BAYESIAN ESTIMATION OF MEASLES VACCINATION COVERAGE UNDER RANKED SET SAMPLING

Radhakanta Das[1], Vivek Verma[2], Dilip C. Nath[3]

ABSTRACT

The present article is concerned with the problem of estimating an unknown population proportion p, say, of a certain population characteristic in a dichotomous population using the data collected through ranked set sampling (RSS) strategy. Here, it is assumed that the proportion p is not fixed but a random quantity. A Bayes estimator of p is proposed under squared error loss function assuming that the prior density of p belongs to the family of Beta distributions. The performance of the proposed RSS-based Bayes estimator is compared with that of the corresponding classical version estimator based on maximum likelihood principle. The proposed procedure is used to estimate measles vaccination coverage probability among the children of age group 12-23 months in India using the real-life epidemiological data from National Family Health Survey-III.

Key words: Bayes estimator, maximum likelihood principle, square error loss, risk function and immunization coverage.

1. Introduction

Ranked Set Sampling (RSS) was first introduced by McIntyre (1952). This is an alternative method of sampling procedure that is used to achieve the greater efficiency in estimating the population characteristics. Generally the most appropriate situation for employing RSS is where the exact measurement of sampling units is expansive in time or effort; but the sample units can be readily ranked either through subjective judgement or *via* the use of relevant concomitant variables. The most basic version of RSS is balanced RSS where the same number of observations is drawn corresponding to each judgement order statistic. In order to draw a balanced ranked set sample of size n, first an integer s is chosen such that $n = ms$, for some positive

---

[1]Department of Statistics, Presidency University, 86/1 College Street Kolkata, 700073, West Bengal, India. E-mail: radhakanta.stat@presiuniv.ac.in

[2]Corresponding Author. Department of Statistics, Gauhati University, Guwahati, 781014, Assam, India. E-mail: viv_verma456@yahoo.com

[3]Department of Statistics, Gauhati University, Guwahati, 781014, Assam, India. E-mail: dilipc.nath@gmail.com

integer m. Then we select s^2 units from the population at random and the units are divided into sets of s units each. Within each set, s units are ranked according to the characteristic of interest by judgement or with the help of one or more auxiliary variables. From the ith set $(i = 1, 2, \cdots, s)$ we observe the actual measurement corresponding to only the ith ordered unit in that set. This entire procedure, which may be called a *cycle*, is repeated m times independently to obtain a ranked set sample of size $n = ms$.

Let $X_{[i]j}$ denote the quantified i^{th} judgement order statistic from the jth cycle. Thus, the sampling scheme yields the following ranked set sample of size n.

$$
\begin{matrix}
X_{[1]1}, \cdots, X_{[1]j}, \cdots, X_{[1]m} \\
\vdots \qquad \vdots \qquad \vdots \\
X_{[i]1}, \cdots, X_{[i]j}, \cdots, X_{[i]m} \\
\vdots \qquad \vdots \qquad \vdots \\
X_{[s]1}, \cdots, X_{[s]j}, \cdots, X_{[s]m}
\end{matrix}
\qquad (1.1)
$$

It is obvious that the observations within each row of above observation matrix are independently and identically distributed (iid), and the observations within any column are independently but not identically distributed. To acquire depth in theories and logistics of RSS methodology one can go through the book by Chen et al.(2004).

In the present investigation we assume that the variable of interest is binary; that is, there are only two possible outcomes, generally called success (denoted as 1) and failure (denoted as 0). Thus, the study variable is supposed to follow Bernoulli distribution with success probability p $(0 < p < 1)$, say. Here, the ranking of s binary observations in each set, where there are only 0 and 1 runs in the series, is done systematically as discussed by Terpstra and Nelson (2005). For instance, suppose $s = 4$ and the observations are, say, $X_1 = 1; X_2 = 0; X_3 = 1$, and $X_4 = 0$. Then, a possible ordered arrangement of the observations might be (X_2, X_4, X_1, X_3), or (X_4, X_2, X_1, X_3) or (X_2, X_4, X_3, X_1) or (X_4, X_2, X_3, X_1). But for the sake of uniqueness we take the arrangement (X_2, X_4, X_1, X_3) where the suffix of X in each run is in increasing order and hence we get the ordered statistics as $X_{(1)} = X_2; X_{(2)} = X_4; X_{(3)} = X_1$ and $X_{(4)} = X_3$. The same systematic rule can easily be extended in the ranking of a polytomous variable also. For a binary population the success probability p can be viewed as a proportion of individuals possessing certain known characteristic in the population. In classical inference on a population proportion, the ranked set sampling with binary data has already been introduced and used by many researchers like, among others, Lacayo et al. (2002), Kvam (2003), Terpstra (2004), Terpstra

and Liudahl (2004), Chen et al. (2005, 2006, 2007, 2009), Terpstra and Nelson (2005), Terpstra and Miller (2006), Chen (2008), Gemayel etal.,(2012) used RSS for auditing purpose, Wolfe (2010, 2012) and Zamanzade and Mahdizadeh (2017, 2017) discussed application of RSS to air quality monitoring. Jozani and Mirkamali (2010, 2011) used ranked set sample for binary data in the context of control charts for attributes. In earlier works, estimation of p using ranked set samples is based on the assumption that the parameter p is an unknown but a fixed quantity. But there may be situations where some prior knowledge on p may be available in terms of its changing pattern over time or with respect to other factors, which amounts to treat p as a random quantity. In this article a Bayesian estimation of p in the domain of ranked set sample is considered.

We organize the paper in the following way. In section 2, a Bayes estimator of the population proportion is proposed. As a natural competitor of the proposed estimator, a classical version estimator based on maximum likelihood principle is discussed in section 3. Section 4 provides an efficiency comparison of the estimators in terms of risk under square error loss function. In section 5 the proposed procedure is used to estimate measles vaccination coverage probability among the children of age group 12-23 months in India using the real-life epidemiological data from National Family Health Survey-III. Lastly, section 6 gives a brief concluding remark.

