Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.4, No.4, 2014 www.iiste.org

On The Comparative Analysis Of Beta And Kumaraswamy Priors Using Stigmatized Attributes

Tijani, I. A. and Adewara A. A.
Department of Statistics, University of Ilorin, Ilorin, Nigeria
Email: ismailtijani2012@gmail.com and aaadewara@gmail.com

ABSTRACT
Hussain et al 2011 compared bayes estimator of population proportion of a stigmatized attribute with the Maximum Likelihood Estimator (MLE) but this paper compares two priors: beta and kumaraswamy in the Bayesian analysis using various values of n, x, p1 and p2 and it was observed that across the sample sizes considered over different values of x's, p1 and p2 using R package software for simulation, the mean square error and bias of the Kumaraswamy prior are smaller than that of Beta prior when the proportion of stigmatized attribute increases from 0.1 to 0.4
Keyword: Prior, Beta, kumaraswamy, Stigmatized, Attribute

INTRODUCTION

Randomized response is a research method used in structured survey interview. It was first proposed by Warner in 1965 and later modified by Greenberg in 1969. It allows respondents to respond to sensitive issues (such as criminal behavior or sexuality) while maintaining confidentiality. Chance decides, unknown to the interviewer, whether the question is to be answered truthfully, or "yes", regardless of the truth (Source: Hussain et al, 2011).
For example, social scientists have used it to ask people whether they use drugs, whether they have illegally installed telephones, or whether they have evaded paying taxes. Before abortions were legal, social scientists used the method to ask women whether they had had abortions.
Ask a man whether he had sex with a prostitute this month. Before he answers ask him to flip a coin. Instruct him to answer "yes" if the coin comes up tails, and truthfully, if it comes up heads. Only he knows whether his answer reflects the toss of the coin or his true experience
(Source: Hussain et al, 2011).
Half the people-or half the questionnaire population—who have not had sex with a prostitute get tails and the other half get heads when they flip the coin. Therefore, half of those who have not had sex with a prostitute will answer "yes" even though they have not done it. So whatever proportion of the group said "no", the true number who did not have sex with a prostitute is double that. For example, if 20% of the population surveyed said "no", then the true fraction that did not have sex with a prostitute is 40% (Source: Hussain et al, 2011).
Randomized Response (RR) techniques were developed for the purpose of protecting surveyees privacy and avoiding answer bias mainly. They were introduced by Warner (1965) as a technique to estimate the percentage of people in a population *U* that has a stigmatizing attribute *A.* In such cases respondents may decide not to reply at all or to incorrectly answer. The usual problem faced by researchers is to encourage participants to respond, and then to provide truthful response in surveys. The RR technique was designed to reduce both response bias and non-response bias, in surveys which ask sensitive questions. It uses probability theory to protect the privacy of an individual's response, and has been used successfully in several sensitive research areas, such as abortion, drugs and assault. The basic idea of RR is to scramble the data in such a way that the real status of the respondent cannot be identified (Source:Hussain et al,2011).
Many researchers like Warner (1965), Greenberg (1969), Hussain et al (2011), Hussain and Shabbir (2009), Abid, M, et al (2011), Hussain, Z. et al (2011a, 2011b) etc had contributed in no small measure in this research area.

AIM AND OBJECTIVES

The main purpose of this work is to compare the competence of two priors (beta and kumaraswamy) in the estimation of stigmatized attribute.

ESTIMATION OF POPULATION PROPORTION OF A STIGMATIZED CHARACTERISTIC USING:

(i). Beta Prior

Hussain and Shabbir (2009) developed the Bayes estimator of the population proportion of a sensitive characteristic when data are obtained through the randomized response technique (RRT) proposed by Same in 2007. As far Bayesian estimation based on the data gathered by this technique and assuming a Beta prior distribution with parameters a and b inorder to estimate the parameters. The prior distribution of π is given by

$$f(\pi)=\frac{1}{\beta(a,b)}\pi^{a-1}(1-\pi)^{b-1}, \qquad 0<\pi<1 \quad , \qquad a,b > o \qquad (1)$$

The bias of the estimator of π is B(π)=E(π)-π (2)

and

$$MSE(\pi)==\frac{\sum_{i=0}^{x}\sum_{j=0}^{n-x}\frac{(n-x)!}{j!(n-x-j)!}\frac{x!}{(x-i)!i!}f^{x-i}h^{n-x-j}B(a+i+2,b+j)}{\sum_{i=0}^{x}\sum_{j=0}^{n-x}\frac{(n-x)!}{j!(n-x-j)!}\frac{x!}{(x-i)!i!}f^{x-i}h^{n-x-j}B(a+i,b+j)}-E\left(\pi/x\right)^2+B\left(\left(\pi/x\right)^2\right.$$

