Journal of Official Statistics, Vol. 11, No. 4, 1995, pp. 379–390

Comparison Between Maximum Likelihood and Bayes Methods for Estimation of Binomial Probability with Sample Compositing

Yogendra P. Chaubey and Weiming Li[1]

This article focuses on the Bayesian approach to estimating a population prevalence rate through the method of sample compositing which has important applications in environmental sampling as well as in estimating the prevalence of certain diseases, etc. This method requires random samples of fixed size k, which is determined before the experimentation based on cost consideration as well as the target error in the form of the mean squared error of the estimator. Thus, two choices of prior may be available to the experimenter, (i) the prior on P, the population proportion, (ii) the prior on $P' = 1 - (1 - P)^k$, the group prevalence proportion. These two choices are considered in this article and their performance has been evaluated in comparison with the maximum likelihood estimator. It is observed that the Bayes methodology offers different choices to the experimenter with possible reduction in cost as well as error.

Key words: Batch sampling; composite sample; Bayes estimate; MLE; bias correction.

1. Introduction

Batch sampling, also known as composite sampling, involves taking samples in batches, and classifying them into one of two mutually exclusive categories, positive and negative or defective and nondefective, say. Each batch is supposed to consist of an equal number of sampling units and may be natural or artificially constructed. A composite sample of size k is formed by combining k individual test portions (units) and the composite is judged positive or negative based on a single measure. Alternatively, a sample of size n may be divided into m batches, each consisting of k units such that $n = mk$. Thus, in batch sampling, a batch is considered positive if at least one of its members has the positive characteristic. Traditionally, each unit of a sample is classified and therefore, batch sampling is more cost effective as compared to the traditional approach.

Sample compositing has been used in various situations (see Garner, Stapanian, Yfantis, and Williams (1989) for various applications), however, we concentrate on applications which involve estimation of proportions in a population, e.g., proportions of people with AIDS (or other disease), proportion of polluted samples,

[1] Department of Mathematics and Statistics, Concordia University, Montreal, Canada.
Acknowledgement: This research was partially supported by a research grant from NSERC of Canada. The authors would also like to thank Daniel Thorburn and a referee for some constructive suggestions which have improved the article substantially.

© Statistics Sweden

proportion of defectives produced in an industry. Dorfman (1943) applied this method for estimating the proportion of people with venereal disease in a large population and more recently Garner, Stapanian, and Williams (1987) advocated its use in environmental monitoring. The reader may further be referred to the papers by Boswell and Patil (1987) and Boswell, Gore, Patil, and Taillie (1992) for the use of the above technique in classification of polluted and nonpolluted samples.

In Garner et al. (1989) the maximum likelihood estimator (MLE) of the population prevalence is considered for the case of composite sampling and is compared with the estimator obtained in the case where each sample unit is classified. In the present article, we investigate a Bayesian perspective of the above situation.

Section 2 presents various estimators considered including the MLE whereas Section 3 gives the form of the asymptotic distribution of the estimators. The estimators are compared in Section 4. Section 4.1 presents the comparison based on the Bayes risk whereas Section 4.2 presents the comparison based on bias and mean squared error criteria. Some practical considerations are discussed in Section 5.

2. Methods of Estimation

2.1. *Maximum likelihood*

As in Garner et al. (1989) we summarize the sample compositing technique for estimating the probability of presence of a characteristic in a population as follows: Suppose random samples of size k called *groups* or *batches*, are chosen from the sample of n individual test portions ($n \gg k$), and each group of test portions is composited and analyzed. The group size is $m = n/k$. Let P denote the population prevalence. A group is declared to be positive (or defective) if it has at least one defective item. Let P' denote the probability of a group being classified as defective and $X_{(k)}$ be the number of groups which are defective. Then $X_{(k)} \sim Binom(m, P')$, i.e.,

$$P[X_{(k)} = x] = \binom{m}{x} P'^{x}(1 - P')^{m-x} \tag{2.1}$$

where $x = 0, 1, \ldots, m$; and P' is the probability that a group is defective, i.e.,

$$P' = 1 - (1 - P)^k$$

or

$$P = 1 - (1 - P')^{1/k}. \tag{2.2}$$

Table 2.1. "Optimal" k for given P and m

	m						
P	10	20	30	40	50	70	100
0.25	3	3	4	4	5	5	5
0.10	5	8	10	12	12	13	13
0.05	9	15	20	25	25	25	25
0.01	35	60	70	110	120	140	140

