On Bayesian Analysis of a General Class of Randomized Response Models In Social Surveys about Stigmatized Traits

Zawar Hussain[1*], Muhammad Abid[1], Javid Shabbir[1] and Nasir Abbas[2]

[1]Department of Statistics, Quaid-i-Azam University
[2] Department of Statistics, Government Post graduate College Jhang

Abstract: While conducting a social survey on stigmatized/sensitive traits, obtaining efficient (truthful) data is an intricate issue and estimates are generally biased in such surveys. To obtain trustworthy data and to reduce false response bias, a technique, known as randomized response technique, is now being used in many surveys. In this study, we performed a Bayesian analysis of a general class of randomized response models. Suitable simple Beta prior and mixture of Beta priors are used in a common prior structure to obtain the Bayes estimates for the proportion of a stigmatized/sensitive attributes in the population of interest. We also extended our proposal to stratified random sampling. The Bayes and the maximum likelihood estimators are compared. For further understanding of variability, we have also compared the prior and posterior distributions for different values of the design constants through graphs and credible intervals. The condition to develop a new randomized response model is also discussed.

Key words: Bayesian Estimation, Randomized Response, Credible Intervals, Mean Square Error, Beta Distribution, Mixture Prior, Stratified Random Sampling, Sensitive surveys.

1. Introduction

Although sensitive (stigmatized or socially unacceptable) characteristics (induced abortion, tax evasion, cheating in the exams, drug usage, illicit behaviors, etc.) are possessed by individuals in every society, yet there are a few studies on estimating the true incidence rate of sensitive traits. This shortage in the research on such sensitive characteristics is mainly due to legal, moral and ethical issues related to them as these issues make it difficult to conduct a study on different aspects of sensitive traits, and also affect the quality of the information obtained through surveys. Collecting good information on such characteristics may be difficult, because it requires asking respondents to articulate the often complex and sensitive process that led them to the decision of possessing illicit behaviors. The eminence of the data is badly affected by the method of data collection used in these types of social surveys. For example, in a survey on induced abortion, if

---

* Corresponding author.

a woman is asked to report whether she has had abortion(s) in an official form or if she must answer questions presented by medical staff, she may be less willing (or unwilling) to report the truth than if she is questioned informally or indirectly. Data quality can also be affected by the method of data collection (i.e., gathered through a personal interview by a trained interviewer, collected by means of a self-administered questionnaire, or assembled by medical service providers and entered into official records). Finally, the legal status of sensitive characteristics may affect respondents' willingness to report truthfully on ones status about the possession of sensitive characteristics. And, regardless of the legal climate, respondents may (consciously or unconsciously) give socially acceptable answers rather than the actual status.

All of these possible biases, which can fluctuate from one study to another, might affect the worth of the data obtained. In brief, the differences in the conclusions of different surveys could be due to inconsistency of data. Among other complications, the cooperation of the study population is inconsistent and unreliable; and the use of questionnaires and other standardized survey instruments may be insufficient for revealing sensitive characteristics. Jones and Forrest [6] reported that in longitudinal and retrospective studies, approximately 50 percent of the women that have had an abortion will conceal their past abortion(s) from interviewers. Adler [19] concluded that even in short-term follow-up studies, there are high sample abrasion rates, typically in between 20 to 60 percent. Lazarus and Stern [2] were of the view that the efficiency of regular (or standardized) questionnaires might be doubted because this survey technique may not be trustworthy in finding the prevalence of abortions. Given such complexities in the data collected through retrospective surveys, it can be understood that research on sensitive issues may be insufficient to draw general conclusions about the prevalence of sensitive traits at population level. The many limitations of the studies mentioned above must be borne in mind when the results are interpreted.

It should also be noted that the research difficulties (obstacles) discussed above– concealment, attrition, time variance, and untrustworthy survey tools – are likely to result in the underestimation of incidence rate of stigmatizing (or socially unacceptable) traits in any study. Therefore, estimates of rate of incidence of sensitive traits must always be interpreted with great care.

To avoid these difficulties, an ingenious interviewing procedure, known as the Randomized Response Technique (*RRT*), for eliciting information on sensitive characteristic was introduced by Warner [27]. Its aim was to reduce the frequency of false (misleading) answers, increase respondents' cooperation and obtaining trustworthy (truthful) responses by giving the respondents two questions, one of which is sensitive and other is non-sensitive. With the help of a randomization device (spinning a roulette wheel, drawing a card from deck, rolling a die, etc.) the respondents choose one of the two questions. It makes the interviewees less likely to give a false answer. Several research workers, such as Greenberg et al. [4], Moors [11], Folsom et al. [25], Mangat [20], Singh and Horn [28], and Christofides [29], have contributed to the further development of the Warner [27] model. The interested readers may refer to Chaudhuri and Mukerjee [1], Tracy and Mangat [5], Perri [23], Yu et al. [14], Hussain and Shabbir [31, 32, 33 and 34], Odumade and Singh [22] and the references listed therein.

Once the data have been collected through *RRT*, then the stage of estimating the parameter(s) follows. For the estimation purposes, in general, we come across two recognized methods, namely, the classical method and the Bayesian method. The utility of the Bayesian method cannot be disregarded in the studies where prior information about the unknown parameter(s) is on hand and can be used along with the sample information for estimating the unknown population parameter(s). The Bayesian skeleton provides an accepted way to study and interpret situations such as randomized response sampling, where only partial information is available. Although the studies on the Bayesian analysis of *RRT*s have been done, yet a small quantity of texts on the Bayesian analysis of the *RRT*s is available. Winkler and Franklin [26], Pitz [8], Spurrier and Padgett [13], O'Hagan [3], Oh [18], Migon and Tachibana [9], Unnikrishnan and Kunte [21], Bar-Lev et al. [7], Barabesi and Marcheselli [15,16], Kim et al. [17], Hussain and Shabbir [32,33], Hussain et al. [35] and Song and Kim [12] are the major references on the Bayesian analysis of the *RRT*s.

To the best of our knowledge, Winkler and Franklin [26] were the first to study the *RRT* in the Bayesian structure. Using a Beta prior distribution, they were able to show that the posterior distributions were a mixture of Beta distributions. They suggested to use conjugate priors but did not use it and this suggestion was used by Bar-Lev at al. [7] assuming the truncated Beta distributions. O'Hagan [3] presented the Bayes linear estimator for estimating the proportion of the variable of interest. He showed that the Bayes linear estimators are simple and robust. Migon and Tachibana [9] suggested some approximations and Unnikrishnan and Kunte [21] recommended a Gibbs sample approximation. Later, Kim et al. [17] did a Bayesian analysis of Mangat [20] *RRT*. More recently, Barabesi and Marcheselli [16] provided a Bayesian approach to jointly estimate the population proportion and sensitivity level of sensitive traits by introducing a two-stage randomized response procedure. Now, taking the Winkler and Franklin [26] idea of specifying prior information in full and analyzing the posterior mixture distribution (as done in Kim et al. [17], and Barabesi and Marcheselli [16]), we plan to study a *general* class of *RRT*s yielding the probability of a *yes* response given by:

$$P(yes) = \delta = c\pi_A + g,$$ (1.1)

where c and g are *RRT* dependent real numbers, π_A is the true but known population proportion of individuals with sensitive trait. Now, from (1.1), we have

$$\pi_A = \frac{\delta - g}{c}, \quad 0 \leq \pi_A \leq 1.$$ (1.2)

The Maximum Likelihood estimator (MLE) as well as the moment estimator of π_A is given by:

$$\hat{\pi}_A = \frac{\hat{\delta} - g}{c},$$ (1.3)

where $\hat{\delta} = \frac{n_1}{n}$, the sample proportion of *yes* responses in a sample of size *n*.

Its variance is given by

$$Var(\hat{\pi}_A) = \frac{\delta(1-\delta)}{nc^2}. \tag{1.4}$$

In the frequentist approach, the moment estimators or the maximum likelihood estimators are generally implemented. But the major disadvantage associated with them is that they can assume values outside the interval (0, 1) (cf. Kim et al. [17], and Barabesi and Marcheselli [16]) because from (1.3) it is apparent that $\pi_A < 0$ if $(\hat{\delta} - g) < 0$. That is, we may have an unsatisfactory estimator. Therefore the Bayesian estimation technique may be applied in social sensitive surveys when prior information is on hand to cut the damage of information due to the use of randomization devices. Another advantage of using the Bayes estimators is that they take the values in the interval (0, 1). So, applying the Bayesian estimation technique may be justified and advantageous. The Bayesian estimation may be applied in stratified random sampling (STRS) when it is suitable to stratify the population into sub-populations. Similarly, the situations may arise when different researchers have different prior beliefs. In this paper, we plan to study the Bayesian estimation under simple random sampling with replacement (SRSWR) and stratified random sampling (STRS). We also plan to study the affect of using mixture prior distribution. We have arranged our paper as follows. In the section 2, we present the Bayesian estimation in case of SRSWR and STRS assuming simple Beta prior and apply mixture prior only in case of SRSWR. Section 3 contains the efficiency comparisons. In Section 4, we apply the suggested estimation on a real dataset. We conclude our study in Section5. We assume squared error loss function throughout the paper.

2. Bayesian Estimation of π_A Using General Randomized Response Model

While planning a survey study, situation may arise that the investigator has some prior knowledge about the parameter of interest. His/her prior knowledge may be in form of a point guess or a probability distribution. We assume here that prior knowledge is quantified in the form of a probability distribution and more specifically, it is a Beta distribution. The reason of using Beta prior is two-fold. The Beta distribution is a conjugate prior and covers the case of uniform prior when both the hyperparameters are equal to one. Also with Beta prior, we have a closed form expression of the Bayes estimator.

Under the assumption made above the prior distribution of π_A is given by:

$$f(\pi_A) = \frac{1}{\beta(a,b)} \pi_A^{a-1} (1-\pi_A)^{b-1}, \quad 0 < \pi_A < 1, \ a,b > 0. \tag{2.1}$$

where a and b are the hyperparameters and $\beta(a,b)$ is the usual Beta function. Using the prior distribution in (2.1), we now develop Bayesian estimation for a general class of RRT yielding the probability of a yes answer as in (1.1), under three different cases: (i) simple random sampling and single prior, (ii) STRS and single prior, and (iii) SRSWR and mixture prior.

(i) Case of SRSWR and single prior

For a given model (belonging to general class) with probability of *yes* as defined in (1.1), let $t = \sum_{i=1}^{n} x_i$ be the total number of *yes* responses in a sample of size n drawn from the population with simple random sampling with replacement. Here, $x_i = 1$ with probability δ, and $x_i = 0$ with probability $(1-\delta)$. Then, the conditional distribution of t given π_A is given by:

$$f(t|\pi_A) = \binom{n}{t} \delta^t (1-\delta)^{n-t}. \tag{2.2}$$

Substituting (1.1) in (2.2), we get:

$$f(t|\pi_A) = \frac{n!}{t!(n-t)!}(c\pi_A + g)^t (1 - c\pi_A - g)^{n-t}. \tag{2.3}$$

On simplification, we get:

$$f(t|\pi_A) = \frac{n!(c)^n}{t!(n-t)!}\sum_{i=0}^{t}\sum_{j=0}^{n-t}\frac{(n-t)!}{j!(n-t-j)!}\frac{t!}{i!(t-i)!}d^{t-i}h^{n-t-j}\pi_A^{i}(1-\pi_A)^{j} \tag{2.4}$$

where $d = \dfrac{g}{c}$, and $h = \dfrac{1-c-g}{c}$, $0 < g, c < 1, (g+c) < 1$. The condition $(g+c) < 1$, guarantees the existence of Bayesian estimator assuming positive values.