(3)

(ii). Kumaraswamy prior distribution

In this section, we shall go for Bayesian estimation based on the data gathering by this technique, we assume a Kumaraswamy prior distribution with parameters a and b, for the parameter to be estimated. That is prior distribution of π is given by;

$$f(\pi)=ab\pi^{a-1}\left(1-\pi^{a}\right)^{b-1}, 0<\pi<1 and\ a,b>0 \qquad (4)$$

The bias is given as:

$$\beta(\pi)=E[\pi/x]-\pi$$
$$=\beta-\pi \qquad (5)$$

and

$$MSE(\pi)=\frac{\sum_{i=0}^{x}\sum_{j=0}^{n-x}\binom{n-x}{j}\binom{x}{j}f^{x-i}h^{n-x-j}\beta(i+3,b+j)}{\sum_{i=0}^{x}\sum_{j=0}^{n-x}\binom{n-x}{j}\binom{x}{j}f^{x-i}h^{n-x-j}\beta(i+1,b+j)}$$

$$-2\pi\frac{\sum_{i=0}^{x}\sum_{j=0}^{n-x}\binom{n-x}{j}\binom{x}{j}f^{x-i}h^{n-x-j}\beta(i+2,b+j)}{\sum_{i=0}^{x}\sum_{j=0}^{n-x}\binom{n-x}{j}\binom{x}{j}f^{x-i}h^{n-x-j}\beta(i+1,b+j)+\pi^{2}} \qquad (6)$$

. **RESULTS**

In this section, attempts will be made to obtain both the mean square error and bias of the proportion of the stigmatized attribute at various values of sample sizes(n), x, p1 and p2.

(n:25,50,75,100,150. x=5,15,20,30,35. p1=0.6 and p2=0.4). The statistical software R will be used in implementing our results through simulation as shown in the tables and graphs below with varied values of π**:**

Estimation of Mean Square Error and Bias at various values of n, x, p1 and p2 with varied values of π:

Table 1: n=25, x=5, p1=0.6, p2=0.4

MSE			BIAS	
Π	BETA PRIOR	KUMA PRIOR	BETA PRIOR	KUMA PRIOR
0.1	8.706003E-10	1.46357E-10	0.51116728	0.209585176
0.2	1.018681E-9	7.236103E-11	0.41116728	0.109585176
0.3	1.023107E-9	9.708084E-13	0.31116728	0.009585176
0.4	8.024361E-9	1.471081E-10	0.21116728	0.090414
0.5	3.68302E-10	1.080567E-10	0.11116728	0.190414824
0.6	5.992426E-12	4.052708E-9	0.01116728	0.290414824
0.7	5.957948E-10	1.150809E-8	-0.08883272	0.390414824
0.8	4.125575E-9	2.782641E-8	-0.18883272	0.490414824
0.9	1.44352E-8	6.031767E-8	-0.28883272	0.590414824

Table 2: n=50, x=15, p1=0.6, p2=0.4

MSE			**BIAS**	
Π	BETA PRIOR	KUMA PRIOR	BETA PRIOR	KUMA PRIOR
0.1	3.024865E-12	4.185519E-13	0.52121647	0.193883039
0.2	2.66351E-12	1.323177E-13	0.42121647	0.093883039
0.3	2.007143E-12	7.278705E-16	0.32121647	0.006116961
0.4	1.186721E-12	2.730759E-13	0.22121647	0.106116961
0.5	4.276352E-13	1.236453E-12	0.12121647	0.206116961
0.6	1.51441E-14	3.152623E-12	0.02121647	0.306116961
0.7	6.179558E-12	2.325552E-13	0.07878353	0.406116961
0.8	1.284917E-12	1.029726E-11	0.17878353	0.506116961
0.9	3.228616E-12	1.526142E-11	0.27878353	0.606116961

Table 3: n=75, x=20, p1=0.6, p2=0.4

MSE			**BIAS**	
Π	BETA PRIOR	KUMA PRIOR	BETA PRIOR	KUMA PRIOR
0.1	1.844584E-13	1.85642E-14	0.53532821	0.0698279
0.2	1.420033E-13	3.653616E-15	0.43532821	0.0698279
0.3	9.385285E-14	7.59334E-16	0.33532821	0.0301721
0.4	4.929921E-14	1.508439E-14	0.23532821	0.1301721
0.5	1.665623E-14	4.81843E-14	0.13532821	0.2301721
0.6	1.111054E-15	9.70449E-14	0.03532821	0.3301721
0.7	1.544514E-13	3.490898E-15	0.06467179	0.4301721
0.8	2.031905E-14	2.106195E-13	0.16467179	0.5301721
0.9	4.50893E-14	2.556092E-13	0.26467179	0.6301721