The MLE of P' as in Garner et al. (1989) is given by $\hat{P}'_k = X_{(k)}/m$ and thus the MLE of P is given by

$$\hat{P}_k = 1 - (1 - X_{(k)}/m)^{1/k}. \tag{2.3}$$

They suggested that k should be chosen so as to give the smallest MSE, where $MSE = V(\hat{P}_k) + B^2$, $B = E(\hat{P}_k) - P$ and gave a series of tables for choosing such an "optimal" k. This is summarized in Table 2.1.

The above procedure suffers from two important problems: (i) The optimality criterion does not give any importance to the bias and (ii) The optimal k depends on the unknown value P. The first problem requires some consideration of bias whereas the second problem warrants a Bayesian solution.

The bias of MLE is computed for a selection of values of P, k and m and is displayed in Table 2.2. From this table we note that MLE may have serious bias, especially when the value of k is far from its optimal value. We find also that the bias of MLE of P is always positive and increases as k increases, although MLE gives lower MSE than that of the traditional estimator (requiring the knowledge of presence or absence of the characteristic in each unit).

Hence, we consider the possibility of a bias correction. Even though the explicit form of bias as a function of P is not available, a reasonable approximation may be provided using the standard methods. We have, for large m,

$$E(\hat{P}_k) \approx P + \frac{k-1}{2mk^2}\left[\frac{1-(1-P)^k}{(1-P)^{k-1}}\right]. \tag{2.4}$$

The comparison of approximate bias using the formula (2.4) with the exact bias shows that the above approximation is adequate for our purposes (i.e., when P is small, see Table 2.2). Thus, we propose, a bias corrected estimator, $\hat{P}_{(k)C}$, given by

$$\hat{P}_{k(C)} = \hat{P}_k - \frac{k-1}{2mk^2}\frac{[X_{(k)}/m]}{[1 - X_{(k)}/m]^{(k-1)/k}} \tag{2.5}$$

when $X_{(k)} < m$, otherwise $\hat{P}_{k(C)} = 1$.

Table 2.3 displays the ratio of the bias of corrected MLE to that of uncorrected MLE from which we find that the bias correction is effective for small k as well as large k. Substantial reduction in bias is observed for small values of P.

In the next section we consider the Bayesian estimation procedure where $X_{(k)}$ is replaced by X for notational convenience.

2.2. *Bayes method*

To develop a Bayes estimator for P we have two choices, namely, by considering a *prior* on P or a *prior* on P'. The *prior* on P may be more meaningful since P' depends on k, whose optimal value is desired for producing the smallest MSE using sample compositing. However, when the value of k is fixed in advance such as in fixed size