2. Bayes Estimators of p

Let X be the variable of interest assumed to follow Bernoulli (p) distribution with p being the success probability. It has been found in the literature (e.g. Stokes (1977)) that the use of a single concomitant variable for ranking is effective regardless of whether the association of the concomitant variable of interest is positive or negative. Suppose, after applying judgement ranking made on the basis of a readily available auxiliary variable, say Y, we have a ranked set sample $\{X_{[i]j}, i = 1(1)s, j = 1(1)m\}$ of size $n = ms$, where $X_{[i]j}$ denotes the quantified i^{th} judgement order statistics in the jth cycle. It can easily be justified that, for each $i = 1(1)s$, the observations in the ith ranking group $X_{[i]1}, \cdots, X_{[i]j}, \cdots, X_{[i]m}$ constitute a simple random sample (SRS) of size m from Bernoulli distribution with success probability denoted by $p_{[i]}$, say. So, for each $i = 1(1)s$, $p_{[i]}$ represents the probability of assuming the value 1 (which corresponds to success) for the ith judgement order statistic $X_{[i]1}$. Immediately we get the following result.

Result 2.1: *Suppose an observation with a higher judgement order is more likely to*

be a 'success'. Then, we have

$$p_{[i]} = I_p(s-i+1,i), \quad for \ each \ i = 1,2,\cdots,s \tag{2.1}$$

and

$$\frac{1}{s}\sum_{i=1}^{s} p_{[i]} = p, \tag{2.2}$$

where $I_x(a,b)$, $x \in (0,1)$, is the standard incomplete beta integral given by

$$I_x(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1}(1-t)^{b-1}dt.$$

The above result is standard (see Tepstra (2004)) and hence omitted.

Note 2.1: If an observation with a lower judgement order is more likely to be a success, then, for every $i = 1,2,\cdots,s$, that

$$p_{[i]} = I_p(i,s-i+1) \tag{2.3}$$

and (2.2) also holds in this case.

In this section the proportion parameter p is assumed to be a random variable and the randomness is quantified in terms of suitable prior density, say $\tau(p)$ of p over the interval $[0,1]$. Here, we derive a Bayesian estimator of p by incorporating the available prior information on p along with the information provided by the ranked set sample data. By the virtue of ranked set sampling all the observations $X_{[i]j}, \forall i = 1(1)s, \forall j = 1(1)m$ are independent. Let us define the variables

$$Z_i = \sum_{j=1}^{m} X_{[i]j}, \ \forall i = 1(1)s.$$

Obviously, the variables $Z_1, Z_2, \cdots, Z_s$ are independently distributed as $Z_i \sim$ Binomial $(m, p_{[i]})$. For each i, $1 \leq i \leq s$, $p_{[i]}$ is a function of the basic parameter p, so we denote it as $p_{[i]}(p)$. With this notation one can easily write the likelihood function of p, given the ranked set sample data $\mathbf{z} = (z_1, z_2, \cdots, z_s)$ as

$$
\begin{aligned}
L(p|\mathbf{z}) &= \prod_{i=1}^{s} P[Z_i = z_i | p_{[i]}(p)] \\
&= \prod_{i=1}^{s} \binom{m}{z_i} [p_{[i]}(p)]^{z_i} [1-p_{[i]}(p)]^{(m-z_i)}
\end{aligned}
\tag{2.4}
$$

where **z** is a particular realization of the random vector $Z = (Z_1, Z_2, \cdots, Z_s)$. Thus, the posterior density of p, given $\mathbf{Z} = \mathbf{z}$, with respect to the prior $\tau(p)$ for p is given by

$$
\begin{aligned}
h(p|\mathbf{z}) &= \frac{L(p|\mathbf{z})\tau(p)}{\int_0^1 L(p|\mathbf{z})\tau(p)dp} \\
\Leftrightarrow h(p|\mathbf{z}) &\propto L(p|\mathbf{z})\tau(p) \\
\Leftrightarrow h(p|\mathbf{z}) &\propto \prod_{i=1}^{s}[p_{[i]}(p)]^{z_i}\,[1 - p_{[i]}(p)]^{(m-z_i)}\tau(p).
\end{aligned} \tag{2.5}
$$

The information regarding unknown parameter is upgraded in the light of the observed data and is quantified through the posterior distribution $h(p|\mathbf{z})$ w.r.t. the prior $\tau(p)$ and hence any statistical inference regarding p is made on the basis of its posterior distribution given the ranked set sample data. Here, the posterior distribution does not have any standard form. In such a situation, to make any statistical inference on p one should use Monte Carlo simulation technique which provides a great deal of computational facilities. According to this method, a sufficiently large number, say N of observations are drawn at random independently from the posterior distribution $h(p|\mathbf{z})$ and let it be denoted as $p^{(1)}, p^{(2)}, \cdots, p^{(N)}$. Then the posterior mean and variance of p can be approximated as

$$
\begin{aligned}
E(p|\mathbf{z}) &= \int_0^1 ph(p|\mathbf{z})dp \\
&\simeq \frac{1}{N}\sum_{j=1}^{N}p^{(j)}
\end{aligned} \tag{2.6}
$$

and

$$
\begin{aligned}
V(p|\mathbf{z}) &= \int_0^1 \{p - E(p|\mathbf{z})\}^2 h(p|\mathbf{z})dp \\
&\simeq \frac{1}{N}\sum_{j=1}^{N}[p^{(j)}]^2 - \left[\frac{1}{N}\sum_{j=1}^{N}p^{(j)}\right]^2
\end{aligned} \tag{2.7}
$$