Table 4: n=100, x=30, p1=0.6, p2=0.4

MSE			BIAS	
Π	BETA PRIOR	KUMA PRIOR	BETA PRIOR	KUMA PRIOR
0.1	6.432135E-16	4.943282E-17	0.54683963	0.15159682
0.2	3.727441E-16	4.969964E-18	0.44683963	0.05159682
0.3	1.848477E-16	3.600016E-18	0.34683963	0.04840318
0.4	7.309341E-16	2.642013E-17	0.24683963	0.14840318
0.5	1.915079E-17	5.480429E-17	0.14683963	0.24840318
0.6	1.36744E-18	7.56564E-17	0.04683963	0.34840318
0.7	8.328159E-17	1.170544E-18	0.05316037	0.44840318
0.8	6.107277E-18	7.829873E-17	0.15316037	0.54840318
0.9	9.903798E-18	6.496817E-17	0.25316037	0.64840318

Table 5: n=150, x=35, p1=0.6, p2=0.4

MSE			BIAS	
Π	BETA PRIOR	KUMA PRIOR	BETA PRIOR	KUMA PRIOR
0.1	4.262126E-17	1.224111E-18	0.58546547	0.0992198286
0.2	2.196066E-17	5.671642E-23	0.48546547	0.0007801714
0.3	9.796078E-18	6.696247E-19	0.38546547	0.1007801714
0.4	3.588673E-18	1.775286E-18	0.28546547	0.2007801714
0.5	9.547208E-19	2.511011E-18	0.18546547	0.3007801714
0.6	1.204641E-19	2.649041E-18	0.08546547	0.4007801714
0.7	2.313972E-18	1.949242E-21	0.01453453	0.5007801714
0.8	6.366088E-20	1.751584E-18	0.11453453	0.6007801714
0.9	1.101998E-19	1.175845E-18	0.21453453	0.7007801714