Thus, the joint distribution of t and π_A is given by:

$$f(t,\pi_A) = \frac{n!(c)^n}{\beta(a,b)t!(n-t)!}\sum_{i=0}^{t}\sum_{j=0}^{n-t}\frac{(n-t)!}{j!(n-t-j)!}\frac{t!}{i!(t-i)!}d^{t-i}h^{n-t-j}\pi_A^{a+i-1}(1-\pi_A)^{b+j-1} \tag{2.5}$$

Now the posterior distribution of π_A given t may be obtained as:

$$f(\pi_A|t) = \frac{\displaystyle\sum_{i=0}^{t}\sum_{j=0}^{n-t}\frac{(n-t)!}{j!(n-t-j)!}\frac{t!}{i!(t-i)!}d^{t-i}h^{n-t-j}\pi_A^{a+i-1}(1-\pi_A)^{b+j-1}}{\displaystyle\sum_{i=0}^{t}\sum_{j=0}^{n-t}\frac{(n-t)!}{j!(n-t-j)!}\frac{t!}{i!(t-i)!}d^{t-i}h^{n-t-j}\beta(a+i,b+j)}, \quad (0 < \pi_A < 1). \tag{2.6}$$

Under the squared error loss function, the Bayes estimator of π_A is given by

$$\hat{\pi}_{A(Bayes)} = \frac{\displaystyle\sum_{i=0}^{t}\sum_{j=0}^{n-t}\frac{(n-t)!}{j!(n-t-j)!}\frac{t!}{i!(t-i)!}d^{t-i}h^{n-t-j}\beta(a+i+1,b+j)}{\displaystyle\sum_{i=0}^{t}\sum_{j=0}^{n-t}\frac{(n-t)!}{j!(n-t-j)!}\frac{t!}{i!(t-i)!}d^{t-i}h^{n-t-j}\beta(a+i,b+j)}. \tag{2.7}$$

(i) Case of STRS and single prior

Let the population under study is divided into L strata of sizes $N_1, N_2, ..., N_L$. Let $n_l \ (l = 1, 2, ..., L)$ be the optimum sample size from the lth stratum such that $\sum_{l=1}^{L} n_l = n$ and $W_l = \dfrac{N_l}{N}$ be the lth stratum weight. Assuming $Beta(a_l, b_l), (l = 1, 2, ..., L)$ as the prior distribution in l^{th} stratum, the stratified Bayes estimator of π_A is written as

$$\hat{\pi}_{A(Str.Bayes)} = \sum_{l=1}^{L} W_l \hat{\pi}_{A(Bayes)_l}, \text{ where}$$

$$\hat{\pi}_{A(Bayes)_l} = \frac{\displaystyle\sum_{i=0}^{t_l} \sum_{j=0}^{n_l - t_l} \frac{(n_l - t_l)!}{j!(n_l - t_l - j)!} \frac{t_l!}{i!(t_l - i)!} d^{t_l - i} h^{n_l - t_l - j} \beta(a_l + i + 1, b_l + j)}{\displaystyle\sum_{i=0}^{t_l} \sum_{j=0}^{n_l - t_l} \frac{(n_l - t_l)!}{j!(n_l - t_l - j)!} \frac{t_l!}{i!(t_l - i)!} d^{t_l - i} h^{n_l - t_l - j} \beta(a_l + i, b_l + j)}. \tag{2.8}$$

Similarly, when *MLE* is computed from each stratum, the stratified *MLE* is given by

$$\hat{\pi}_{A(ML)} = \sum_{l=1}^{L} W_l \hat{\pi}_{A_l}, \tag{2.9}$$

where

$$\hat{\pi}_{A_l} = \frac{\hat{\delta}_l - g}{c}. \tag{2.10}$$

(ii) Case of SRSWR and mixture prior

We may use mixture prior when different researchers have different beliefs about the prior distribution. Let $f_m(\pi_A)$ be the prior belief (in the form of a probability distribution function) of the $m^{th} \ (m = 1, 2, ..., M)$ researcher, then a mixture of the priors $f = \sum_{m=1}^{M} q_m f_m$ should preferably be used in order to incorporate all the beliefs, where q_m is the weight attached to the prior belief of the mth researcher. Let $f_m(\pi_A) = Beta(a_m, b_m)$, where $Beta(a_m, b_m)$ is a Beta distribution with parameters a_m and b_m, then a mixture prior is given by $f = \sum_{m=1}^{M} q_m Beta(a_m, b_m)$.

Under the mixture prior and simple random sampling with replacement, the Bayes estimator can be derived following the steps in the case (i) above. The Bayes estimator, now, is given by

$$\hat{\pi}_{A(Bayes)_{Mixture}} = \frac{\sum_{m=1}^{M}\left[\dfrac{q_m}{\beta(a_m,b_m)}\left\{\sum_{i=0}^{t}\sum_{j=0}^{n-t}\dfrac{(n-t)!}{j!(n-t-j)!}\dfrac{t!}{i!(t-i)!}d^{t-i}h^{n-t-j}\beta(a_m+i+1,b_m+j)\right\}\right]}{\sum_{m=1}^{M}\left[\dfrac{q_m}{\beta(a_m,b_m)}\left\{\sum_{i=0}^{t}\sum_{j=0}^{n-t}\dfrac{(n-t)!}{j!(n-t-j)!}\dfrac{t!}{i!(t-i)!}d^{t-i}h^{n-t-j}\beta(a_m+i,b_m+j)\right\}\right]} \qquad (2.11)$$

3. Efficiency Comparisons and Discussion

Prior to performing a comparison study of the maximum likelihood and the Bayes estimators, it is to be mentioned that the Bayes estimator is computationally more difficult when the sample size and/or the number of *yes* responses increases. We used *R* software to handle this computational issue.

Once a Bayes estimator is calculated, the task of a Bayesian is more or less done but he/she may be attracted to compare his/her estimator with competing estimators (e.g., the *MLEs*). He/she may also be looking to understand that how his/her prior estimate (prior mean) has been affected (updated) by the addition of sample information. On the other hand, a frequentist by and large makes comparisons of the competitor estimators in terms of *MSE*. Analysis of the posterior distribution can give a picture of the effect of sample information on the prior estimate. That is, the posterior mean, the posterior variance, the posterior median and the credible intervals clearly show the effect of sample information. For the interest of Bayesians, the posterior means, standard deviations (*SD*) and 95 % credible intervals (*CI*), using the posterior distribution defined in (2.6), are given in Tables 1-4 below and the Tables S1-S2 in supplementary material.

Table 1: Description of posterior distribution for $n = 15$, $t = 3$

G	c	Beta (2,4)			Beta (10,20)		
		Mean	SD	95% C.I	Mean	SD	95% C.I
0.1	0.1	0.359202	0.180027	0.0300, 0.6900	0.33944	0.0848848	0.1765, 0.5050
	0.2	0.35212	0.173009	0.0545,0.7050	0.338241	0.0838906	0.1865, 0.5150
	0.3	0.330576	0.162666	0.0820, 0.7550	0.333177	0.0823902	0.1965, 0.5350
	0.4	0.303455	0.150332	0.0840, 0.7700	0.325778	0.0805497	0.1980, 0.5390
	0.5	0.27561	0.137172	0.0760, 0.8250	0.316848	0.0784177	0.1850, 0.5100
	0.6	0.249608	0.12434	0.0700, 0.8250	0.306897	0.0760238	0.1800, 0.4900
	0.7	0.226495	0.112616	0.0650, 0.8290	0.296293	0.0734063	0.1600, 0.4450
	0.8	0.206419	0.102302	0.0590, 0.8790	0.28533	0.0706171	0.1500, 0.4250
0.3	0.1	0.305564	0.169928	0.0650, 0.8090	0.326928	0.0837094	0.1850, 0.5250
	0.2	0.271896	0.156669	0.0580, 0.8890	0.318309	0.0821497	0.1880, 0.5850
	0.3	0.237933	0.14046	0.0500, 0.8890	0.307945	0.0800367	0.1830, 0.8350
	0.4	0.207343	0.123882	0.0430, 0.9090	0.296262	0.0774161	0.1770, 0.8200
	0.5	0.18154	0.108854	0.0380, 0.9290	0.283674	0.0743582	0.1690, 0.8300

	0.6	0.160385	0.0961091	0.0340, 0.9300	0.27059	0.0709668	0.1610, 0.8500
0.7	0.1	0.228555	0.140329	0.0460, 0.9490	0.305453	0.0803889	0.1800, 0.8800
	0.2	0.15833	0.101161	0.0310, 0.9690	0.274388	0.0738412	0.1600, 0.8800

Table 2: Description of posterior distribution for $n = 15$, $t = 6$

G	c	Beta (2,4)			Beta (10,20)		
		Mean	SD	95% C.I	Mean	SD	95% C.I
0.1	0.1	0.443357	0.185444	0.0935,0.7891	0.358265	0.0863934	0.1865,0.5225
	0.2	0.479163	0.17611	0.1405,0.8099	0.368809	0.0857952	0.1975,0.5316
	0.3	0.480954	0.167986	0.1600,0.8201	0.371781	0.084598	0.2098,0.5385
	0.4	0.466067	0.160582	0.1590,0.7752	0.370322	0.0831581	0.2210,0.5477
	0.5	0.441898	0.152432	0.1515,0.7380	0.365957	0.0815214	0.2190,0.5388
	0.6	0.412906	0.142946	0.1445,0.6955	0.359511	0.0796667	0.2115,0.5220
	0.7	0.382472	0.132477	0.1330,0.6437	0.351498	0.0775693	0.2055,0.5073
	0.8	0.353021	0.121887	0.1240,0.5939	0.342288	0.0752216	0.2000,0.4926
0.3	0.1	0.346763	0.179659	0.0330,0.6865	0.336509	0.0848436	0.1755,0.5040
	0.2	0.343635	0.175277	0.0365,0.6733	0.336217	0.0842715	0.1770,0.5034
	0.3	0.328108	0.166559	0.0395,0.6512	0.333003	0.0830934	0.1780,0.5001
	0.4	0.304965	0.154534	0.0483,0.6300	0.327319	0.0813721	0.1790,0.4955
	0.5	0.27881	0.140679	0.0160,0.4594	0.319566	0.079138	0.1775,0.4865
	0.6	0.25314	0.126764	0.0248,0.4300	0.310133	0.0764268	0.1670,0.4634
0.7	0.1	0.261647	0.154084	0.0550, 0.9690	0.315542	0.0819488	0.1880, 0.9000
	0.2	0.197579	0.121236	0.0400, 0.9720	0.293192	0.0773045	0.1740, 0.9200