Thus, under square error loss function the Bayes estimate $(\hat{p}_B)$ of p w.r.t. the prior $\tau(p)$ is given by the mean of the posterior distribution $h(p|\mathbf{z})$, that is, for sufficiently

large N,

$$\hat{p}_B \simeq \frac{1}{N} \sum_{j=1}^{N} p^{(j)}. \qquad (2.8)$$

Alternative Bayes Esimator of p:

An alternative Bayes estimator of p can easily be constructed as the average of the Bayes estimators of $p_{[i]}$'s by assuming that the probabilities $p_{[1]}, p_{[2]}, \cdots, p_{[s]}$ are all unknown parameters, although all of them are the functions of the basic parameter p, satisfying the relation (2.2). Suppose $p_{[1]}, \cdots, p_{[s]}$ are independently and identically distributed with common prior density given below.

$$\tau(\theta) = \frac{1}{\mathscr{B}(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} , \; 0 < \theta < 1, \; \alpha > 0, \; \beta > 0. \qquad (2.9)$$

Then, under squared error loss function the Bayes estimator of $p_{[i]}$, based on Z_i, can easily be obtained as (see Ferguson (2014))

$$\hat{p}_B^*[i] = \frac{Z_i + \alpha}{m + \alpha + \beta}, \quad \text{for } i = 1, ..s. \qquad (2.10)$$

After having the estimators $\hat{p}_B^*[i]; i = 1, 2, \cdots, s$, we are in a position to construct, by virtue of the relation (2.2), a Bayesian estimator of p as

$$\hat{p}_B^* = \frac{1}{s} \sum_{i=1}^{s} \hat{p}_B^*[i] \qquad (2.11)$$

which, by using (2.10), takes the form

$$\hat{p}_B^* = \frac{m\bar{X} + \alpha}{m + \alpha + \beta}, \qquad (2.12)$$

where $\bar{X} \left(= \frac{1}{ms} \sum\limits_{i=1}^{s} \sum\limits_{j=1}^{m} X_{[i]j} \right)$ is the grand mean of ms sample observations. Note that, under square error loss, the risk function of $\hat{p}_B^*$ is given as

$$
\begin{aligned}
R_{\hat{p}_B^*}(p) &= E(\hat{p}_B^* - p)^2 \\
&= E\left(\frac{m\bar{X} + \alpha}{m + \alpha + \beta} - p \right)^2 \\
&= (m + \alpha + \beta)^{-2} E\{(m\bar{X} - mp) + \alpha - p(\alpha + \beta)\}^2 \\
&= (m + \alpha + \beta)^{-2} \left[\frac{1}{s^2} \sum\limits_{i=1}^{s} mp_{[i]}(1 - p_{[i]}) + \{\alpha - p(\alpha + \beta)\}^2 \right] \\
&= \frac{m}{(m + \alpha + \beta)^2 s^2} \sum\limits_{i=1}^{s} p_{[i]}(1 - p_{[i]}) + \frac{\{\alpha - p(\alpha + \beta)\}^2}{(m + \alpha + \beta)^2}. \quad (2.13)
\end{aligned}
$$

3. Estimator of p based on Maximum Likelihood Principle

Here, we briefly describe a classical version estimator of p based on the maximum likelihood (ML) principle. For this we first assume that the proportion parameter p is an unknown fixed number lying between 0 and 1. Now, all the observations in the ranked set sample are independently distributed, the likelihood function of p based on the given ranked set sample $\mathbf{X} = \mathbf{x}$ can be expressed as

$$
\begin{aligned}
L_1(p|\mathbf{x}) &= \prod\limits_{i=1}^{s} \prod\limits_{j=1}^{m} [p_{[i]}(p)]^{x_{[i]j}} [1 - p_{[i]}(p)]^{1 - x_{[i]j}} \\
&= \prod\limits_{i=1}^{s} [p_{[i]}(p)]^{z_i} [1 - p_{[i]}(p)]^{m - z_i} \\
&= \prod\limits_{i=1}^{s} \{I_p(s - i + 1, i)\}^{z_i} \{I_{1-p}(i, s - i + 1)\}^{m - z_i}. \quad (3.1)
\end{aligned}
$$

Equivalently, the log-likelihood function of p is given as

$$
l(p|\mathbf{x}) = \log_e L_1(p|\mathbf{x}) = \sum\limits_{i=1}^{s} z_i \log_e I_p(s - i + 1, i) + \sum\limits_{i=1}^{s} (m - z_i) \log_e I_{1-p}(i, s - i + 1),
$$

and hence the likelihood equation for determining MLE of p is obtained as

$$\frac{d}{dp}l(p|\mathbf{x}) = 0$$

$$\Leftrightarrow \sum_{i=1}^{s}\frac{z_ib(p|s-i+1,i)}{I_p(s-i+1,i)} = \sum_{i=1}^{s}\frac{(m-z_i)b(1-p|i,s-i+1)}{I_{1-p}(i,s-i+1)} \qquad (3.2)$$

where $b(x|\alpha,\beta)$ represents the probability density function of $Beta(\alpha,\beta)$ distribution. Due to the complicated nature of the above likelihood equation it is difficult to get an explicit solution for p and hence the RSS-based MLE of p does not have any closed form.