Table 2.2. *Exact bias and approximate bias of MLE*

		m					
		10		20		30	
	k	exact	approx.	exact	approx.	exact	approx.
$P = 0.25$							
	2	0.0081	0.0073	0.0038	0.0036	0.0025	0.0024
	5	0.0541	0.0193	0.0133	0.0096	0.0073	0.0064
	10	0.3940	0.0566	0.2209	0.0283	0.1273	0.0189
	15	0.6417	0.1723	0.5556	0.0861	0.4828	0.0574
	20	0.7221	0.5599	0.6970	0.2800	0.6732	0.1866
$P = 0.10$							
	2	0.0028	0.0026	0.0014	0.0013	0.0009	0.0009
	5	0.0057	0.0050	0.0026	0.0025	0.0017	0.0017
	10	0.0190	0.0076	0.0044	0.0038	0.0027	0.0025
	15	0.0905	0.0108	0.0142	0.0054	0.0049	0.0036
	20	0.2391	0.0154	0.0695	0.0072	0.0227	0.0051
$P = 0.05$							
	2	0.0014	0.0013	0.0007	0.0006	0.0004	0.0004
	5	0.0024	0.0022	0.0012	0.0011	0.0008	0.0007
	10	0.0036	0.0029	0.0015	0.0014	0.0010	0.0010
	15	0.0057	0.0034	0.0019	0.0017	0.0012	0.0011
	20	0.0145	0.0040	0.0024	0.0020	0.0015	0.0013
$P = 0.01$							
	2	0.0003	0.0003	0.00013	0.00013	0.00009	0.00008
	5	0.0004	0.0004	0.00021	0.00020	0.00014	0.00014
	10	0.0005	0.0005	0.00024	0.00024	0.00016	0.00016
	15	0.0005	0.0005	0.00026	0.00025	0.00017	0.00017
	20	0.0006	0.0005	0.00027	0.00025	0.00018	0.00017
$P = 0.005$							
	2	0.00013	0.00013	0.00006	0.00006	0.00004	0.00004
	5	0.00022	0.00020	0.00010	0.00010	0.00007	0.00007
	10	0.00025	0.00023	0.00012	0.00012	0.00008	0.00008
	15	0.00026	0.00024	0.00013	0.00012	0.00008	0.00008
	20	0.00027	0.00025	0.00013	0.00012	0.00009	0.00008

lots and the experimenter has some knowledge of P', the alternative prior may be more meaningful. For completeness, both cases are considered in this article.

2.2.1. Prior on P

Consider a $Beta(\alpha, \beta)$ prior on P which can accommodate a variety of *prior* information about P to develop a Bayes estimator, i.e., we consider a probability distribution on P given by

$$f(p) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)}. \tag{2.6}$$

The joint distribution of (X, P) is given by

$$f(x,p) = \binom{m}{x} \{B(\alpha,\beta)\}^{-1} p^{\alpha-1}(1-p)^{km-kx+\beta-1}[1-(1-p)^k]^x.$$

Table 2.3. Ratio of bias of adjusted MLE to the exact bias

	k	m = 10	20	30	50	100
$P = 0.25$						
	1	1	1	1	1	1
	2	0.0098	0.0024	0.0016	0.0010	0.0005
	5	0.6048	0.1241	−0.0012	−0.0092	−0.0042
	10	0.9688	0.9264	0.8714	0.7130	0.1809
	15	0.9954	0.9904	0.9851	0.9735	0.9371
	20	0.9992	0.9983	0.9974	0.9957	0.9912
$P = 0.10$						
	1	1	1	1	1	1
	2	0.0009	0.0006	0.0004	0.0003	0.0001
	5	−0.0191	−0.0135	−0.0086	−0.0050	−0.0025
	10	0.4870	−0.0021	−0.0182	−0.0104	−0.0049
	15	0.8729	0.5116	0.1177	−0.0165	−0.0086
	20	0.9590	0.8727	0.7094	0.2193	−0.0147
$P = 0.05$						
	1	1	1	1	1	1
	2	8.4E-6	0.0002	0.0002	0.0001	0.0001
	5	−0.0252	−0.0116	−0.0076	−0.0045	−0.0022
	10	−0.0200	−0.0190	−0.0122	−0.0071	−0.0035
	15	0.2334	−0.0245	−0.0162	−0.0093	−0.0045
	20	0.6437	0.0128	−0.0207	−0.0120	−0.0057
$P = 0.01$						
	1	1	1	1	1	1
	2	−0.0006	−0.0001	−0.00003	1.2E-7	6.44E-6
	5	−0.0225	−0.0107	−0.0070	−0.0041	−0.0021
	10	−0.0310	−0.0147	−0.0096	−0.0057	−0.0028
	15	−0.0349	−0.0164	−0.0107	−0.0063	−0.0031
	20	−0.0338	−0.0176	−0.0115	−0.0068	−0.0033
$P = 0.005$						
	1	1	1	1	1	1
	2	−0.0006	−0.0001	−0.0001	−0.0000	1.5E-9
	5	−0.0222	−0.0105	−0.0069	−0.0041	−0.0020
	10	−0.0300	−0.0143	−0.0094	−0.0056	−0.0027
	15	−0.0331	−0.0157	−0.0103	−0.0061	−0.0030
	20	−0.0351	−0.0166	−0.0108	−0.0064	−0.0032