Table 3: Description of posterior distribution for *n* = 75, *t* = 15

G	c	Beta (2,4)			Beta (10,20)		
		Mean	SD	95% C.I	Mean	SD	95% C.I
0.1	0.1	0.453409	0.175151	0.0950, 0.7670	0.363361	0.0852568	0.2200, 0.8050
	0.2	0.403827	0.150018	0.1200, 0.6970	0.355623	0.0806513	0.2280, 0.8050
	0.3	0.328836	0.122554	0.1350, 0.6970	0.332806	0.0747305	0.2150, 0.8150
	0.4	0.268406	0.0995718	0.1150, 0.7070	0.305414	0.0683501	0.1980, 0.8050
	0.5	0.224326	0.0826783	0.0970, 0.7270	0.278251	0.06208	0.1800, 0.8450
	0.6	0.191904	0.0703257	0.0840, 0.7270	0.25333	0.0562941	0.1650, 0.8550
	0.7	0.167367	0.0610517	0.0740, 0.7470	0.231257	0.051151	0.1510, 0.8550
	0.8	0.148258	0.0538815	0.0655, 0.7500	0.211994	0.0466643	0.1390, 0.8550
0.3	0.1	0.221947	0.136287	0.0450, 0.7950	0.302968	0.0797796	0.1790, 0.8950
	0.2	0.149876	0.0954572	0.0290, 0.8450	0.268327	0.0724182	0.1570, 0.8950
	0.3	0.110291	0.0705613	0.0220, 0.8650	0.235298	0.0642853	0.1360, 0.8950
	0.4	0.0867038	0.0554298	0.0170, 0.8650	0.206577	0.0566459	0.1200, 0.8950
	0.5	0.0712812	0.0455048	0.0130, 0.8850	0.182601	0.0500283	0.1070, 0.8950
	0.6	0.060465	0.0385476	0.0110, 0.8850	0.162833	0.0444853	0.0950, 0.9050
0.7	0.1	0.0896485	0.0613047	0.0150, 0.9550	0.222669	0.0636439	0.1270, 0.9050
	0.2	0.0486872	0.0334514	0.0090, 0.9550	0.15619	0.0457419	0.0880, 0.9050

Table 4: Description of posterior distribution for $n = 75$, $t = 30$

	c	Beta (2,4)			Beta (10,20)		
		Mean	SD	95% C.I	Mean	SD	95% C.I
0.1	0.1	0.719864	0.121937	0.1030, 0.8950	0.456502	0.0866918	0.3050,0.7050
	0.2	0.73481	0.10775	0.1130, 0.8950	0.489614	0.0806587	0.3490, 0.6850
	0.3	0.693705	0.108237	0.1590, 0.8610	0.487178	0.0764079	0.3590, 0.6850
	0.4	0.623087	0.106012	0.1690, 0.7950	0.469232	0.0726278	0.3500, 0.7250
	0.5	0.543684	0.0968202	0.2550, 0.7050	0.443624	0.0685436	0.3320, 0.7250
	0.6	0.472821	0.0853447	0.2350, 0.6150	0.414783	0.0640513	0.3120, 0.7250
	0.7	0.41496	0.075058	0.2250, 0.5450	0.385576	0.0593874	0.2900, 0.7250
	0.8	0.368451	0.0665721	0.2050, 0.4850	0.357714	0.0548398	0.2700, 0.7250
0.3	0.1	0.399779	0.181685	0.0400,0.7220	0.349178	0.0854085	0.2150, 0.7150
	0.2	0.377945	0.161968	0.0740,0.6875	0.347043	0.0825741	0.2160, 0.7150
	0.3	0.321062	0.134867	0.0765,0.5925	0.332063	0.0775623	0.2090, 0.7150
	0.4	0.267509	0.110373	0.0915,0.5715	0.310092	0.0713675	0.1970, 0.7150
	0.5	0.225937	0.0918319	0.0650,0.4222	0.285785	0.0648506	0.1820, 0.7150
	0.6	0.194462	0.0781404	0.0570,0.3610	0.262022	0.0586391	0.1700, 0.7150
0.7	0.1	0.128391	0.0853877	0.0230, 0.7050	0.256768	0.070964	0.1490, 0.8250
	0.2	0.0730347	0.0489989	0.0130, 0.7050	0.195609	0.0553926	0.1120, 0.8250

It is obvious that description of posterior distribution is not helpful in making comparisons of the Bayes estimator and the Classical estimator (*MLE*). These two estimators may be compared either in the terms of variance and/or the *MSE*. As we tend to compare the Bayes estimator and the *MLE*, we must be conscious of the method of comparison. We follow the method used by Hussain et al. [35] to compare the *MLE* and the Bayes estimators. As a first approach, we compare them on Bayesian grounds. Since Bayes estimator is based on a single sample, so the posterior variance is a measure of its variability. For this single (observed) sample, we can have the *MLE* of π_A from (1.3) and this value of the *MLE* can be substituted in (1.4) to obtain an estimated variance of the *MLE*. Finally, we can compare the posterior variance and the estimated variance of the *MLE* and calculate the relative efficiency (*RE*). The results of these comparisons (in terms of *RE*) are provided in Table 5.

Table 5: *RE* of Bayes estimator relative to *MLE* for $n = 15, \ t = 6$.

g	c	$a = 10, \ b = 20$			$a = 2, \ b = 4$		
		$SD\left(\hat{\pi}_{Bayes}\right)$	$SD\left(\hat{\pi}_A\right)$	*RE*	$SD\left(\hat{\pi}_{Bayes}\right)$	$SD\left(\hat{\pi}_A\right)$	*RE*
0.1	0.1	0.0864	1.2649	14.641	0.0848	1.2649	14.9086
	0.2	0.0858	0.6324	7.3717	0.0842	0.6324	7.5049
	0.3	0.0846	0.4216	4.9840	0.0830	0.4216	5.0742
	0.4	0.0832	0.3162	3.8027	0.0813	0.3162	3.8861
	0.5	0.0815	0.2530	3.1033	0.0791	0.2529	3.1967
	0.6	0.0797	0.2108	2.6462	0.0764	0.2108	2.7584
	0.7	0.0776	0.1807	2.3296	0.0743	0.1764	2.3741
	0.8	0.0752	0.1581	2.1020	0.0710	0.1345	1.8943
0.3	0.1	0.1796	1.2649	7.0406	0.1816	0.5656	3.1135
	0.2	0.1752	0.6324	3.6083	0.1619	0.2828	1.7462
	0.3	0.1665	0.4216	2.5314	0.1348	0.1885	1.3981
	0.4	0.1545	0.3162	2.0463	0.1103	0.1414	1.2813
	0.5	0.1406	0.2529	1.7982	0.0918	0.1131	1.2320
	0.6	0.1267	0.2108	1.6630	0.0781	0.0942	1.2065