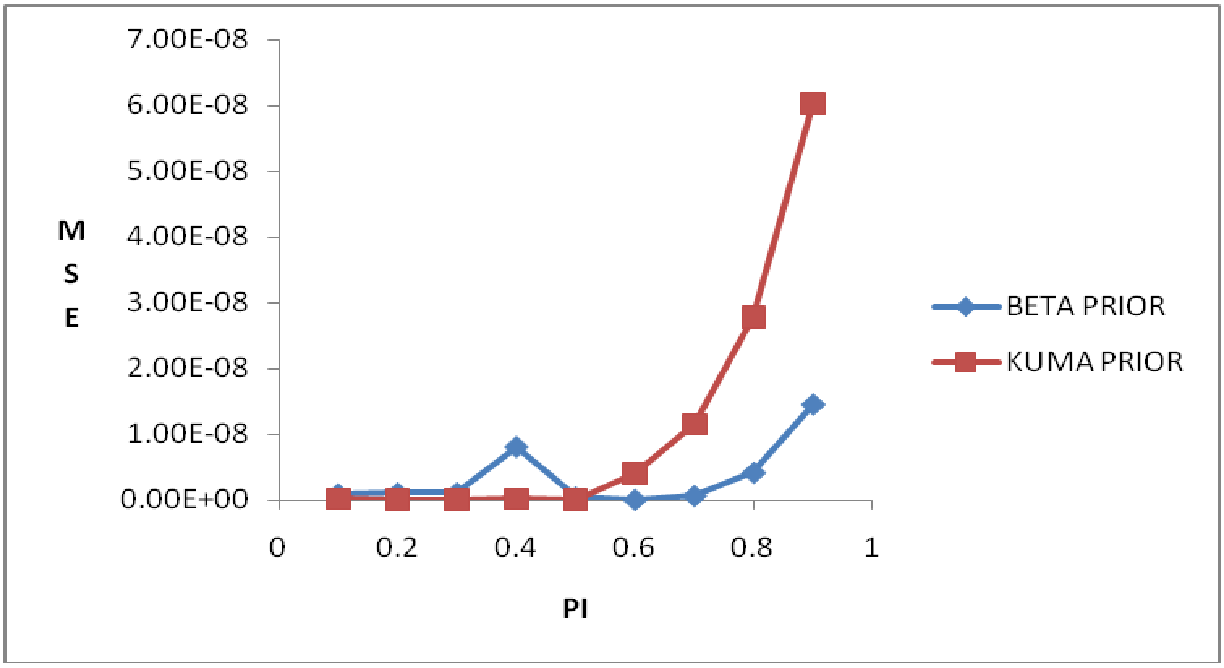

FIG 1a: THE MSE PLOT OF PRIORS WHEN n=25

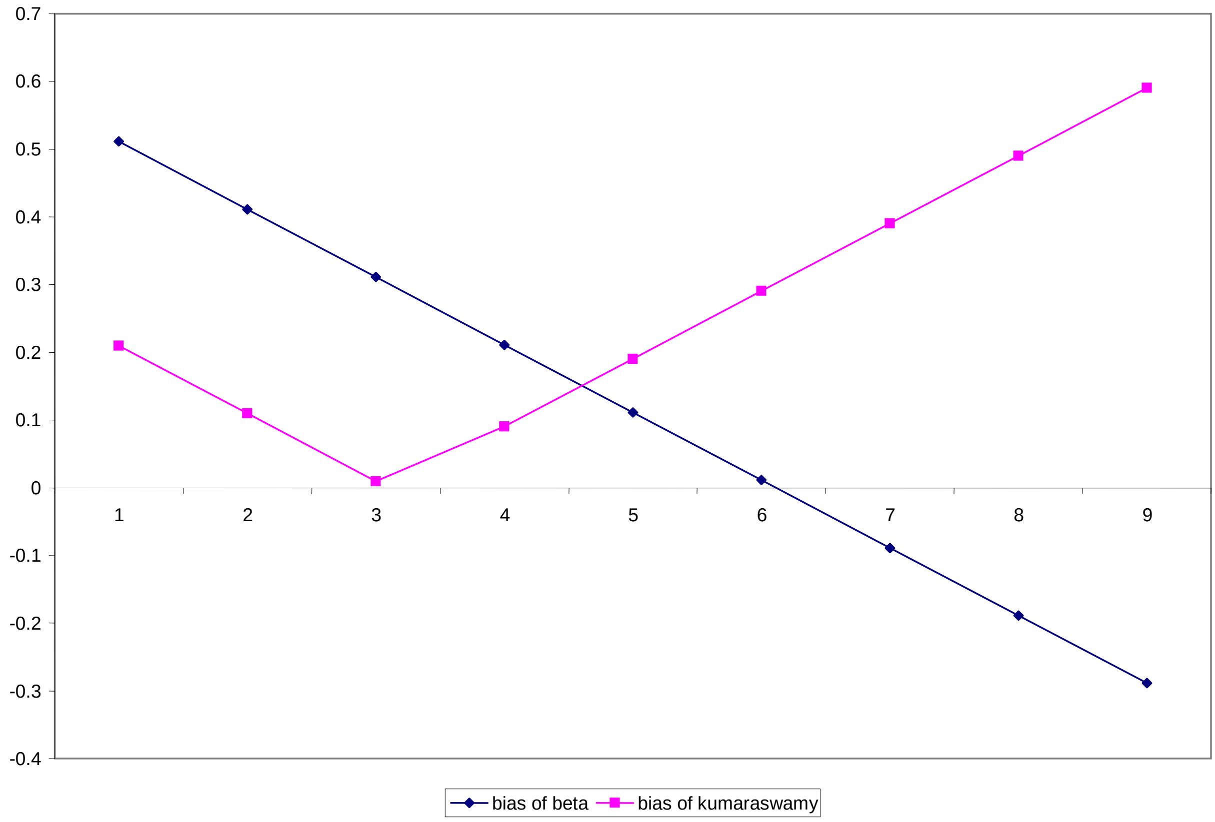

FIG 1b: THE BIAS PLOT OF PRIORS WHEN n=25