The marginal probability density $f(x)$ of X is given by

$$f(x) = \int_0^1 \binom{m}{x} \{B(\alpha, \beta)\}^{-1} p^{\alpha-1}(1-p)^{km-kx+\beta-1}[1-(1-p)^k]^x dp$$

$$= \binom{m}{x} \{B(\alpha, \beta)\}^{-1} \sum_{j=0}^{x} \binom{x}{j}(-1)^j B(\alpha, kj + km - kx + \beta).$$

The posterior distribution of P can therefore be evaluated as

$$f(p|x) = f(x,p)/f(x) = p^{\alpha-1}(1-p)^{km-kx+\beta-1} \times [1-(1-p)^k]^x \left[\sum_{j=0}^{x} \binom{x}{j}(-1)^j B(\alpha, kj+km-kx+\beta)\right]^{-1}.$$

Then the Bayes estimator of P is given by

$$\hat{P}_{k(B)} = E(P|X) = \frac{\sum_{j=0}^{X} \binom{X}{j} (-1)^j B(\alpha+1, kj+km-kX+\beta)}{\sum_{j=0}^{X} \binom{X}{j} (-1)^j B(\alpha, kj+km-kX+\beta).} \tag{2.7}$$

2.2.2. Prior on P'

If we put a *Beta*(α, β) *prior* on P', the *posterior* distribution of P' can be similarly computed as above and is given by

$$f(p'|x) = \frac{p'^{x+\alpha-1}(1-p')^{m-x+\beta-1}}{B(x+\alpha, m-x+\beta)}.$$

The above yields the *posterior* of P through the technique of transformation and we can compute the mean of this distribution giving us the Bayes estimator under the square error loss. However, we can directly compute this, without computing the *posterior* of P as

$$\begin{aligned}\hat{P}_{k(B)'} &= \int_0^1 [1-(1-p')^{1/k}] f(p'|X)dp' \\ &= 1 - \frac{\Gamma(m+\alpha+\beta)\Gamma(m-X+\beta+1/k)}{\Gamma(m-X+\beta)\Gamma(m+\alpha+\beta+1/k)}.\end{aligned} \tag{2.8}$$

An ad hoc but simple estimation procedure would be to consider the posterior mean of P' and use it to consider an estimator of P by the virtue of the relation between P and P'. This procedure will be called *indirect Bayesian* procedure which is presented in the next section.

2.2.3 Ad hoc method (approximation to Bayes method)

Choosing a *Beta*(α, β) *prior* on P' gives the Bayes estimator of P' as

$$\hat{P}' = \frac{X+\alpha}{m+\alpha+\beta}. \tag{2.9}$$

Substituting above into (2.2) in place of P', we have *indirect* Bayesian estimator

$$\hat{P}_{k(IB)} = 1 - \left(1 - \frac{X+\alpha}{m+\alpha+\beta}\right)^{1/k}. \tag{2.10}$$

The estimator given in (2.10) can be given the following justification. For large N

we have from Abramowitz and Stegun (1960, eqn 6.1.46)

$$\frac{\Gamma(N+a)}{\Gamma(N+b)} \simeq N^{a-b}. \tag{2.11}$$

Therefore for large m, from (2.8) we find that

$$\hat{P}_{k(B)'} \simeq 1 - \left(\frac{m-X+\beta}{m+\alpha+\beta}\right)^{1/k} = \hat{P}_{k(IB)} \tag{2.12}$$

which is easier to compute, and therefore may be preferred in practice.

