = 10$



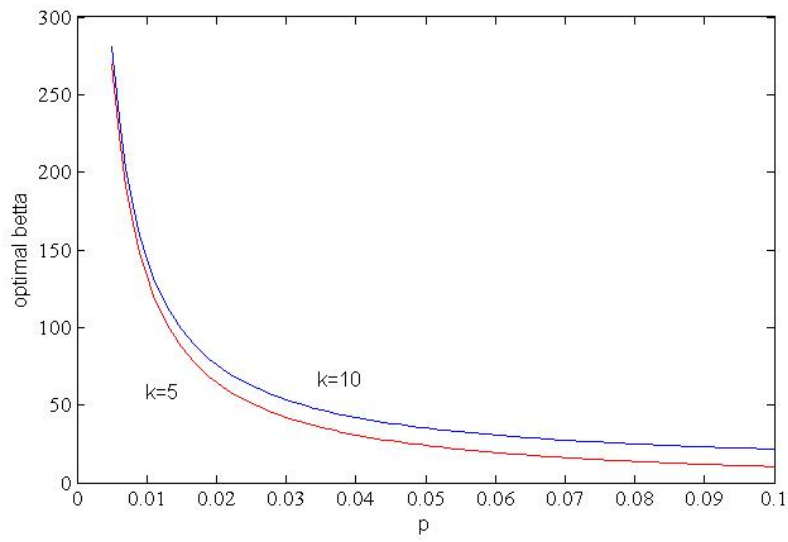


Figure 5:  $K = 5$  or  $10$ ,  $m = 10$

## 5 Conclusion

According to the analytical and numerical results obtained above, We know that the  $MSE$  of Bayes estimators and approximation to Bayes estimators are both smaller than that of  $MLE$  for small  $p$ . Indirect Bayes estimators performs very well too. We could draw a conclusion that Bayes estimators are much better than  $MLE$  when  $p$  is small.

## References

- [1] Bhattacharyya, F. K., Karandinos, M. G., Defoliart, G. R., 1979. Point estimates and confidence interval for infection rates using pooled organisms in epidemiologic studies. *American Journal of Epidemiology*. **109**, 124-131.
- [2] Chaubey, Y., Li, W., 1995. Comparison between maximum likelihood and Bayes methods for estimation of binomial probability with sample compositing. *Journal of Official Statistics*. **11**, 379-390.
- [3] Chiang, C. L., Reeves, W. C., 1962. Statistical estimation of virus infection rates in mosquito vector populations. *American Journal of Hygiene*. **75**, 377-391.
- [4] Chick, S., 1996. Bayesian models for limiting dilution assay and group test data. *Biometrics*. **52**, 1055-1062,
- [5] Dorfman, R., 1943. The detection of defective members of large populations. *Annals of Mathematical Statistics*. **14**, 436-440.
- [6] Gastwirth, J., Johnson, W., Reneau, D., 1991. Bayesian analysis of screening data: Application to AIDS in blood donors. *The Canadian Journal of Statistics*. **19**, 135-150.
- [7] Gastwirth, J., 2000. The efficiency of pooling in the detection of rare mutations. *American Journal of Human Genetics*. **67**, 1035-1039.
- [8] Gibbs, A. J., Gower, J. C., 1960. The use of a multiple-transfer method in plant virus transmission studies-some statistical points arising in the analysis of results. *Annals of Applied Biology*. **48**, 75-83.
- [9] Hepworth, G., 1996. Exact confidence intervals for proportions estimated by group testing. *Biometrics*. **52**, 1134-1146.

- [10] Katholi, C. R., Unnasch, T. R., 2006. Important experimental parameters for determining infection rates in arthropod vectors using pool screening approaches. *The American Society of Tropical Medicine and Hygiene*. **74**, 779-785.
- [11] Kline, R. L., Brothers, T. A., Brookmeyer, R., Zeger, S., Quinn, T. C., 1989. Evaluation of human immunodeficiency virus seroprevalence in population surveys using pooled sera. *Journal of Clinical Microbiology*. **27**, 1449-1452.
- [12] Radoni, B. C., Hepworth, G., Richardson, C., Moran, J. R., 1994. The use of a sequential batch testing procedure and ELISA to determine the incidence of five viruses in Victorian cut-flower sim Carnations. *Aust. J. Agric. Res.* **45**, 223-230.
- [13] Swallow, W. H., 1985. Group testing for estimating infection rates and probabilities of disease transmission. *Phytopathology*. **75**, 882-889.
- [14] Worlund, D. D., Taylor, G., 1983. Estimation of disease incidence in fish populations. *Canadian Journal of Fisheries and Aquatic Science*. **40**, 2194-2197.
- [15] Xie, M., Tatsuoka, K., Sacks, J., Young, S., 2001. Group testing with blockers and synergism. *J. Am. Statist. Assoc.* **96**, 92-102.