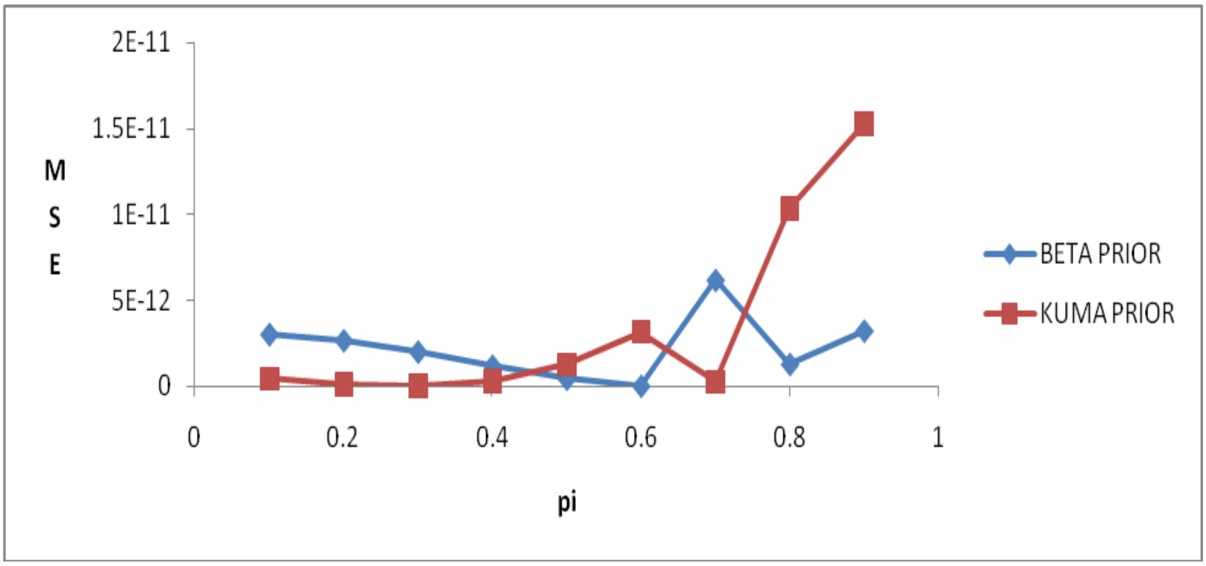

FIG 2a: THE MSE PLOT OF PRIORS WHEN n =50

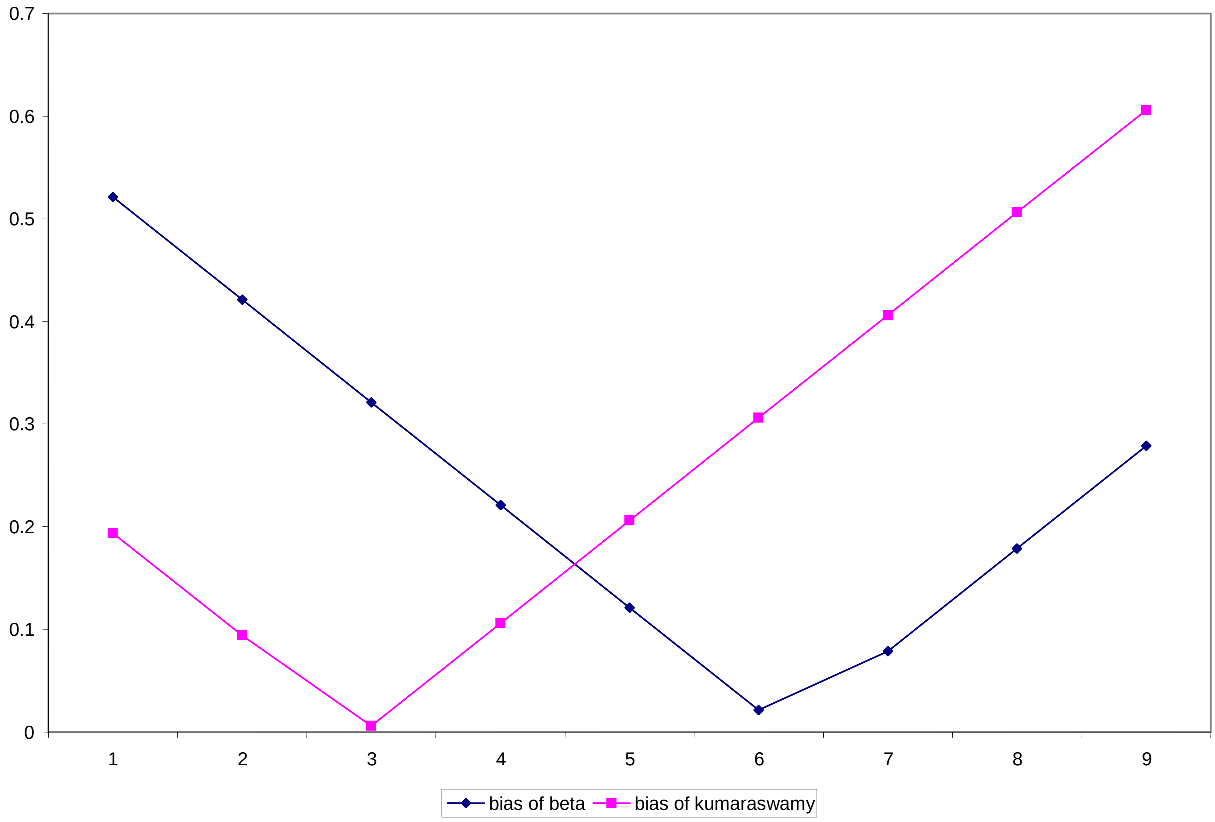