3. Asymptotic Distribution

In practice the values of n and m are fairly large, so that large sample distributions of the estimators can be used for inference purpose. The distributions of Bayesian estimators are not easily tractable, however, the asymptotic distributions of the maximum likelihood estimator and that of the indirect Bayesian estimator follow from the general result which obeys the Mann-Wald Theorem (see Rao 1973, p. 426), for the function g given by

$$g(P') = 1 - \left(1 - \frac{P'+b}{1+c}\right)^a \tag{3.1}$$

then

$$\sqrt{m}\frac{g(\hat{P}'_k) - g(P')}{P'(1-P')} \sim N\left(0, \left[\frac{a}{1+c}\left(1-\frac{P'+b}{1+c}\right)^{a-1}\right]^2\right). \tag{3.2}$$

For the maximum likelihood estimator, we have $g(\hat{P}'_k) = \hat{P}_k$ with $b = c = 0$ and $a = 1/k$. For the indirect Bayes estimator, $g(\hat{P}'_k) = \hat{P}_{k(IB)}$ with $b = \alpha/m$, $c = (\alpha+\beta)/m$ and $a = 1/k$. These results may be used in setting up confidence limits for P or for setting up large sample standard errors for the corresponding estimators.

The large sample confidence intervals can be obtained using the above large sample distributions using MLE or indirect Bayes. For Bayes estimator based on the prior on P one may approximate the prior on P' by a beta density and use the large sample result (3.2). It can be pointed out that all the distributions discussed in the above section have the same asymptotic distribution. However, the use of the asymptotic distribution is not investigated in this article.

In the next section we present comparisons of the estimators presented above.

4. Comparisons of Estimators

4.1. Bayesian comparisons of different estimators

In general, the Bayes estimator $\delta_B(X)$ is the "optimal" estimator of the parameter θ in the sense that under the given prior and loss function it has the smallest risk, where the risk $r(\tau, \delta)$ for a decision rule (estimator) $\delta(X)$ with respect to the prior τ is given by

$$r(\tau, \delta) = E_\theta E_{X|\theta}[\delta(X) - \theta]^2 \tag{4.1}$$

under the squared error loss. Therefore, the task of a Bayesian is done once the Bayes estimator is computed. However, he/she may still be interested in assessing the closeness of other estimators to the Bayes estimator in terms of the Bayes risk. A frequentist, on the other hand, is generally interested in comparing the estimators based on bias and *MSE*. For this purpose, first we compare the Bayes risks of $\hat{P}_{k(B)}$ and $\hat{P}_{k(B')}$ with that of the MLE. We define the Bayes Relative Efficiency (BRE) of an estimator δ by

$$BRE(\delta) = \frac{r(\tau, \delta_B)}{r(\tau, \delta)}. \tag{4.2}$$

The computations in the following tables have been done using the following decomposition of the risk function

$$r(\tau, \delta) = E_X(\delta(X) - \delta_B(X))^2 + E_X Var(\theta|X). \tag{4.3}$$

It is clear from the above formula that the Bayes estimator $\delta_B(X) = E(\theta|X)$ is the optimal under the squared error loss.

For comparing the Bayes estimator corresponding to a *Beta*(α, β) *prior* on P, the *prior* on P' is chosen to be *Beta*(α', β'), where α' and β' are obtained by equating the first two moments. Thus we have

$$\alpha' = \frac{(1-A)(B-A)}{A^2 - B} \quad and \quad \beta' = \frac{A(B-A)}{A^2 - B} \tag{4.4}$$

where

$$A = \Gamma(\alpha+\beta)\Gamma(\beta+k)/\Gamma(\beta)\Gamma(\alpha+\beta+k) \quad \text{and}$$
$$B = \Gamma(\alpha+\beta)\Gamma(\beta+2k)/\Gamma(\beta)\Gamma(\alpha+\beta+2k).$$

For the reason of simplification, we take $\alpha = 1$ in the following computations. It is interesting to find that $\alpha' = 1$ and $\beta' = \beta/k$ in this case.

Table 4.1 displays the values of BRE for various estimators corresponding to the case $m = 10$ for several values of $\bar{P}$ and k, where $\bar{P}$ represents a prior guess for P and is taken to be the mean of the prior on P. Similar computations were performed for other values of $m = 5(5)100$, however, only the $m = 10$ case is reported here for the reason of space.