From Table 5, it can be concluded that the in case of two different considered prior distributions, Bayesian estimators outperform the *MLE* for all permissible values of c and g. For the same purpose, graphs of the prior and posterior distributions are provided in Figures 1-12 below and Figures S1-S6 in the supplementary material for different priors and other constants (c and g).
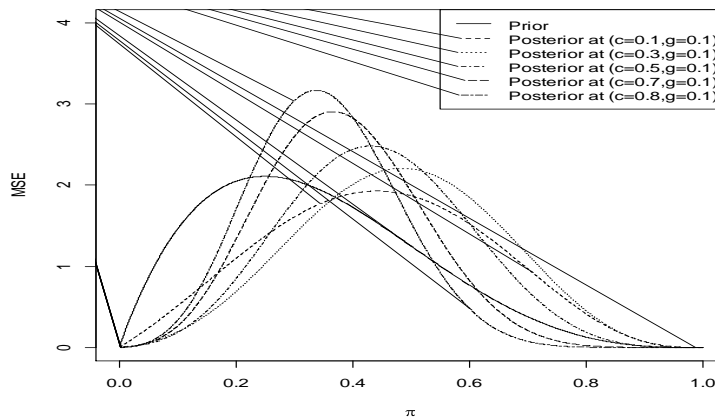
Figure 1: Posterior and prior distributions for $a = 2$, $b = 4$, $n = 15$ and $t = 6$
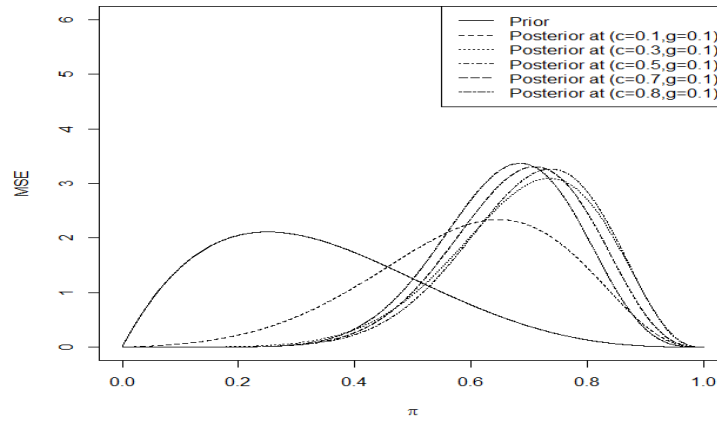
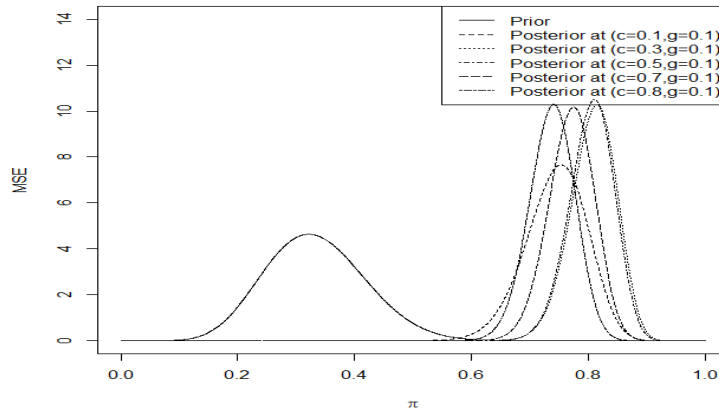Figure 2: Posterior and prior distributions for $a = 2$, $b = 4$, $n = 15$ and $t = 12$
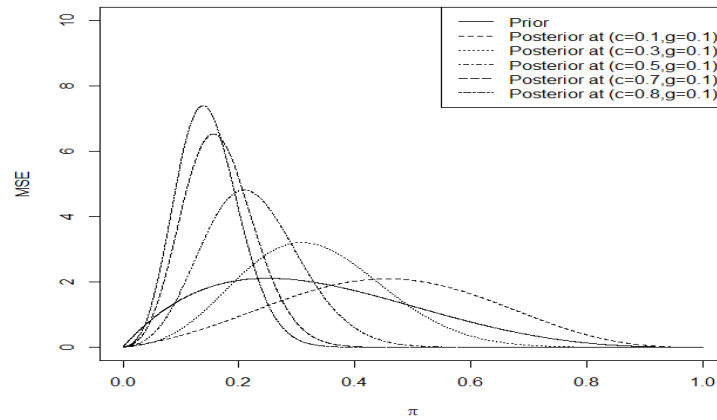
Figure 3: Posterior and prior distributions for $a = 2$, $b = 4$, $n = 75$ and $t = 15$
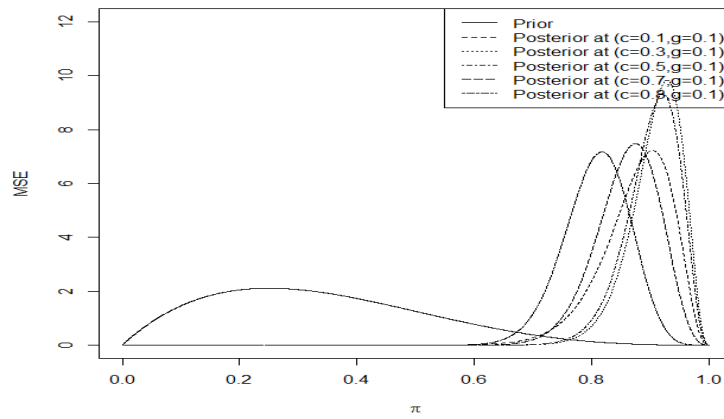
Figure 4: Posterior and prior distributions for $a = 2$, $b = 4$, $n = 75$ and $t = 60$
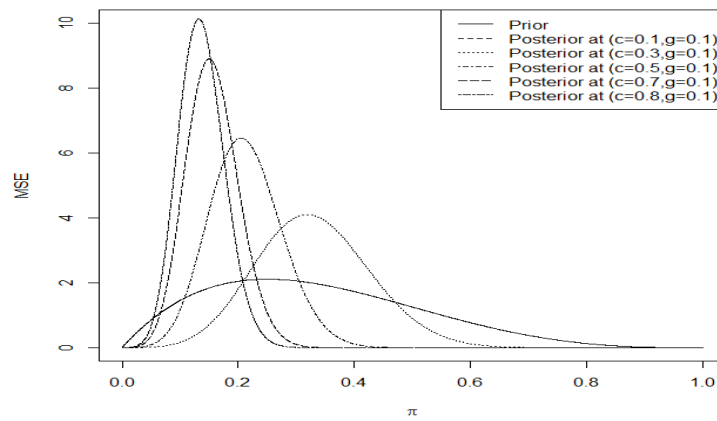
Figure 5: Posterior and prior distributions for $a = 2$, $b = 4$, $n = 150$ and $t = 30$
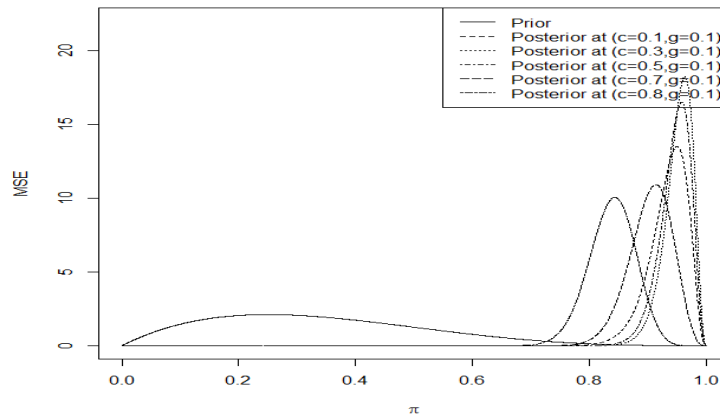
Figure 6: Posterior and prior distributions for $a = 2$, $b = 4$, $n = 150$ and $t = 120$
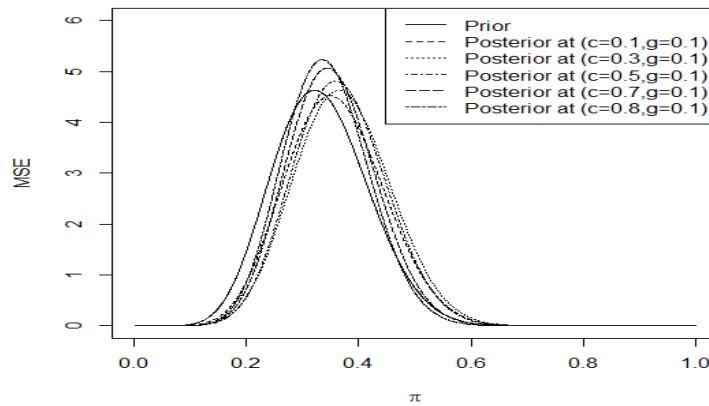
Figure 7: Posterior and prior distributions for $a = 10$, $b = 20$, $n = 15$ and $t = 6$
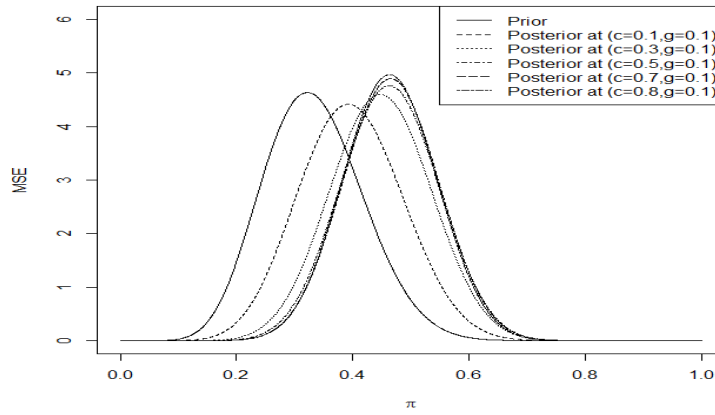
Figure 8: Posterior and prior distributions for $a = 10$, $b = 20$, $n = 15$ and $t = 12$
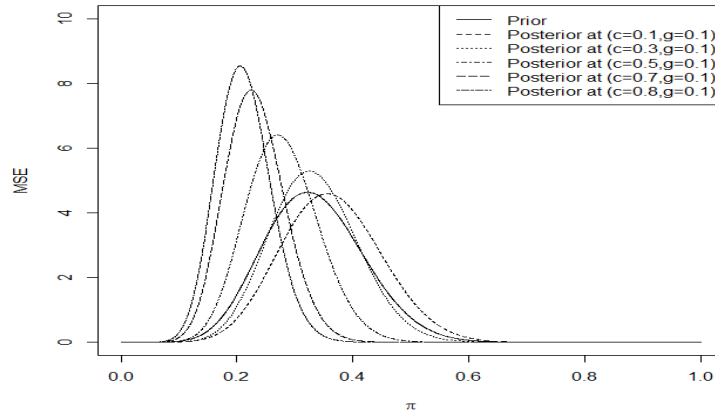
Figure 9: Posterior and prior distributions for $a = 10$, $b = 20$, $n = 75$ and $t = 15$
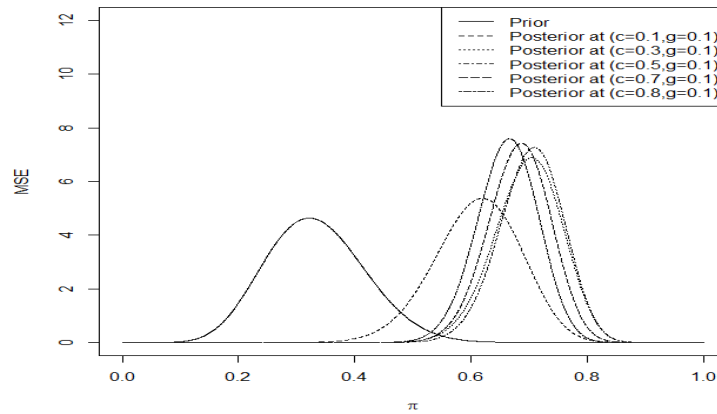
Figure 10: Posterior and prior distributions for $a = 10$, $b = 20$, $n = 75$ and $t = 60$
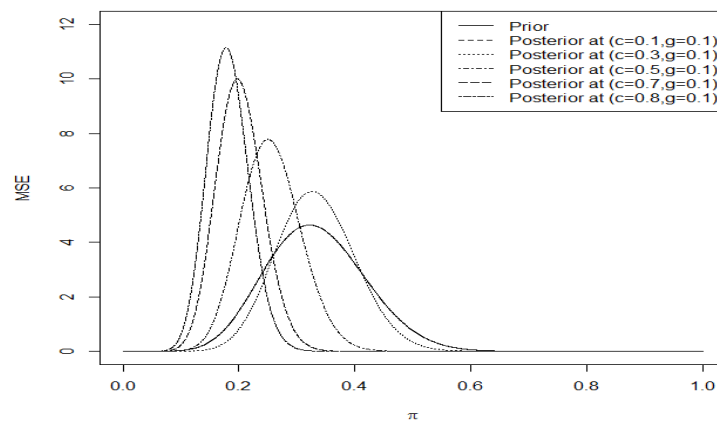
Figure 11: Posterior and prior distributions for $a = 10$, $b = 20$, $n = 150$ and $t = 30$

Figure 12: Posterior and prior distributions for $a = 10$, $b = 20$, $n = 150$ and $t = 120$

Secondly, we compare the *MLE* and the Bayes estimator on frequentist grounds in terms of *MSE*. To do this, we use the approach by Chaubey and Li [30] to compare the Bayesian and the classical estimators and do not use loss function. This approach, in fact, uses loss function only to obtain the Bayes estimator and there is no further role of loss function. This approach has also been used by Kim et al. [17] and Hussain et al. [35]. We also intend to use it in our study.

The *MSEs* of the Bayes estimators with simple Beta priors and the classical estimators under simple random sampling are defined as:

$$MSE(\hat{\pi}_{A_{(Bayes)}}) = E(\pi_{A_{(Bayes)}} - \pi_A)^2 = \sum_{t=0}^{n} (\hat{\pi}_{A_{(Bayes)}} - \pi_A)^2 \frac{n!}{t!(n-t)!} \delta^t (1-\delta)^{n-t}, \qquad (3.1)$$

and

$$MSE(\hat{\pi}_A) = E(\pi_A - \pi_A)^2 = \sum_{t=0}^{n} (\hat{\pi}_A - \pi_A)^2 \frac{n!}{t!(n-t)!} \delta^t (1-\delta)^{n-t}. \qquad (3.2)$$

Similarly, the *MSEs* of the Bayes estimator with mixture prior and the stratified Bayes estimator are respectively defined as

$$MSE\left(\hat{\pi}_{A(Bayes)_{Mixture}}\right) = E\left(\pi_{A(Bayes)_{Mixture}} - \pi_A\right)^2$$

$$= \sum_{t=0}^{n} \left(\hat{\pi}_{A(Bayes)_{Mixture}} - \pi_A\right)^2 \frac{n!}{t!(n-t)!} \delta^t (1-\delta)^{n-t}. \qquad (3.3)$$

$$MSE\left(\hat{\pi}_{A(Str.Bayes)}\right) = E\left(\pi_{A(Str.Bayes)} - \left(\pi_A\right)_l\right)^2$$

$$= \sum_{l=1}^{L} W_l^2 \left\{ \sum_{t_l=0}^{n_l} \left(\hat{\pi}_{A(Bayes)_l} - \left(\pi_A\right)_l\right)^2 \frac{n_l!}{t_l!(n_l-t_l)!} \delta_l^{t_l} (1-\delta_l)^{n_l-t_l} \right\}$$

$$+ \left[\sum_{l=1}^{L} W_l \left\{ \sum_{t_l=0}^{n_l} \left(\hat{\pi}_{A(Bayes)_l} - \left(\pi_A\right)_l\right) \frac{n_l!}{t_l!(n_l-t_l)!} \delta_l^{t_l} (1-\delta_l)^{n_l-t_l} \right\} \right]^2 . \qquad (3.4)$$

The *MSE* of *MLE* under STRS is defined as

$$MSE\left(\hat{\pi}_{A(ML)}\right) = E\left(\pi_{A(ML)} - \pi_A\right)^2 = \sum_{l=1}^{L} W_l^2 \left\{ \sum_{t_l=0}^{n_l} \left(\hat{\pi}_{A_l} - \left(\pi_A\right)_l\right)^2 \frac{n_l!}{t_l!(n_l-t_l)!} \delta_l^{t_l} (1-\delta_l)^{n_l-t_l} \right\}$$

$$+ \left[\sum_{l=1}^{L} W_l \left\{ \sum_{t_l=0}^{n_l} \left(\hat{\pi}_{A_l} - \left(\pi_A\right)_l\right) \frac{n_l!}{t_l!(n_l-t_l)!} \delta_l^{t_l} (1-\delta_l)^{n_l-t_l} \right\} \right]^2 . \qquad (3.5)$$

Given any randomized response model using (1.1), we can compare the Bayes estimators and the *MLE* in terms of *MSEs*.