FIG 2b: THE BIAS PLOT OF PRIORS WHEN n=50

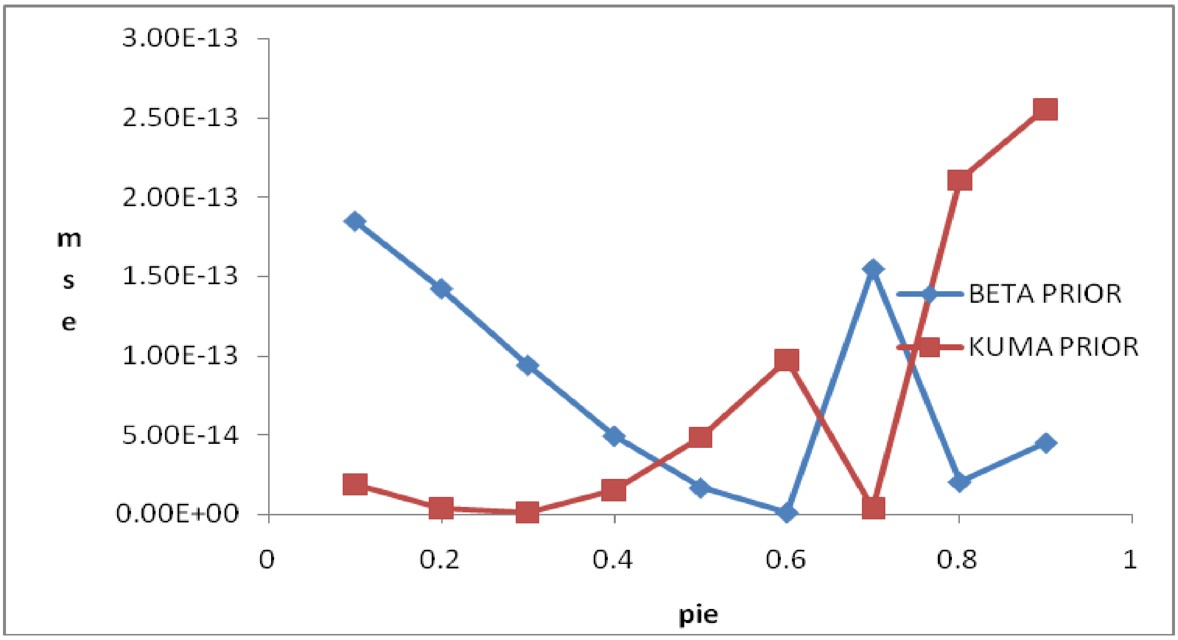

FIG 3a: THE MSE PLOT OF PRIORS WHEN n=75

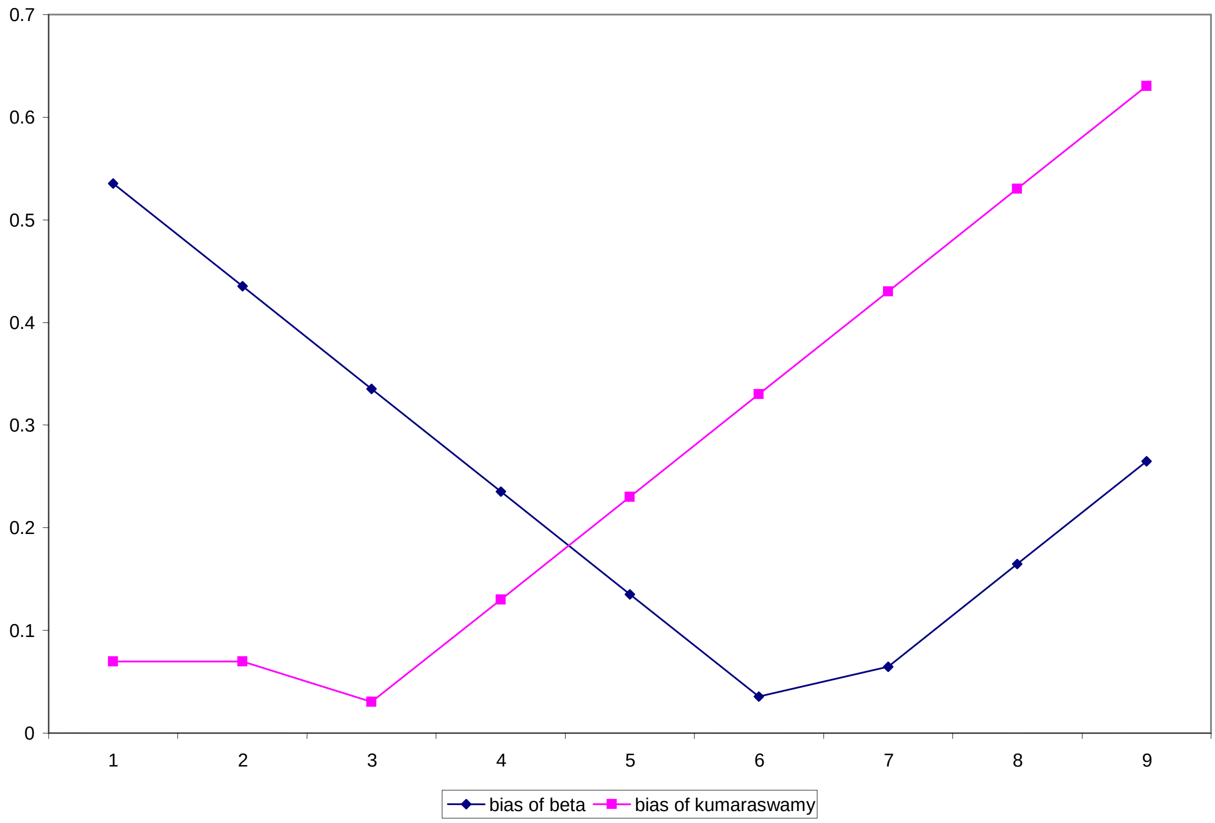