As an alternative way out we can obtain an estimate of p by indirectly using the maximum likelihood principle. For this we first consider the probabilities $p_{[1]}, p_{[2]}, \cdots, p_{[s]}$ as the unknown parameters, although all of them are the functions of the basic parameter p. Then we determine the maximum likelihood estimates of those parameters separately and substitute these estimates in the relation (2.2) of Result 2.1 to get an estimate of p. Given the ranked set sample $\mathbf{X} = \mathbf{x}$, the likelihood function of the parameters $p_{[1]}, p_{[2]}, \cdots, p_{[s]}$ is written as

$$L_1(p_{[1]}, p_{[2]}, \cdots, p_{[s]}|\mathbf{x}) = \prod_{i=1}^{s}\prod_{j=1}^{m}[p_{[i]}]^{x_{[i]j}}[1-p_{[i]}]^{1-x_{[i]j}}$$

$$= \prod_{i=1}^{s}[p_{[i]}]^{z_i}[1-p_{[i]}]^{m-z_i} \qquad (3.3)$$

and the corresponding log-likelihood function is given by

$$l(p_{[1]}, p_{[2]}, \cdots, p_{[s]}|\mathbf{x}) = \sum_{i=1}^{s}z_i\log_e p_{[i]} + \sum_{i=1}^{s}(m-z_i)\log_e(1-p_{[i]}).$$

Thus, by solving s maximum likelihood equations, $\frac{\partial}{\partial p_{[i]}}l(p_{[1]}, p_{[2]}, \cdots, p_{[s]}|\mathbf{x}) = 0$, for $i = 1, 2, \cdots, s$, we easily get the MLE of $p_{[i]}$ as

$$\hat{p}_{[i]} = \frac{Z_i}{m}, \quad i = 1, 2, \cdots, s.$$

Then, after replacing $p_{[i]}$'s by $\hat{p}_{[i]}$'s in the relation (2.2) we get an estimate of p as

$$\hat{p}_M = \frac{1}{s}\sum_{i=1}^{s}\hat{p}_{[i]}. \qquad (3.4)$$

Note 3.1: It is easy to argue that the ML estimates $\hat{p}_{[1]}, \hat{p}_{[2]}, \cdots, \hat{p}_{[s]}$ are statistically independent as the variables Z_i's are independently distributed. Again, substituting the value $\frac{Z_i}{m}$ of $\hat{p}_{[i]}$ in the equation (3.4), the estimate $\hat{p}_M$ can be shown to be identical with the overall mean of the given ranked set sample. It is also readily verified that $\hat{p}_M$ is an unbiased estimator of p.

4. Comparison Between $\hat{p}_B$, $\hat{p}_B^*$ and $\hat{p}_M$

The goal of this section is to compare the estimators of p derived in sections 2 and 3. Since the posterior mean, by definition, minimizes the Bayes risk under squared error loss function, it is not surprising that a Bayes estimator of an unknown parameter is often superior to the corresponding MLE in respect of mean squared error (MSE). However, MLE neither requires any specification of prior distribution for the parameter nor it involves any particular loss function. Thus, the comparison should be made on the basis of a criterion which does not bother about the particular nature of prior information regarding unknown parameter. However, as MSE of an estimator can be regarded as risk under squared error loss, one can use risk function for comparison purpose. The expressions for risk functions of the estimators are described below. Under square error loss the risk of $\hat{p}_B$ is given as

$$R_{\hat{p}_B}(p) = E(\hat{p}_B - p)^2, \tag{4.1}$$

which cannot be further simplified analytically. On the other hand, the risk of $\hat{p}_M$ has a theoretical expression obtained as

$$
\begin{aligned}
R_{\hat{p}_M}(p) &= E(\hat{p}_M - p)^2 \\
&= V\left(\frac{1}{s}\sum_{i=1}^{s} \hat{p}_{[i]}\right) \\
&= \frac{1}{ms^2}\sum_{i=1}^{s} p_{[i]}(1 - p_{[i]}), \tag{4.2}
\end{aligned}
$$

after using the fact that $\hat{p}_{[i]}$'s are independent with $V(\hat{p}_{[i]}) = \frac{1}{m}p_{[i]}(1 - p_{[i]})$.

Result 4.1: The risk function $R_{\hat{p}_M}(p)$ of the estimator $\hat{p}_M$ is symmetric around '$p = \frac{1}{2}$'.

Proof: By (2.1) we rewrite $R_{\hat{p}_M}(p)$ as

$$
\begin{aligned}
R_{\hat{p}_M}(p) &= \frac{1}{ms^2} \sum_{i=1}^{s} I_p(s-i+1,i)[1-I_p(s-i+1,i)] \\
&= \frac{1}{ms^2} \sum_{i=1}^{s} I_p(s-i+1,i) I_{1-p}(i,s-i+1), \ \ \text{for } p \in [0,1].
\end{aligned}
$$

Now, for any $\xi \in [0,1]$, we see that

$$
\begin{aligned}
ms^2 R_{\hat{p}_M}\left(\frac{1}{2}+\xi\right) &= \sum_{i=1}^{s} I_{\frac{1}{2}+\xi}(s-i+1,i)\, I_{\frac{1}{2}-\xi}(i,s-i+1) \\
&= \sum_{j=1}^{s} I_{\frac{1}{2}+\xi}(j,s-j+1)\, I_{\frac{1}{2}-\xi}(s-j+1,j), \ \ \text{putting } j=s-i+1 \\
&= \sum_{j=1}^{s} I_{\frac{1}{2}-\xi}(s-j+1,j)\, I_{\frac{1}{2}+\xi}(j,s-j+1) \\
&= ms^2 R_{\hat{p}_M}\left(\frac{1}{2}-\xi\right),
\end{aligned}
$$

and hence the required proof follows.

In the present situation we compare the performances of the estimators $\hat{p}_B$, $\hat{p}_B^*$ and $\hat{p}_M$ by plotting their risk functions in the same co-ordinate axes. The estimator $\hat{p}_B$ performs uniformly better than the estimator $\hat{p}_M$ if

$$
R_{\hat{p}_B}(p) \leq R_{\hat{p}_M}(p),
$$

for all $p \in [0,1]$ with strict inequality for at least one value of p. Here, we conveniently choose $\text{Beta}(\alpha,\beta)$ distribution as a prior for p, that is,

$$
\tau(p) = \{\mathscr{B}(\alpha,\beta)\}^{-1} p^{\alpha-1}(1-p)^{\beta-1}, \ \ 0 < p < 1, \alpha > 0, \beta > 0.
$$