To compare the *MLE* and the Bayes estimator with single prior, from classical point of view, using (3.1) and (3.2), the graphical comparisons are provided in Figures 13-37 below and the Figures S7-S25 in the supplementary material. Similarly, using (3.2) and (3.3) the Bayes estimator under mixture prior is compared with the *MLE*. The results are displayed graphically in Figures 38-39 and the Figure S26 in the supplementary material. To compare the *MLE* and Bayes estimator, under STRS, we have used the optimum allocation of total sample size into different strata. Assuming that study population is divided into two strata, using (3.4) and (3.5), the *MLE* and the Bayes estimators are compared and *RE* results are given in Table 6.

From Figures 13-37, it is observed that the Bayes estimators with different Beta priors perform better than the *MLE* over a wide range of population parameter π. Also from Figures 38-39, the superiority of the Bayes estimator with the mixture prior is apparent. From Figures 13-39, it is also observed that the relative efficiency of the Bayes estimator is not much affected by the sample size and the choice of prior, but it is affected by the constants c and g .

In case of STRS, from Table 6, we can see that the Bayes estimator performs well. We observed that when prior distribution is correctly specified then any value of the constants c and g can be chosen. On the other hand, if prior is mis-specified then $(g-c)$ must be positive and large to obtain the maximum efficiency. Similar pattern is observed for the other sample sizes but the extent of the relative efficiency is decreased when the sample size is increased.

Table 6: *RE* for Bayes estimator relative to *MLE* under stratified random sampling using optimum sample sizes