FIG 3b: THE BIAS PLOT OF PRIORS WHEN n=75

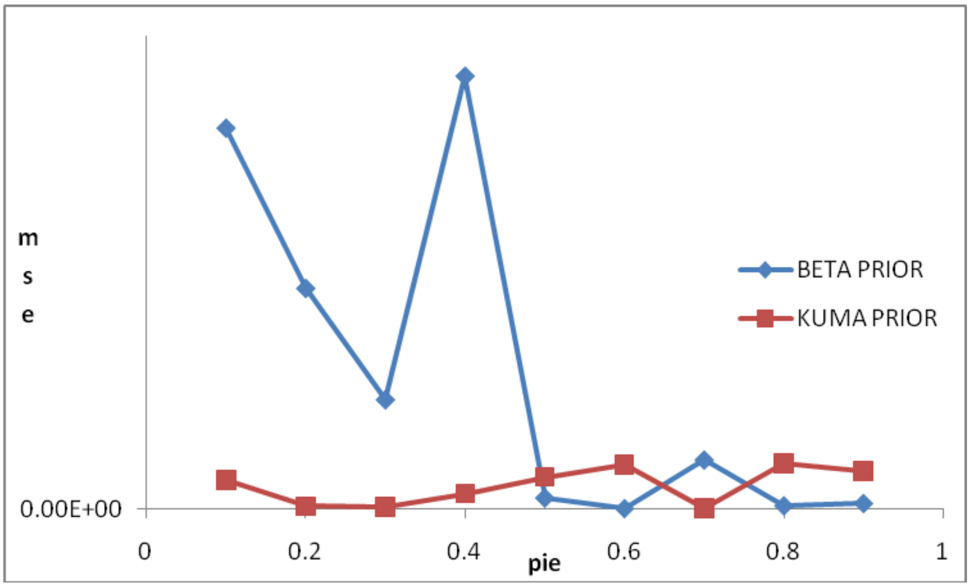

FIG 4a: PLOT OF MSE OF PRIORS WHEN n=100