## APPENDIX

I list the equations in this study here for the convenience of finding them.

① Bayes Inferences from Prior on p.

Loss function  $L_1(p, a) = (p - a)^2$ .

$$\hat{p}_{B1} = \frac{\sum_{j=0}^y \binom{y}{j} (-1)^j B(\alpha + 1, km + kj - ky + \beta)}{\sum_{j=0}^y \binom{y}{j} (-1)^j B(\alpha, km + kj - ky + \beta)}. \quad y \geq 1 \quad \text{Equation 2.4 page 8}$$

Approximation to the Bayes estimator  $\hat{p}_{B1}$ .

$$\hat{p}_{B'1} = \frac{\alpha \sum_{j=0}^y (-1)^j \binom{y}{j} (km + kj - ky + \beta)^{-\alpha-1}}{\sum_{j=0}^y (-1)^j \binom{y}{j} (km + kj - ky + \beta)^{-\alpha}}. \quad y \geq 1 \quad \text{Equation 2.10 page 12}$$

Loss function  $L_2(p, a) = p^{-1}(p - a)^2$ .

$$\hat{p}_{B2} = \begin{cases} \frac{\sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, km + kj + \beta - ky)}{\sum_{i=0}^{k-1} \sum_{j=0}^{y-1} (-1)^j \binom{y-1}{j} B(km + kj + \beta + i - k, \alpha)} & \text{if } y \geq 1 \\ \frac{\sum_{j=0}^y (-1)^j \binom{y}{j} B(\alpha, km + \beta + kj - ky)}{B(\alpha - 1, km + \beta)} & \text{if } y = 0 \text{ and } \alpha > 1. \end{cases} \quad \text{Theorem 1 page 10}$$

Approximation to the Bayes estimator  $\hat{p}_{B2}$ .

$$\hat{p}_{B'2} = \begin{cases} \frac{\sum_{j=0}^y (-1)^j \binom{y}{j} (km + kj - ky + \beta)^{-\alpha}}{\sum_{i=0}^{k-1} \sum_{j=0}^{y-1} (-1)^j \binom{y-1}{j} (km + kj + i + \beta - ky)^{-\alpha}}, & \text{if } y \geq 1; \\ \frac{\alpha - 1}{n + \alpha + \beta - 1}, & \text{if } y = 0 \text{ and } \alpha > 1. \end{cases} \quad \text{Equation 2.13 page 13}$$

② Bayes Estimators of  $p$  from Prior on  $\pi$ .

Loss function  $L_1(p, a) = (p - a)^2$ .

$$\hat{p}_{IB1} = 1 - \frac{\Gamma(m + \alpha + \beta)\Gamma(m - y + \beta + 1/k)}{\Gamma(m - y + \beta)\Gamma(m + \alpha + \beta + 1/k)} \quad \text{Equation 3.1 page 18}$$

Approximation to the Bayes estimator  $\hat{p}_{IB1}$ .

$$\hat{p}_{IB'1} = 1 - \left( \frac{m + \beta - y}{m + \alpha + \beta} \right)^{1/k} \quad \text{Equation 3.2 page 18}$$

Loss function  $L_2(p, a) = p^{-1}(p - a)^2$ .

$$\hat{p}_{IB2} = (y + \alpha - 1) \sum_{i=0}^{k-1} \frac{\Gamma(m - y + \beta)\Gamma(m + \alpha + \beta + \frac{i}{k} - 1)}{\Gamma(m - y + \beta + \frac{i}{k})\Gamma(m + \alpha + \beta)} \quad \text{Equation 3.5 page 19}$$

Approximation to the Bayes estimator  $\hat{p}_{IB2}$ .

$$\hat{p}_{IB'2} = \frac{y + \alpha - 1}{m + \alpha + \beta} \frac{1}{\sum_{i=0}^{k-1} \left( \frac{m - y + \beta}{m + \alpha + \beta} \right)^{i/k}} \quad \text{Equation 3.6 page 20}$$