$W_1 = 0.2, W_2 = 0.8$	c	$g = 0.1$	$g = 0.3$	$g = 0.5$	$g = 0.7$	$g = 0.9$
$\pi_1 = 0.02, \pi_2 = 0.18$	0.1	1.723	11.226	39.460	38.211	31.015
	0.3	1.843	13.623	18.829	13.38	2.387
$\pi_1 = 0.05, \pi_2 = 0.15$	0.1	1.686	10.730	40.142	39.446	34.428
	0.3	1.772	12.381	18.320	12.752	2.117
$\pi_1 = 0.15, \pi_2 = 0.05$	0.1	1.682	9.182	42.078	42.709	37.665
	0.3	0.3316	8.498	16.400	10.6346	1.553
$\pi_1 = 0.18, \pi_2 = 0.02$	0.1	1.678	8.752	42.545	43.379	35.789
	0.3	0.335	7.495	15.801	10.057	1.546
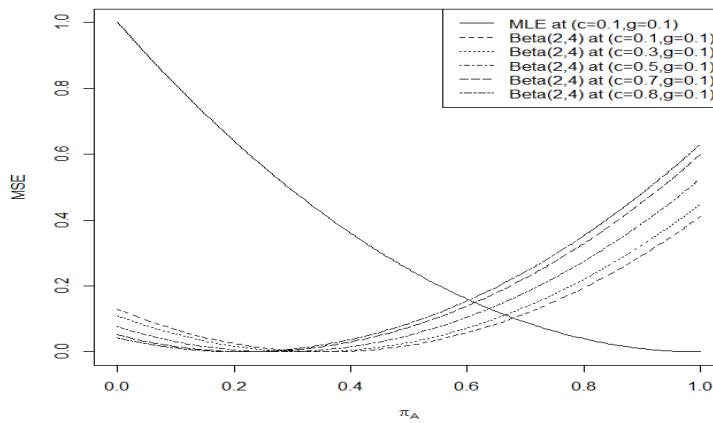
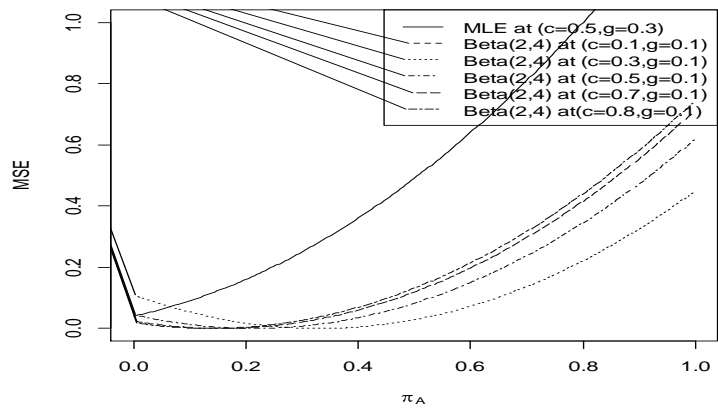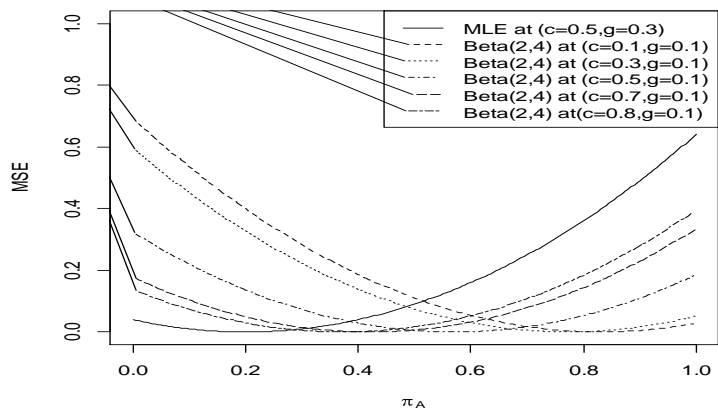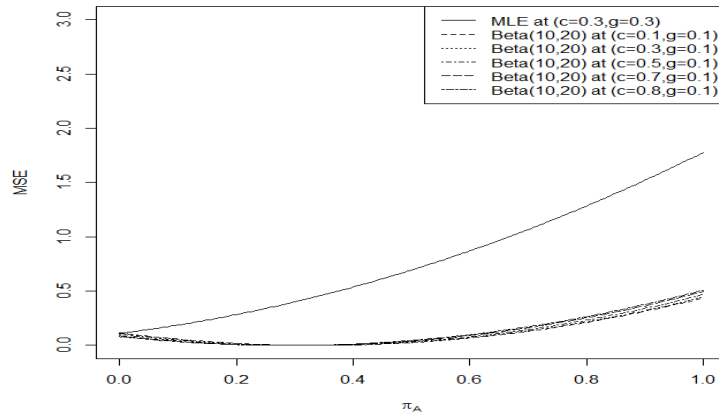Figure 13: *MSE* behavior of *MLE* and Bayes Estimator $n = 15$, $t = 3$
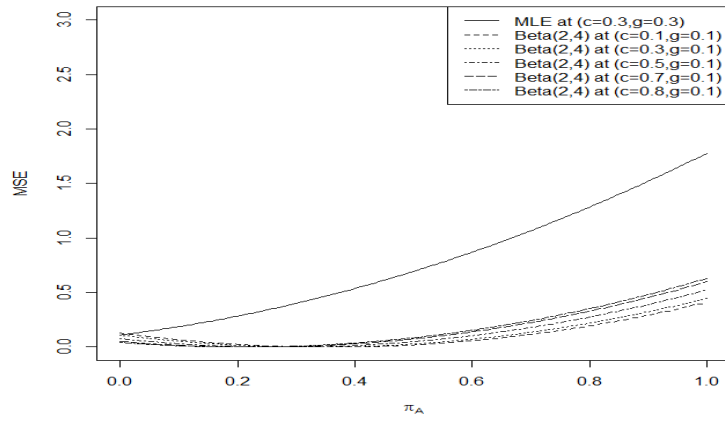
Figure 14: *MSE* behavior of *MLE* and Bayes Estimator $n = 15$, $t = 3$
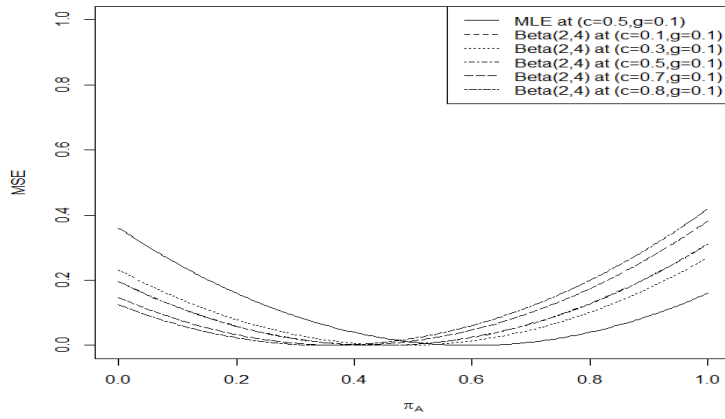
Figure 15: *MSE* behavior of *MLE* and Bayes Estimator $n = 15$, $t = 6$
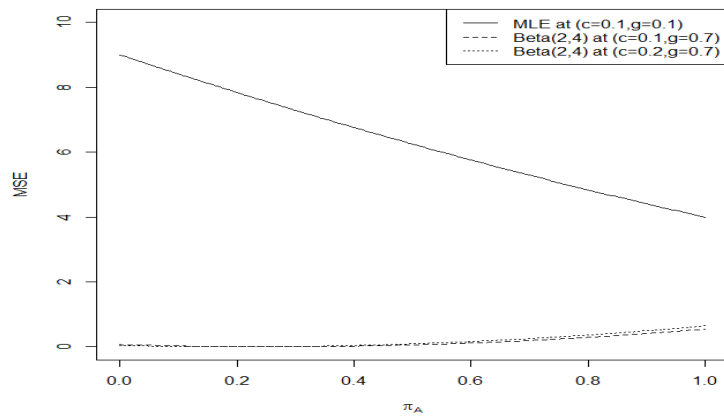
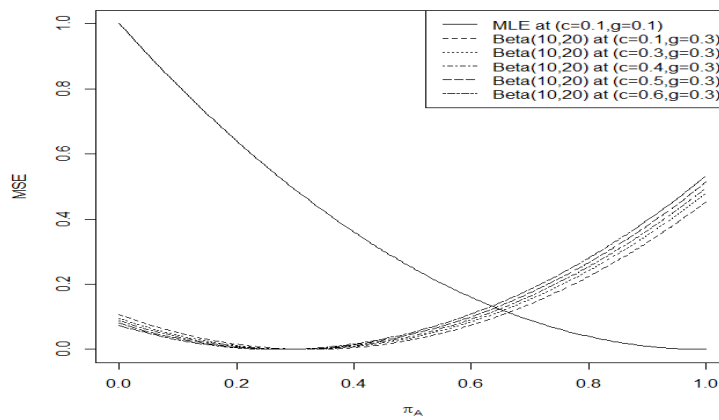Figure 16: *MSE* behavior of *MLE* and Bayes Estimator *n* = 15, *t* = 6
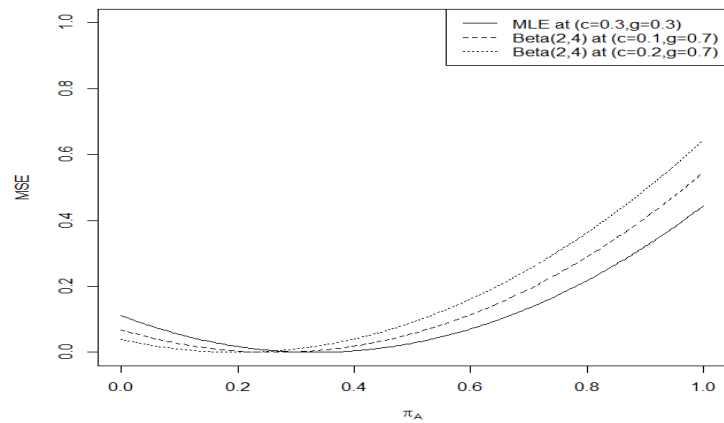
Figure 17: *MSE* behavior of *MLE* and Bayes Estimator *n* = 15, *t* = 6
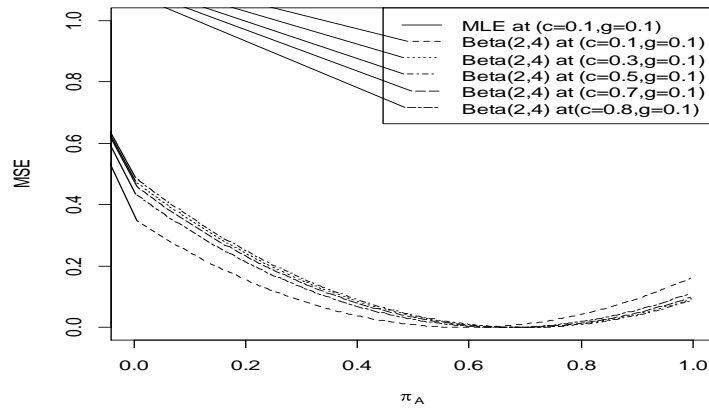
Figure 18: *MSE* behavior of *MLE* and Bayes Estimator $n = 15$, $t = 12$
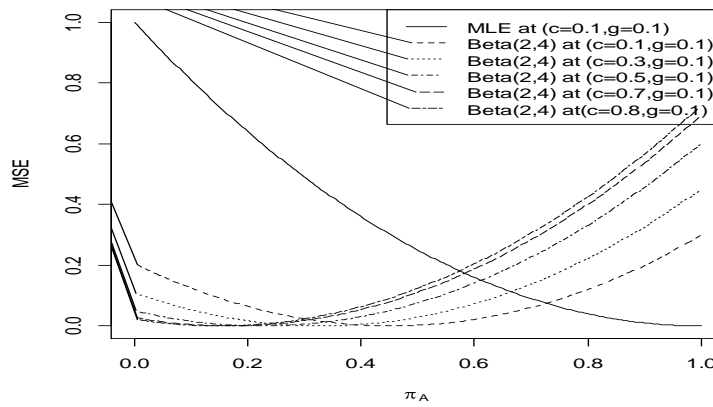
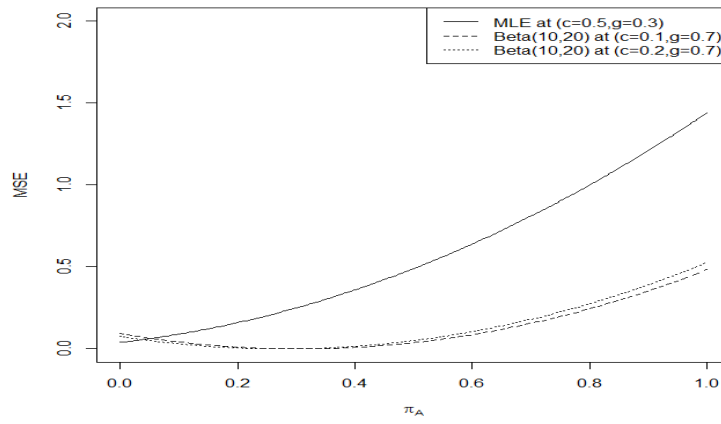Figure 19: *MSE* behavior of *MLE* and Bayes Estimator $n = 75$, $t = 15$
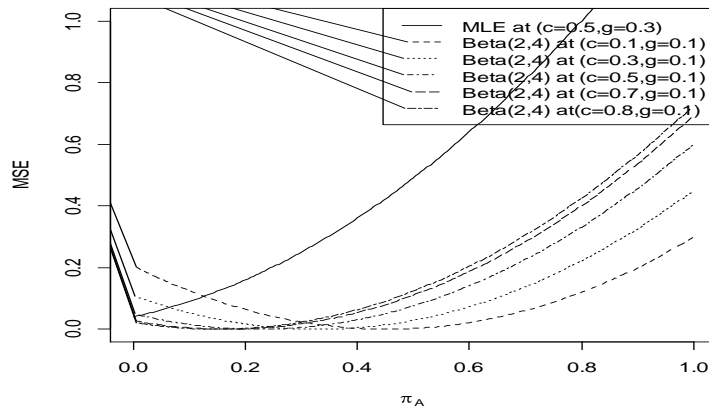
Figure 20: *MSE* behavior of *MLE* and Bayes Estimator *n* = 75, *t* = 15
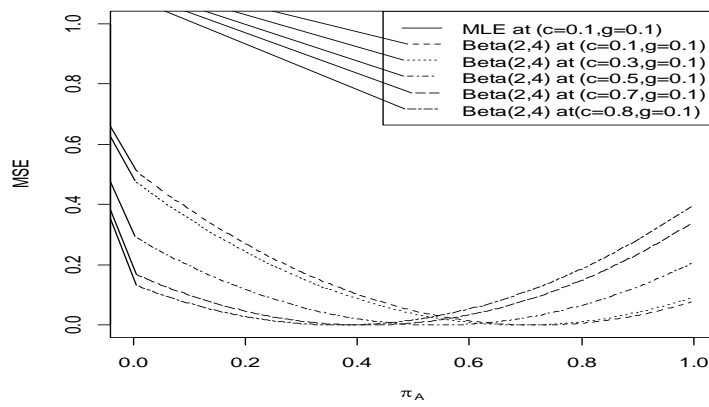
Figure 21: *MSE* behavior of *MLE* and Bayes Estimator *n* = 75, *t* = 30
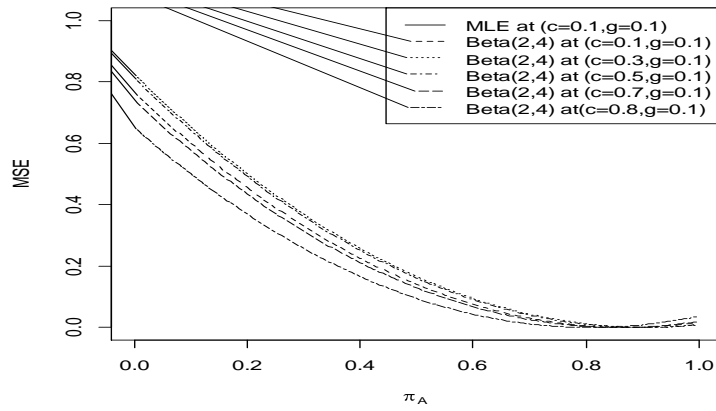
Figure 22: *MSE* behavior of *MLE* and Bayes Estimator *n* = 75, *t* = 60
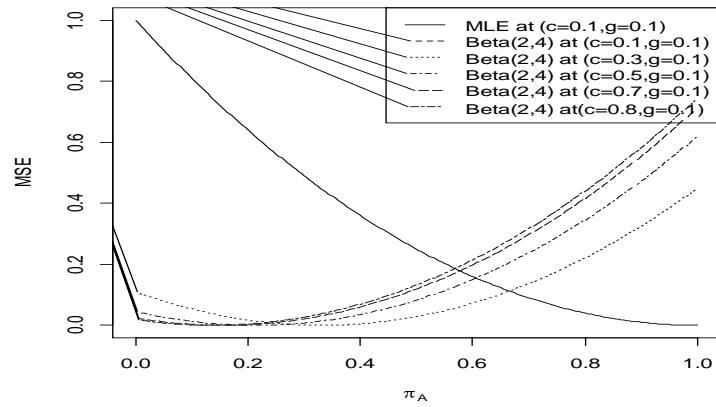
Figure 23: *MSE* behavior of *MLE* and Bayes Estimator *n* = 150, *t* = 30

Figure 24: *MSE* behavior of *MLE* and Bayes Estimator *n* = 150, *t* = 30

Figure 25: *MSE* behavior of *MLE* and Bayes Estimator *n* = 150, *t* = 60

Figure 26: *MSE* behavior of *MLE* and Bayes Estimator *n* = 15, *t* = 3

Figure 27: *MSE* behavior of *MLE* and Bayes Estimator *n* = 15, *t* = 3

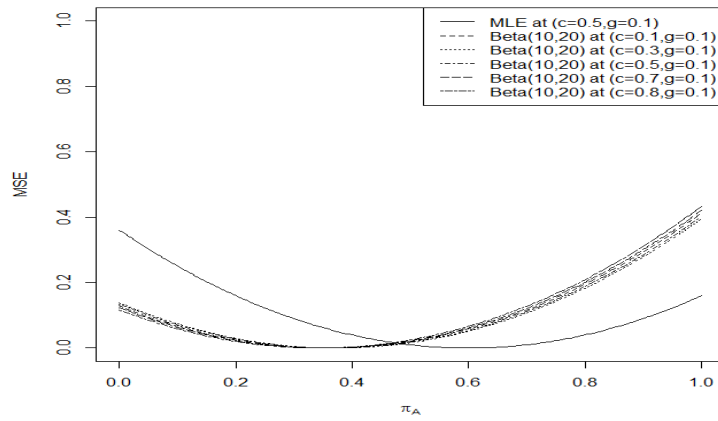Figure 28: *MSE* behavior of *MLE* and Bayes Estimator *n* = 15, *t* = 3