In the numerical computation we take, in particular, $(s,m) = (3,50), (5,30), (3,100), (5,60)$ and $(\alpha,\beta) = (\frac{1}{2},\frac{1}{2}), (2,2), (\frac{1}{2},3), (3,\frac{1}{2})$. Here, we compute the risk values for the Bayes estimator $\hat{p}_B$ by simulation technique with the help of Metropolis-Hasting's algorithm (given in Appendix) and then plot them over the whole range of p. The plotted risk functions are shown in Figures 1-4 given in Appendix. These figures show that the risk curves corresponding to Bayes estimators $\hat{p}_B$ and $\hat{p}_B^*$ completely lie below the risk curve of $\hat{p}_M$ implying that the proposed Bayes estimators are uniformly better than the estimator based on ML principle so far as the given parametric combinations are concerned. Again it is observed that the risk curves

corresponding to Bayes estimators $\hat{p}_B$ and $\hat{p}_B^*$ are not significantly distinct and hence we can conclude that these two estimators are more or less equally good.

5. Estimation of Measles Vaccination Coverage Probability

In public health related studies, the virus of measles is regarded as highly epidemic and is responsible for severe diseases. According to the Medical Dictionary, the virus of measles infects the lungs at childhood which may cause pneumonia and in older children it can lead to inflammation of the brain, called encephalitis, which can cause seizures and brain damage (Perry and Halsey, 2004). As precautionary measures the proper vaccination is introduced from the very beginning of the childhood to acquire the immunity against measles viruses. According to the Integrated Child Development Services (ICDS) program in India, a child should have received basic vaccinations (BCG, polio, DPT and measles) in the 12-23 months of their age.

Here, our objective is to illustrate the proposed procedures for estimating the vaccination coverage of the measles among the children of age group 12-23 months (the age by which children should have received all basic vaccinations) in India 2005-06. The study data has been taken from the website of the Measure DHS-Demographic and Health Surveys (DHS) (http://www.measuredhs.com). DHS provides national and state estimates of fertility, child mortality, the practice of family planning, attention to mother and child and access to services for mothers and children. For this study, data set of National Family Health Survey-III (NFHS-III,2005-2006) for the year 2005-06 of India is considered. Here, the samples of DHS are treated as our population of interest and those children who are in the 12-23 months of their age considered as our study population.

The event of receiving vaccination for a child usually depends on awareness of the child's mother regarding vaccination. The higher the educational qualification of a mother during child bearing period, the higher would be the awareness as expected. Therefore, mother's educational qualification is used as auxiliary variable for ranking purpose in ranked set sampling. The observations are obtained through the following steps.

1. A simple random sample of s^2 units is drawn from the target population and is randomly partitioned into s sets, each having s units.

2. In each of s sets the units are ranked according to the mother's qualification {1 = "No education", 2 = "Primary", 3 = "Secondary", 4 = "Higher"}. The

ranking process could also be based on the individuals' duration (in terms of years) of study. Here, samples in different sets are ranked based on mother's qualification denoted as (1, 2, 3, 4) and duration (in terms of years) of study. Obviously, there is a high chance of having ties. Then in that situation the observations are ordered systematically in the sequence, as discussed by Terpstra and Nelson (2005).

3. From the first set, the unit corresponding to the mother with lowest qualification (or duration of study) is selected. From the second set, the unit corresponding to the mother with the second lowest qualification is selected and so on. Finally, from the sth set, the unit corresponding to the mother with the highest qualification is selected. The remaining $s(s-1)$ sampled units are discarded from the data set.

4. The Steps 1 - 3, called a cycle, are repeated m times to obtain a ranked set sample of size $n = ms$.

Here, in particular, we take $(s, m) = (4, 100)$. Corresponding to each selected mother, information regarding whether her child is administrated with measles vaccination or not is collected. Suppose X is the binary response that takes value '1' if the child is vaccinated and '0' otherwise. With this notation we have the sample $\{X_{[1]1}, X_{[1]2}, \cdots, X_{[4]100}\}$ of size 400, where $X_{[r]j}$ takes the values '1' or '0' accordingly as the jth child in the rth ranking class is vaccinated or not. Obviously, $p_{[r]}$ is the proportion, in r^{th} class, of children who received the vaccination and p is the overall proportion of children receiving the vaccine in entire target population. The implementation of the proposed Bayes approach requires assuming the prior distributions of $p_{[r]}$'s. Here we use Beta (α, β) priors with $(\alpha, \beta) = (0.5, 0.5), (1, 1), (2, 2), (5, 5), (1, 2), (2, 1)$ and $(5, 3)$. With these parametric combinations we compute the estimates $\hat{p}_B$, $\hat{p}_B^*$ and $\hat{p}_M$. We also calculate the estimated relative risk of Bayes estimators w. r. t. $\hat{p}_M$ defined by

$$\hat{\rho}_{\hat{p}_B} = \frac{\hat{R}(\hat{p}_M)}{\hat{R}(\hat{p}_B)}$$

$$\hat{\rho}_{\hat{p}_B^*} = \frac{\hat{R}(\hat{p}_M)}{\hat{R}(\hat{p}_B^*)}$$

and all computed results are summarized in Table 5.1. From the table, it is observed that the Bayes estimate of the proportion of children receiving measles vaccine is very close to that based on ML approach. Also, both the estimates are very close to the value 58.8%, which is the estimated value of p reported by NFHS-III(2005-

Table 5.1: Estimates of the proportion and Relative Efficiency

Estimate	Bayesian approach							ML
	Symmetric prior				Asymmetric prior			
	$\alpha = 0.5$ $\beta = 0.5$	$\alpha = 1$ $\beta = 1$	$\alpha = 2$ $\beta = 2$	$\alpha = 5$ $\beta = 5$	$\alpha = 1$ $\beta = 2$	$\alpha = 2$ $\beta = 1$	$\alpha = 5$ $\beta = 3$	
$\hat{p}_B$	0.562	0.562	0.562	0.562	0.562	0.562	0.564	0.578
$\hat{p}_B^*$	0.577	0.576	0.575	0.570	0.570	0.580	0.581	0.578
$\hat{\rho}_{\hat{p}_B}$	2.13	2.02	2.14	2.25	2.37	2.02	2.15	
$\hat{\rho}_{\hat{p}_B^*}$	1.02	1.03	1.05	1.10	1.03	1.04	1.09	

06). It is also clear that the proposed Bayes procedure, especially the estimator $\hat{p}_B$, shows greater efficiency than the corresponding ML based procedure.