We conclude the following from Table 4.1:

i. The MLE and the bias corrected MLE are much inferior to the Bayes estimators.
ii. The Bayes estimator with matching prior on P' with that on P gives almost matching risk.
iii. The three alternative estimators have 100% BRE for $k = 1$ because these coincide with the Bayes estimator.
iv. The ad hoc estimator (indirect Bayes) seems to be a good approximation to $\hat{P}_{k(B)'}$, especially when P is small.

In the following section we consider the frequentist's concerns regarding the resulting estimators.

Table 4.1. Bayesian comparison ($m = 10$)

	k	(α, β)	BRE($\hat{P}_k$)	BRE($\hat{P}_{k(C)}$)	(α', β')	BRE($\hat{P}_{k(IB)}$)	BRE($\hat{P}_{k(B)'}$)
$\bar{P} = 0.25$		(1, 3)					
	1		0.7143	0.7143	(1, 3.00)	1.0000	1.0000
	5		0.1416	0.1409	(1, 0.60)	0.8682	1.0000
	10		0.0787	0.0785	(1, 0.30)	0.6966	1.0000
	15		0.0693	0.0693	(1, 0.20)	0.6174	1.0000
	20		0.0663	0.0663	(1, 0.15)	0.5730	1.0000
$\bar{P} = 0.10$		(1, 9)					
	1		0.5000	0.5000	(1, 9.00)	1.0000	1.0000
	5		0.1737	0.1793	(1, 1.80)	0.9768	1.0000
	10		0.0320	0.0320	(1, 0.90)	0.8991	1.0000
	15		0.0180	0.0180	(1, 0.60)	0.8225	1.0000
	20		0.0142	0.0142	(1, 0.45)	0.7669	1.0000
$\bar{P} = 0.05$		(1, 19)					
	1		0.3333	0.3333	(1, 19.0)	1.000	1.0000
	5		0.3764	0.4182	(1, 3.80)	0.9937	1.0000
	10		0.0466	0.0471	(1, 1.90)	0.9713	1.0000
	15		0.0144	0.0144	(1, 1.27)	0.9351	1.0000
	20		0.0078	0.0078	(1, 0.95)	0.8945	1.0000
$\bar{P} = 0.01$		(1, 99)					
	1		0.0909	0.0909	(1, 99.0)	1.000	1.0000
	5		0.2938	0.3237	(1, 19.8)	0.9996	1.0000
	10		0.3987	0.4535	(1, 9.90)	0.9984	1.0000
	15		0.2787	0.3073	(1, 6.60)	0.9966	1.0000
	20		0.0975	0.1012	(1, 4.95)	0.9941	1.0000
$\bar{P} = 0.005$		(1, 199)					
	1		0.0476	0.0476	(1, 199.0)	1.000	1.0000
	5		0.1798	0.1968	(1, 39.8)	0.9999	1.0000
	10		0.2883	0.3225	(1, 19.9)	0.9996	1.0000
	15		0.3541	0.4017	(1, 13.3)	0.9990	1.0000
	20		0.3516	0.3979	(1, 9.95)	0.9983	1.0000

4.2. *Classical comparison of different estimators*

In this section we compare various estimators under the criterion of relative efficiency (RE) and relative bias (RB) which are defined as follows. The relative bias of an estimator $\hat{P}$ is defined to be

$$RB(\hat{P}) = \frac{E(\hat{P} - P)}{P} \tag{4.5}$$

whereas the relative efficiency of $\hat{P}$ is given by

$$RE(\hat{P}) = \frac{MSE(\hat{P}_k)}{MSE(\hat{P})}. \tag{4.6}$$

These may be termed as frequentist's comparisons. In this comparison the Bayes method is used only to get the form of the estimator and the loss function does not play any further role. As mentioned before, we want to assess the performance of different estimators by the frequentist's measures.