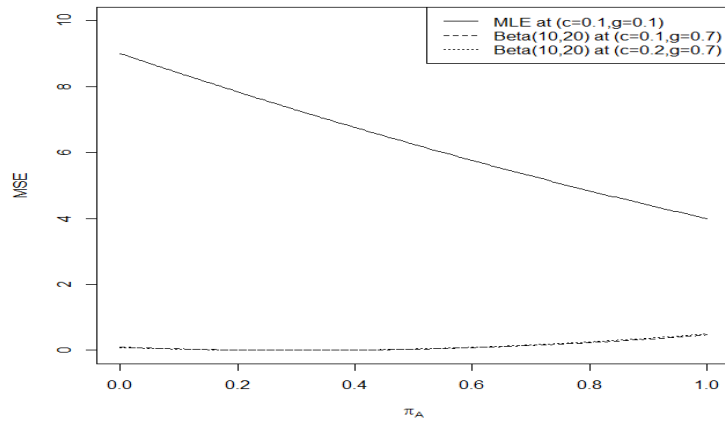Figure 29: *MSE* behavior of *MLE* and Bayes Estimator *n* = 15, *t* = 6

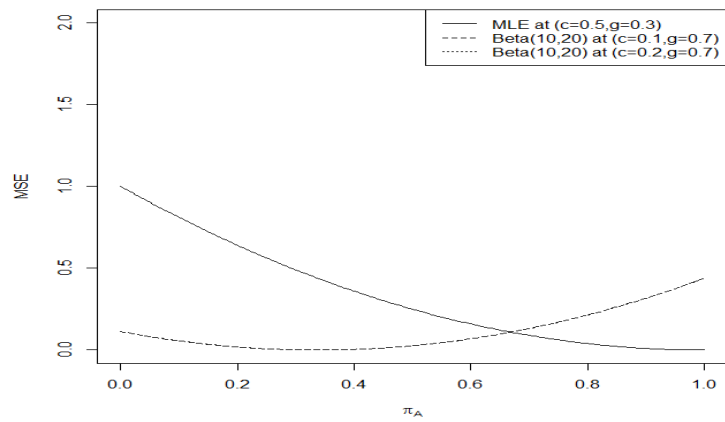Figure 30: *MSE* behavior of *MLE* and Bayes Estimator $n = 15$, $t = 6$

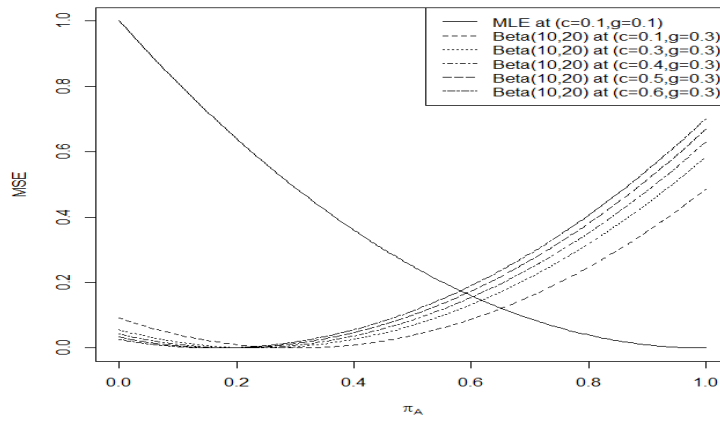Figure 31: *MSE* behavior of *MLE* and Bayes Estimator $n = 15$, $t = 6$

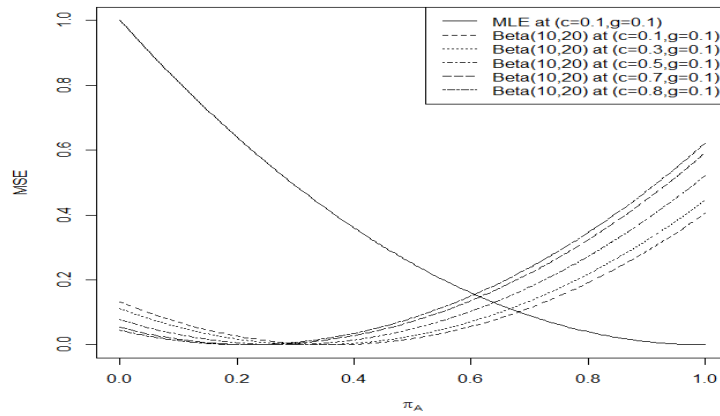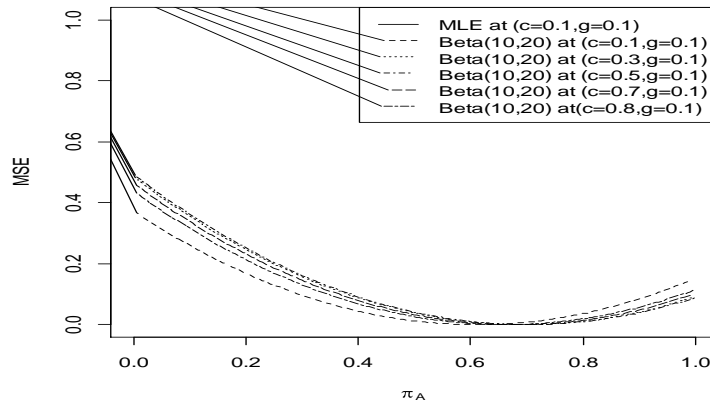Figure 32: *MSE* behavior of *MLE* and Bayes Estimator *n* = 75, *t* = 15
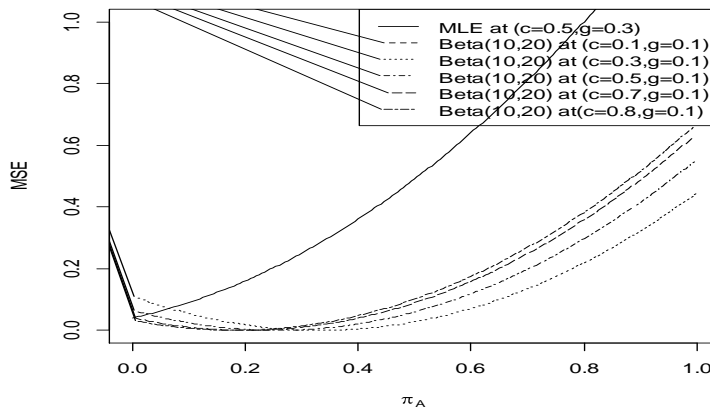

Figure 33: *MSE* behavior of *MLE* and Bayes Estimator *n* = 75, *t* = 30

Figure 34: *MSE* behavior of *MLE* and Bayes Estimator *n* = 75, *t* = 60

Figure 35: *MSE* behavior of *MLE* and Bayes Estimator *n* = 150, *t* = 30

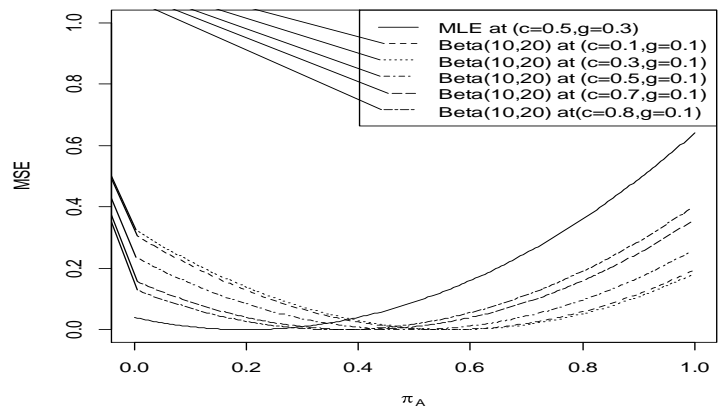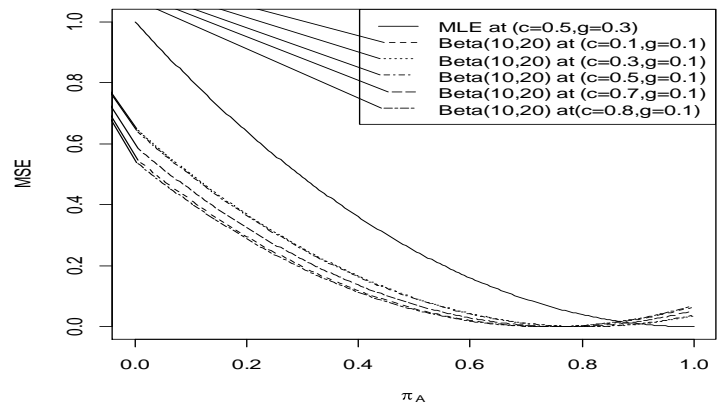Figure 36: *MSE* behavior of *MLE* and Bayes Estimator *n* = 150, *t* = 60

Figure 37: *MSE* behavior of *MLE* and Bayes Estimator *n* = 150, *t* = 120
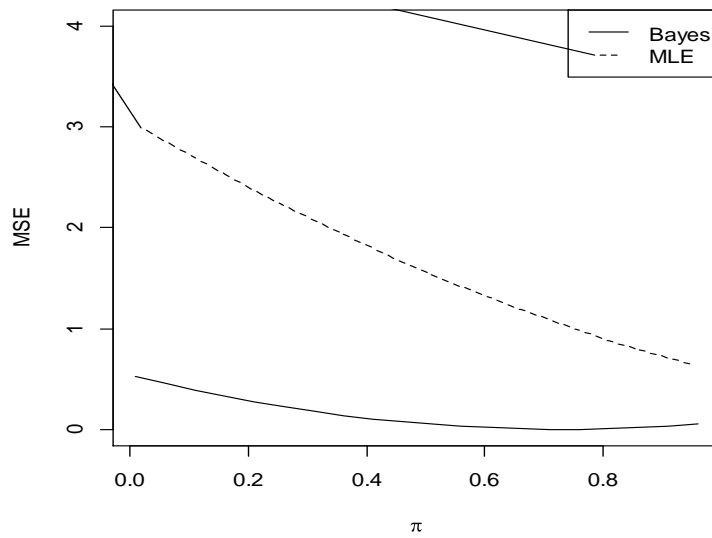
Figure 38: *MSE* behavior of *MLE* and Bayes Estimator for $g = 0.3$, $c = 0.4$, $a_1 = 20, b_1 = 10$, $a_2 = 40, b_2 = 200$, $n = 15$.
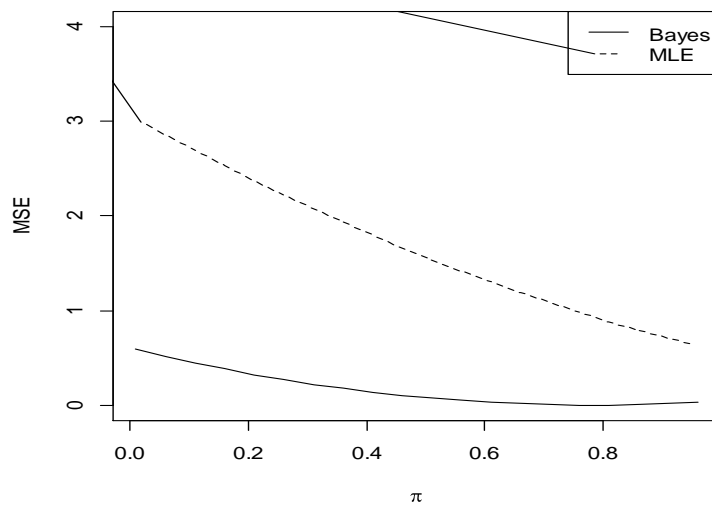

Figure 39: *MSE* behavior of *MLE* and Bayes Estimator for $g = 0.3$, $c = 0.4$, $a_1 = 20, b_1 = 10$, $a_2 = 40, b_2 = 200$, $n = 30$.