6. Concluding Remarks

The present work is concerned with the problem of estimating unknown population proportion p based on ranked set sample (RSS) drawn from a binary population. Since the RSS-based likelihood function of p is complicated, the direct application of Bayes principle (in the context of Bayesian paradigm) or maximum likelihood principle (in connection with classical framework) is not straightforward for estimating p. In Bayesian framework the RSS-based Bayes estimator does not have a simple explicit form even if we choose the simplest distribution, i.e. Uniform$(0,1)$ ($\equiv$ Beta$(1,1)$) as a possible prior for p. The RSS-based likelihood function of p can easily be expressed in the form of polynomial in p. Thus, under the assumption of Beta(α, β) prior for p, the posterior distribution can be shown, through a routine calculation, to be a mixture of several Beta distributions. Also, the explicit form of the Bayes estimator is not so convenient from the computational point of view. Obviously, the posterior distribution does not belong to the Beta-family and hence Beta(α, β) prior is not a conjugate prior in this case. In fact, there does not exist any conjugate prior in standard form due to complexity in the functional form of RSS-based likelihood of p.

As a natural competitor of the Bayes estimator of p, we have used here a very common estimator indirectly based on maximum likelihood principle used in finding MLEs of intermediate parameters $p_{[1]}, p_{[2]}, \cdots, p_{[s]}$. On the other hand, one can directly use the MLE of p as considered by Tepstra (2004) for comparison purpose. However, this estimator does not exist in a closed form but can be computed through

numerical methods. Since our main focus lies in the Bayesian approach of estimation by incorporating available prior information regarding the parameter of interest. We here consider a commonly used estimator for comparison purpose only.

In Bayesian statistics the selection of the prior distribution is crucial to the analysis of data because the final conclusion depends on this particular choice. In our proposed procedure we have considered the Beta prior due to its important features, *viz.,* proper interpretability according to the model (see Paolino (2001)), less computational complexity of posterior distribution (see Gupta and Nadarajah (2004)), having reasonable reflection of prior uncertainty (see Ferrari and Cribari-Neto (2004)), capability to extend to higher dimensions (see Pham-Gia (1994)), etc. However, the Beta-Binomial conjugate analysis may not be adequately robust. Thus, the precision of the prior is important and the sensitivity analysis regarding the prior is necessary. Keeping these in mind one can carry out a robust Bayesian analysis using non-conjugate priors. One such way out might be the use of Cauchy priors after expressing the likelihood of binary data in terms of its exponentially family form, and this Cauchy-Binomial model for binary data might be more robust (see Fuquene et al. (2008)). Several other robust approaches are also discussed in, among others, Berger et al. (1994) and Wang and Blei (2015). The consideration of robust Bayesian approach is beyond the scope of the present work and will be considered in a separate issue.

Acknowledgements

REFERENCES

BERGER, J. O., MORENO, E., PERICCHI, L. R., BAYARRI, M. J., BERNARDO, J. M., CANO, J. A., ... and DASGUPTA, A., (1994). An overview of robust Bayesian analysis. *Test*, 3 (1), pp. 5–124.

CHEN, H., (2008). Alternative Ranked set Sample Estimators for the Variance of a Sample Proportion ,*Applied Statistics Research Progress; Nova Publishers*, pp. 35–38.

CHEN, Z., BAI, Z. D., SINHA, B. K., (2004). Ranked Set Sampling: Theory and Applications, Lecture Notes in Statistics, 176, Springer-Verlag, New York.

CHEN, H., STASNY, E. A., WOLFE, D. A., (2005). Ranked Set Sampling for Efficient Estimation of a Population Proportion *Statistics in Medicine*, 24, pp. 3319–3329.

CHEN, H., STASNY, E. A., WOLFE, D. A., (2006). Unbalanced Ranked Set Sampling for Estimating a Population Proportion, *Biometrics*, 62, pp. 150–158.

CHEN, H., STASNY, E. A., WOLFE, D. A., (2007). Improved Procedures for Estimation of Disease Prevalence Using Ranked Set Sampling, *Biometrical Journal*, 49 (4), pp. 530–538.

CHEN H. , STASNY A. E., WOLFE A. D., MACEACHERN N. S., (2009). Unbalanced Ranked Set Sampling for Estimating A Population Proportion Under Imperfect Rankings, *Communications in Statistics - Theory and Methods*, 38 (12), pp. 2116–2125.

FERRARI, S., CRIBARI-NETO, F., (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31 (7), pp. 799–815.

FERGUSON, T. S., (2014). Mathematical statistics: A decision theoretic approach (Vol. 1) *Academic press.*

FUQUENE P.J.A., COOK, J.D., PERICCHI, L. R., (2008). A Case for Robust Bayesian priors with Applications to Binary Clinical Trials, *http://biostats.bepress.com/mdandersonbiostat/paper44*

GEMAYEL, N. M., STASNY, E. A., TACKETT, J. A., WOLFE, D. A., (2012). Ranked set sampling: an auditing application, *Review of Quantitative Finance and Accounting*, 39 (4), pp. 413–422.

GUPTA, A. K., NADARAJAH, S. (Eds.), (2004). Handbook of beta distribution and its applications. *CRC press.*

JOZANI, M.J., MIRKAMALI, S. J., (2010). Improved attribute acceptance sampling plans based on maxima nomination sampling, *Journal of Statistical Planning and Inference*, 140 (9), pp. 2448–2460.