Table 4.2. Relative bias (RB), $m = 10$

	k	$\hat{P}_k$	$\hat{P}_{k(C)}$	(α, β)	$\hat{P}_{k(B)}$	(α', β')	$\hat{P}_{k(IB)}$	$\hat{P}_{k(B)'}$
$P = 0.25$				(1, 3)				
	1	0	0		0	(1, 3.00)	0	0
	5	0.2166	0.1310		0.0955	(1, 0.60)	0.0125	0.0955
	10	1.5758	1.5268		0.2929	(1, 0.30)	0.0019	0.2929
	15	2.5669	2.5552		0.3739	(1, 0.20)	−0.1085	0.3739
	20	2.8885	2.8861		0.3501	(1, 0.15)	−0.2356	0.3501
$P = 0.10$				(1, 9)				
	1	0	0		0	(1, 9.00)	0	0
	5	0.0573	−0.0011		0.0384	(1, 1.80)	−0.0031	0.0384
	10	0.1905	0.0928		0.0803	(1, 0.90)	0.0049	0.0803
	15	0.9047	0.7897		0.1457	(1, 0.60)	0.0145	0.1457
	20	2.3911	2.2930		0.2348	(1, 0.45)	0.0171	0.2348
$P = 0.05$				(1, 19)				
	1	0	0		0	(1, 19.00)	0	0
	5	0.0485	−0.0012		0.0236	(1, 3.80)	−0.0074	0.0236
	10	0.0671	−0.0013		0.0435	(1, 1.90)	−0.0043	0.0435
	15	0.1144	0.0267		0.0624	(1, 1.27)	−0.0005	0.0624
	20	0.2996	0.1929		0.0849	(1, 0.95)	−0.0039	0.0849
$P = 0.01$				(1, 99)				
	1	0	0		0	(1, 99.00)	0	0
	5	0.0436	−0.0010		0.0045	(1, 19.8)	−0.0087	0.0045
	10	0.0508	−0.0016		0.0118	(1, 9.90)	−0.0110	0.0118
	15	0.0545	−0.0019		0.0182	(1, 6.60)	−0.0110	0.0182
	20	0.0573	−0.0022		0.0237	(1, 4.95)	−0.0103	0.0237
$P = 0.005$				(1, 199)				
	1	0	0		0	(1, 199.0)	0	0
	5	0.0431	−0.0010		0.0016	(1, 39.8)	−0.0063	0.0016
	10	0.0494	−0.0015		0.0051	(1, 19.9)	−0.0098	0.0051
	15	0.0521	−0.0017		0.0089	(1, 13.3)	−0.0112	0.0089
	20	0.0539	−0.0019		0.0125	(1, 9.95)	−0.0117	0.0125

Table 4.2 displays the values of RB while Table 4.3 displays those of RE for the case $m = 10$ for different choices of k and P. As in Section 4.1, these values were computed for $m = 5(5)100$, but we have chosen to present only the case $m = 10$.

From these tables we conclude the following:

i. The biases of the Bayes estimators, $\hat{P}_{k(B)}$, $\hat{P}_{k(IB)}$ and $\hat{P}_{k(B)'}$, are smaller than that of MLE as well as its corrected version.
ii. Indirect Bayes estimator performs very well in the sense of having small bias.
iii. The biases of both the Bayes estimators with matching priors almost coincide.
iv. The MSE's of Bayes estimators, $\hat{P}_{k(B)}$, $\hat{P}_{k(IB)}$ and $\hat{P}_{k(B)'}$, are always smaller than that of the MLE as well as its corrected version.
v. The indirect Bayes estimator performs very well in the sense of having small MSE; it gives an MSE even smaller than that of the bias corrected MLE.

Table 4.3. Relative efficiency (RE)