It is to be noted that in case of single prior, the posterior distribution is easily computed in studying the performance of the suggested estimation method. For example, if we set $a = 2,\ b = 4$, and $a = 10,\ b = 20$, the plots of the posterior distributions are given in Figures 1-12. Under these two priors the respective Bayes estimates are 0.3467 and 0.3365 when we fix $g = 0.3$, $c = 0.1$, n =15, and observe $t = 6$.

4. Analysis of Real Data

To see the performance of the proposed Bayesian method on real data, we consider the data collected by Liu and Chow [24] in their study to estimate the incidence of induced abortion in Taiwan. Liu and Chow [24] used the Warner [27] model with probability of sensitive question equal to 0.7 (i.e. $c = 0.4$ and $g = 0.3$) on 150 married women. The number of *yes* responses in the sample were recorded as 60 yielding the sample proportion of *yes* responses ($\hat{\delta}$) as 0.4. The maximum likelihood estimate obtained by them was 0.25. Later on, same data were analyzed by Winkler and Franklin [26] and Bar-Lev et al. [7] through the Bayesian analysis using complete and truncated prior distributions, respectively. The Bayes estimates (*Mean*), posterior standard deviation (*SD*) and 95% credible intervals (*CI*) computed by Winkler and Franklin [26], Bar-Lev et al. [7] and Barabesi and Marcheselli [16] are given in Table 7.

We assume SRSWR and simple Beta prior in this analysis. Here, we have $c = 0.4$ and $g = 0.3$ and the results are compatible with the findings of Bar-Lev et al. [7] and Barabesi and Marcheselli [16]. For other values of c, g, n and t, these values are given in Tables 1-4. Further, using (3.1) and (3.2), we have compared the suggested Bayes estimator with the *MLE* and the *MSE*-behavior of these two estimators is given in Figures 13-39. For this purpose, we have used R software to run the codes. For comparison purposes, we have used the same data collected by Liu and Chow [24] and two different priors with choices of hyperparameters as $a = 2, b = 4$ and $a = 10, b = 20$.

Table 7: Description of posterior distribution for $c = 0.4$, $g = 0.3$, $n = 150$, $t = 60$

Prior	Winkler and Franklin			Bar-Lev et al (2003)			Barabesi and Marcheselli (2010)		
(a, b)	Mean	SD	95 % CI	Mean	SD	95 % CI	Mean	SD	95 % CI
1,1	0.25	0.10	0.06,0.44	0.25	0.10	0.06,0.44	0.20	0.10	0.00,0.37
2,4	0.26	0.09	0.09,0.43	0.24	0.10	0.05,0.43	0.21	0.09	0.05,0.38
10,20	0.30	0.06	0.16,0.41	0.25	0.09	0.08,0.43	0.28	0.06	0.16,0.40
2,8	0.22	0.08	0.06,0.38	0.22	0.09	0.03,0.39	0.18	0.08	0.03,0.33
10,40	0.21	0.05	0.12,0.31	0.22	0.09	0.05,0.39	0.20	0.05	0.10,0.29

In order to have a clearer picture of the suggested estimators, we have also compared the two estimators assuming other sample sizes by keeping the proportion of *yes* responses fixed at 0.4. It was observed that increasing the sample size does not have significant effect on the relative efficiency of the Bayes estimator, and for given design constants, relative efficiency of the Bayes estimator decreases as the true proportion of sensitive characteristic increases. Also, if a prior distribution with larger variance is assumed, the *MSE* of the Bayes estimator is decreased.

5. Conclusions

An essential conclusion of this study is that, in the Bayesian scenario, the two design constants c and g be chosen such that their difference is large. For a given prior, data and g, the posterior variance as well as the length of credible interval decrease as c increases. The posterior variance increases with the decrease in prior variance. Also, the posterior variance decreases with the increase in sample size. Thus, while developing a *RRT*, statements must be presented to the respondents such that the design constants c and g are away from one another. For example, setting the probability of choosing sensitive question equal to 0.1 gives c and g away from each other for Warner (1965) RRT. In case of the Bayesian estimation, the prior variance should be large in order to have precise estimators and shorter credible intervals. Using mixture prior is not much helpful. As stated above, the Bayes estimator is more efficient in smaller samples; its application in stratification is fruitful in obtaining the precise estimates. On the whole, we may conclude that in conducting a social survey on sensitive traits through the *RRT*, if prior information is available, it must be exploited through the Bayesian estimation method in order to have precise estimates.

References

[1] Chaudhuri, A. and Mukerjee, R. (1998). *Randomized response: Theory and Methods*, Marcel- Decker, New York.

[2] Lazarus,A. and Stern, R. (1986). Psychiatric Aspects of Pregnancy Termination", *Clinical Obesity and Gynecology* **13**,125-34.

[3] O'Hagan, A. (1987). Bayes linear estimators for randomized response models. *Journal of the American Statistical Association* **82**, 580-85.

[4] Greenberg, B., Abul-Ela, A., Simmons, W. and Horvitz, D. (1969). The unrelated question randomized response: theoretical framework. Journal of the American Statistical Association **64**, 529-39.

[5] Tracy, D. and Mangat, N. (1996). Some development in randomized response sampling during the last decade-a follow up of review by Chaudhuri and Mukerjee. *Journal of Applied Statistical Science* **4**, 533-44.

[6] Jones, E. F. and Forrest, J. D. (1992). Underreporting of Abortion in Surveys of U.S. Women: 1976 to 1988, *Demography* **29**(**1**), 113-26.

[7] Bar-Lev, E. S., Bobovich, K. and Boukai, B. (2003). A common conjugate prior structure for several randomized response models. *Test* **12**(**1**), 101-13.

[8] Pitz, G. (1980). Bayesian analysis of randomized response models. J*ournal of Psychological Bulletin* **87**, 209-12.

[9] Migon, H. and Tachibana, V. (1997). Bayesian approximations in randomized response models. *Computational Statistics and Data Analysis* **24**, 401-09.

[10] Fox, J. and Tracy, P. (1986). *Randomized response: theory and techniques*, Marcel-Dekker, New York.

[11] Moors, J. J. A. (1971). Optimizing of the unrelated question randomized response model. *Journal of the American Statistical Association* **66**, 627-29.

[12] Song, J. J. and Kim, J. M. (2012). Bayesian Analysis of Randomized Response Sum Score Variables. *Communications in Statistics-Theory and Methods* **41**(**10**), 1875-84.

[13] Spurrier, J. and Padgett, W. (1980). The application of Bayesian techniques in randomized response. *Sociological Methods* **11**, 533-44.

[14] Yu, Jun-Wu., Tian, Guo-Liang. and. Tang, Man-Lai. (2008). Two new models for survey sampling with sensitive characteristic: design and analysis. *Metrika* **67**, 251-63.

[15] Barabesi, L. and Marcheselli, M. (2006). A practical implementation and Bayesian estimation in Franklin's randomized response procedure. *Communications is Statistics-Simulations and Computation* **35**, 365-73.

[16] Barabesi, L. and Marcheselli, M. (2010). Bayesian estimation of proportion and sensitivity level in randomized response procedures. *Metrika* **72**, 75-88.

[17] Kim, M. J., Tebbs, J. M. and An, S. W. (2006). Extension of Mangat's randomized response model. *Journal of Statistical Planning and Inference* **36**(**4**), 1554-67.

[18] Oh, M. (1994). Bayesian analysis of randomized response models: a Gibbs sampling approach. *Journal of the Korean Statistical Society* **23**, 463-82.

[19] Adler, N. E. (1976). Sample Attrition in Studies of Psychosocial Sequelae of Abortion: How Great A Problem? *Journal of Applied Social Psychology* **6**(**3**), 240-59.

[20] Mangat, N. S. (1994). An improved randomized response strategy. *Journal of Royal Statistical Society B* **56**(**1**), 93-95.

[21] Unnikrishnan, N. and Kunte, S. (1999). Bayesian analysis for randomized response models. *Sankhya* B **61**, 422-32.

[22] Odumade, O. and Singh, S. (2009). Improved Bar-Lev, Bobovich, and Boukai randomized response Models. *Communications in Statistics-Simulations and Computation* **38**(**3**), 473-75.

[23] Perri, P. F. (2008). Modified randomized devices for Simmons' model. *Model Assisted Statistics and Applications* **3**, 233-39.

[24] Liu, P. T. and Chow, L. P. (1976). The efficiency of the multiple trial randomized response technique. *Biometrics* **32**, 607-18.

[25] Folsom, R. E., Greenberg, B. G., Horvitz, D. G. and Abernathy, J. R. (1973). The two alternate question randomized response model for human surveys. *Journal of the American Statistical Association* **68**, 525-30.

[26] Winkler, R. and Franklin, L. (1979). Warner's randomized response model: A Bayesian approach. *Journal of the American Statistical Association* **74**, 207-14.

[27] Warner, S. L. (1965). Randomized response: a survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association* **60**, 63-69.

[28] Singh, S. and Horn, S. (1998). Estimation of stigmatized characteristics of a hidden gang in finite population. *Australian and New Zealand Journal of Statistics* **40**(**3**), 291-97.

[29] Christofides, T. C. (2003). A generalized randomized response technique. *Metrika* **57**, 195-00.

[30] Chaubey, Y. and Li, W. (1995). Comparison between maximum likelihood and Bayes methods of estimation for binomial probability with sample compositing. *Journal of Official Statistics* **11**, 379-90.

[31] Hussain, Z. and Shabbir, J. (2009a). Improved estimation procedures for the mean of sensitive variable using randomized response model. *Pakistan Journal of Statistics* **25**(**2**), 205-20.

[32] Hussain, Z. and Shabbir, J. (2009b). Bayesian estimation of population proportion of a sensitive characteristic using simple Beta prior. *Pakistan Journal of Statistics* **25**(**1**), 27-35.

[33] Hussain, Z. and Shabbir, J. (2009c). Bayesian Estimation of population proportion in Kim and Warde (2005) Mixed Randomized Response using Mixed Prior Distribution. *Journal of Probability and Statistical Science* **7**(**1**), 71-80.

[34] Hussain, Z. and Shabbir, J. (2007). Randomized use of Warner's randomized response model. *InterStat*: April # 7.

[35] Hussain, Z. and Shabbir, J. and Riaz, M. (2011). Bayesian estimation using Warner's randomized response model through simple and mixture prior distributions. *Communications in Statistics-Simulations and Computations* **40**(**1**), 147-64.

Zawar Hussain
Department of Statistics
Quaid-i-Azam University
45320, Islamabad 44000, Pakistan.
zhlangah@yahoo.com

Muhammad Abid
Department of Statistics
Quaid-i-Azam University
45320, Islamabad 44000, Pakistan.

Javid Shabbir
Department of Statistics
Quaid-i-Azam University
45320, Islamabad 44000, Pakistan.
jsqau@yahoo.com

Nasir Abbas
Department of Statistics
Government Postgraduate College Jhang, Pakistan.
zhlangah@yahoo.com