JOZANI, M. J., MIRKAMALI, S. J., (2011). Control charts for attributes with maxima nominated samples, *Journal of Statistical Planning and Inference*, 141, pp. 2386–2398.

KVAM, H. P., (2003). Ranked Set Sampling Based on Binary Water Quality Data With Covariates, *Journal of Agricltural, Biological and Environmental Statistics*, 8 (3), pp. 271–279.

LACAYO, H., NEERCHAL, N. K., SINHA, B. K., (2002). Ranked Set Sampling from a Dichotomous Population, *Journal of Applied Statistical Science*, 11(1), pp. 83–90.

MCINTYRE, G. A., (1952). A method for unbiased selective sampling, using ranked sets, *Australian Journal of Agricultural Research*, 3, pp. 385–390.

NFHS-III., (2005-2006). National Family Health Survey", *International Institute for Population Sciences*, Bombay.

PAOLINO, P., (2001). Maximum likelihood estimation of models with beta-distributed dependent variables, *Political Analysis*, pp. 325–346.

PERRY, T. R., HALSEY, A. N., (2004). The Clinical Significance of Measles: A Review" *The Journal of Infectious Diseases,* 189, Suppl 1: S4–16.

PHAM-GIA, T., (1994). Value of the beta prior information, *Communications in Statistics - Theory and Methods*, 23 (8), pp. 2175–2195.

STOKES, L. S., (1977). Ranked Set Sampling with Concomitant variables, *Communications in Statistics - Theory and Methods*, 6 (12), pp. 1207–1211.

TERPSTRA, J. T., (2004). On Estimating a Population Proportion via Ranked Set Sampling, *Biometrical Journal*, 2, pp. 264–272.

TERPSTRA, J. T., LIUDAHL, L. A., (2004). Concomitant-based Rank Set Sampling Proportion Estimates, *Statistics in Medicine*, 23, pp. 2061–2070.

TERPSTRA, J. T., MILLER, Z. A., (2006). Exact Inference for a Population Proportion Based on a Ranked Set Sample, *Communications in Statistics: Simulation and Computation*, 35 (1), pp. 19–27.

TERPSTRA, J. T., NELSON, E. J., (2005). Optimal Rank Set Sampling Estimates for a Population Proportion, *Journal of Statistical Planning and Inference*, 127, pp. 309–321.

WANG, C., BLEI, D. M., (2015). A general method for robust Bayesian modeling, arXiv preprint arXiv:1510.05078.

WOLFE, D. A., (2010). Ranked set sampling, *Wiley Interdisciplinary Reviews: Computational Statistics*, 2 (4), pp. 460–466.

WOLFE, D. A., (2012). Ranked set sampling: its relevance and impact on statistical inference, *ISRN Probability and Statistics*.

ZAMANZADE, E., MAHDIZADEH, M., (2017). A more efficient proportion estimator in ranked set sampling, *Statistics & Probability Letters*, pp. 28–33.

ZAMANZADE, E., MAHDIZADEH, M., (2017). Estimating the population proportion in pair ranked set sampling with application to air quality monitoring, *Journal of Applied Statistics*, pp. 1–12.

APPENDIX

Algorithm: Metropolis-Hastings algorithm

The purpose of the Metropolis-Hastings (MH) algorithm is to simulate samples from a probability distribution by utilizing the full joint density function and (independent) proposals distributions corresponding to each variable of interest. The steps of Algorithm mainly consist of three components and are given below:

Initialize $x^{(0)} \uparrow q(x)$
Initialize the sample value for each random variable (this value is often sampled from the variable's prior distribution).
for iteration $i = 1, 2, \cdots$ do

Propose: $x^{cand} \uparrow q(x^{(i)}|x^{(i-1)})$
Generate a proposal (or a candidate) sample x^{cand} from the proposal distribution $q(x^{(i)}|x^{(i-1)})$

Acceptance Probability:

$$\alpha\left(x^{cand}|x^{(i-1)}\right) = Min\left\{1, \frac{q(x^{(i)}|x^{cand})\pi(x^{cand})}{q(x^{cand}|x^{(i-1)})\pi(x^{(i-1)})}\right\}$$

$u \sim \text{Uniform}(u; 0, 1)$
Compute the acceptance probability via the acceptance function $\alpha\left(x^{cand}|x^{(i-1)}\right)$ based on the proposal distribution and the full joint density $\pi(.)$

if $u < \alpha$ **then**
Accept the proposal: $x^{(i)} \leftarrow x^{cand}$
else
Reject the proposal: $x^{(i)} \leftarrow x^{(i-1)}$
end if
Accept the candidate sample with probability α, the acceptance probability, or reject it with probability $1 - \alpha$
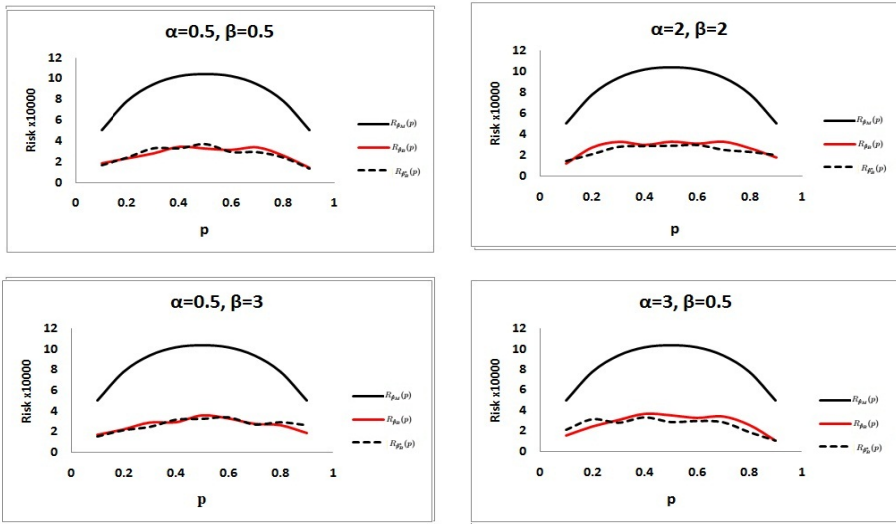
end for