	k	$\hat{P}_{k(C)}$	(α, β)	$\hat{P}_{k(B)}$	(α', β')	$\hat{P}_{k(IB)}$	$\hat{P}_{k(B)'}$
$P = 0.25$			(1, 3)				
	1	1.0000		1.9600	(1, 3.00)	1.9600	1.9600
	5	1.0326		4.6967	(1, 0.60)	7.0430	4.6967
	10	0.9948		19.0334	(1, 0.30)	80.6230	19.0334
	15	0.9986		35.9098	(1, 0.20)	269.945	35.9098
	20	0.9997		59.8529	(1, 0.15)	147.473	59.8529
$P = 0.10$			(1, 9)				
	1	1.0000		4.0000	(1, 9.00)	4.0000	4.0000
	5	1.1792		1.7066	(1, 1.80)	1.9141	1.7066
	10	1.0464		7.8234	(1, 0.90)	10.6963	7.8234
	15	1.0028		35.6663	(1, 0.60)	69.5410	35.6663
	20	0.9992		71.1192	(1, 0.45)	210.2756	71.1192
$P = 0.05$			(1, 19)				
	1	1.0000		9.0000	(1, 19.0)	9.0000	9.0000
	5	1.1249		2.1946	(1, 3.80)	2.3577	2.1946
	10	1.1985		1.8981	(1, 1.90)	2.1710	1.8981
	15	1.0756		5.6804	(1, 1.27)	7.0462	5.6804
	20	1.0158		24.8676	(1, 0.95)	34.7447	24.8676
$P = 0.01$			(1, 99)				
	1	1.0000		121.0000	(1, 99,0)	121.0000	121.0000
	5	1.0948		9.8546	(1, 19.8)	10.1257	9.8546
	10	1.1174		4.4701	(1, 9.90)	4.6918	4.4701
	15	1.1331		3.1422	(1, 6.60)	3.3501	3.1422
	20	1.1478		2.5710	(1, 4.95)	2.7776	2.5710
$P = 0.005$			(1, 199)				
	1	1.0000		441.0000	(1, 199.0)	441.0000	441.0000
	5	1.0919		27.2737	(1, 39.8)	27.7123	27.2737
	10	1.1091		9.9768	(1, 19.9)	10.2871	9.9768
	15	1.1185		6.0878	(1, 13.3)	6.3488	6.0878
	20	1.1259		4.5019	(1, 9.95)	4.7390	4.5019

5. Discussion

Group testing is in general economical in the light of reduced average number of units to be tested, since a group is declared defective as soon as one item is found defective. Based on a given total cost per unit to be tested, a k may be determined in advance and fixed. Next, based on the prior knowledge of the population proportion (the mean or the median value of the *prior*) the above k may be compared to the optimal value for the MLE or that for the *Bayes estimator*. If it is far from the "optimum", a recommendation to revise the budget is suggested.

Choice of priors may be guided by a target value or a preliminary estimate of P, say P_0 along with a measure of its precision V_0. These may be used in determining the constants of a *Beta prior* by equating the first two moments. These values may be updated based on the new estimates.

The direct Bayes method offers a good alternative to the experimenter as he/she may have some general ideas about the population proportion which can be used to obtain the value of optimal k which can be updated in subsequent testing and estimation. Once k is known, a *prior* on P may be transformed into a *prior* on P' and then the alternative Bayes estimator or its approximation (the indirect Bayes estimator) may be used. The indirect Bayes estimator is simple to calculate and hence, may be attractive to users.

Choice of k may be guided by physical considerations or be chosen by min MSE criterion or be determined from the optimal k for MLE. For practical use a fixed k is desirable unless the variability of the prior is very high.

6. References

Abramowitz, M. and Stegun, I.A. (1960). Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Washington, D.C.: Bureau of National Standards.

Bowsell, M.T. and Patil, G.P. (1987). A Perspective on Composite Sampling. Communications in Statistics-Theory and Methods, 16, 3069–3093.

Boswell, M.T., Gore, S.D., Patil, G.P., and Taillie, C. (1992). A Bayesian Approach to Classifying Samples as Polluted or not Polluted. Technical Report No. 92-0801, State College, PA: Department of Statistics, The Pennsylvania State University.

Dorfman, R. (1943). The Detection of Defective Members of Large Populations. Annals of Mathematical Statistics, 14, 436–440.

Garner, F.C., Stapanian, M.A., and Williams, L.R. (1987). Composite Sampling for Environmental Monitoring. In Principles of Environmental Sampling, L.H. Keith (Ed.), Washington, D.C.: American Chemical Society.

Garner, F.C., Stapanian, M.A., Yfantis, E.A., and Williams, L.R. (1989). Probability Estimation with Sample Compositing Techniques. Journal of Official Statistics, 5, 365–374.

Rao, C.R. (1973). Linear Statistical Inference and Its Applications, 2nd edition. New York: John Wiley and Sons.

Received October 1992
Revised November 1995