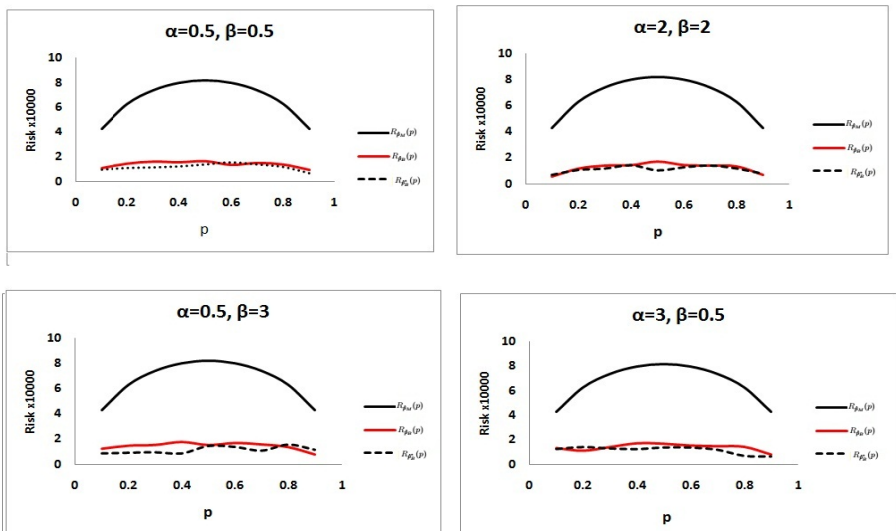Figure 1: Risk curves for the estimators $\hat{p}_M$, $\hat{p}_B$ and $\hat{p}_B^*$ when $s = 3, m = 50$ at different choices of α and β

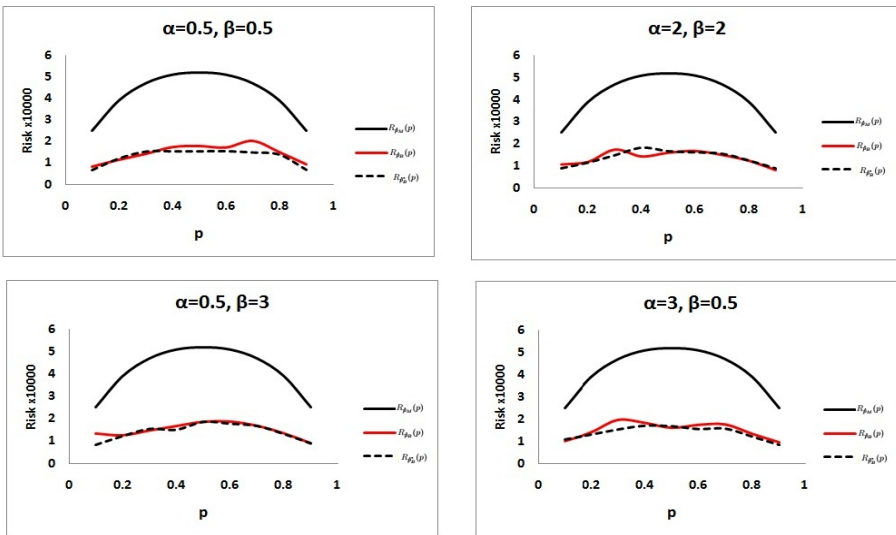Figure 2: Risk curves for the estimators $\hat{p}_M$, $\hat{p}_B$ and $\hat{p}_B^*$ when $s = 5, m = 30$ at different choices of α and β

Figure 3: Risk curves for the estimators $\hat{p}_M$, $\hat{p}_B$ and $\hat{p}_B^*$ when $s = 3, m = 100$ at different choices of α and β
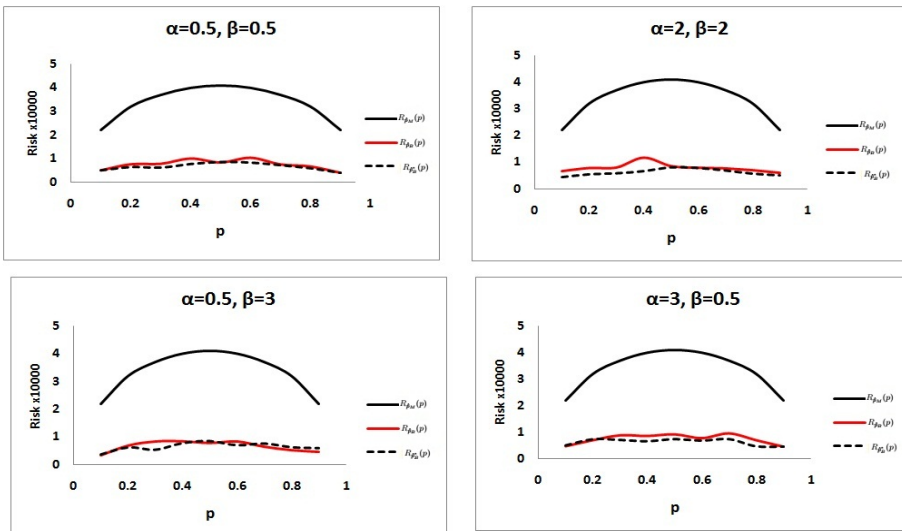

Figure 4: Risk curves for the estimators $\hat{p}_M$, $\hat{p}_B$ and $\hat{p}_B^*$ when $s = 5, m = 60$ at different choices of α and β