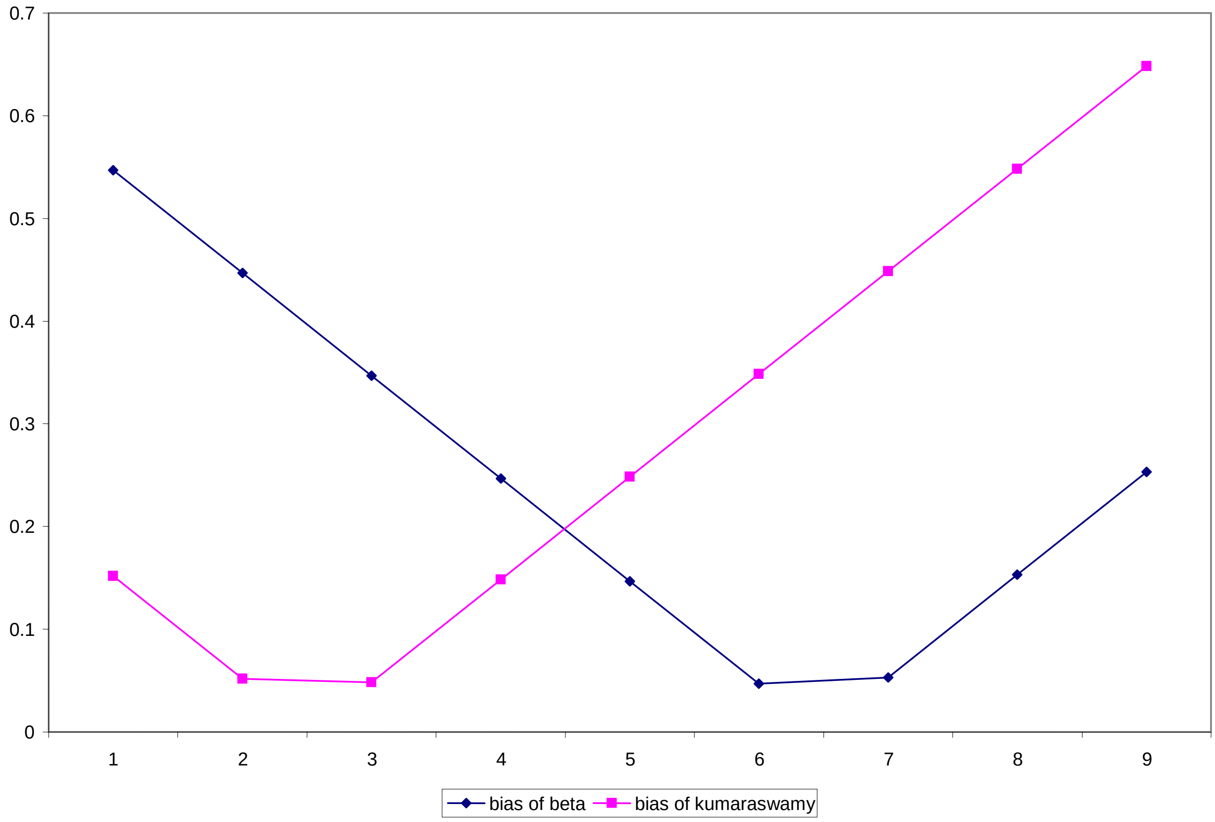

FIG 4b: THE BIAS PLOT OF PRIORS WHEN n=100

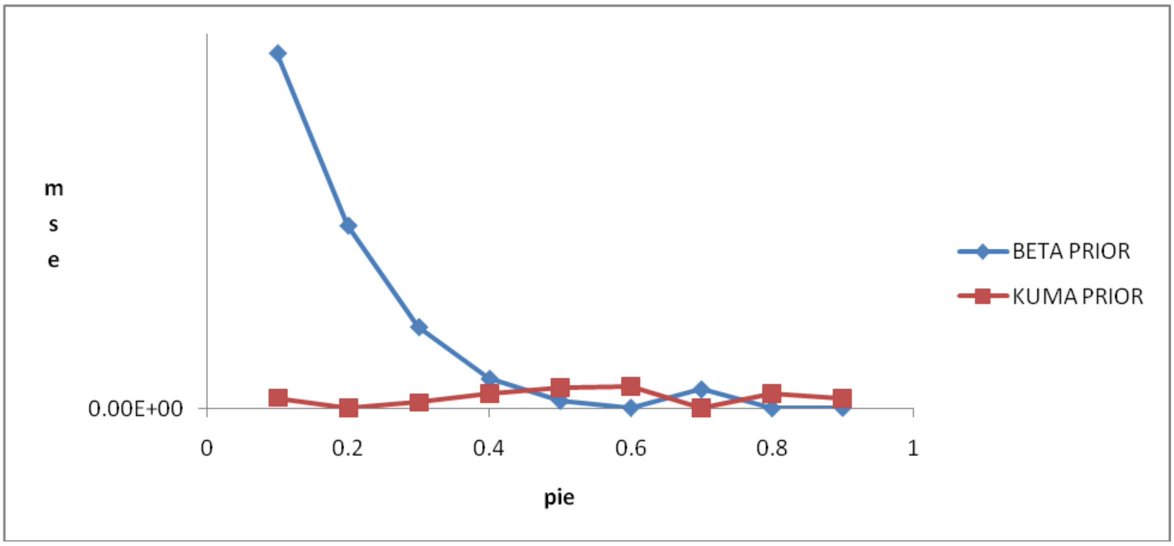

FIG 5a: PLOT OF MSE OF PRIORS WHEN n=150

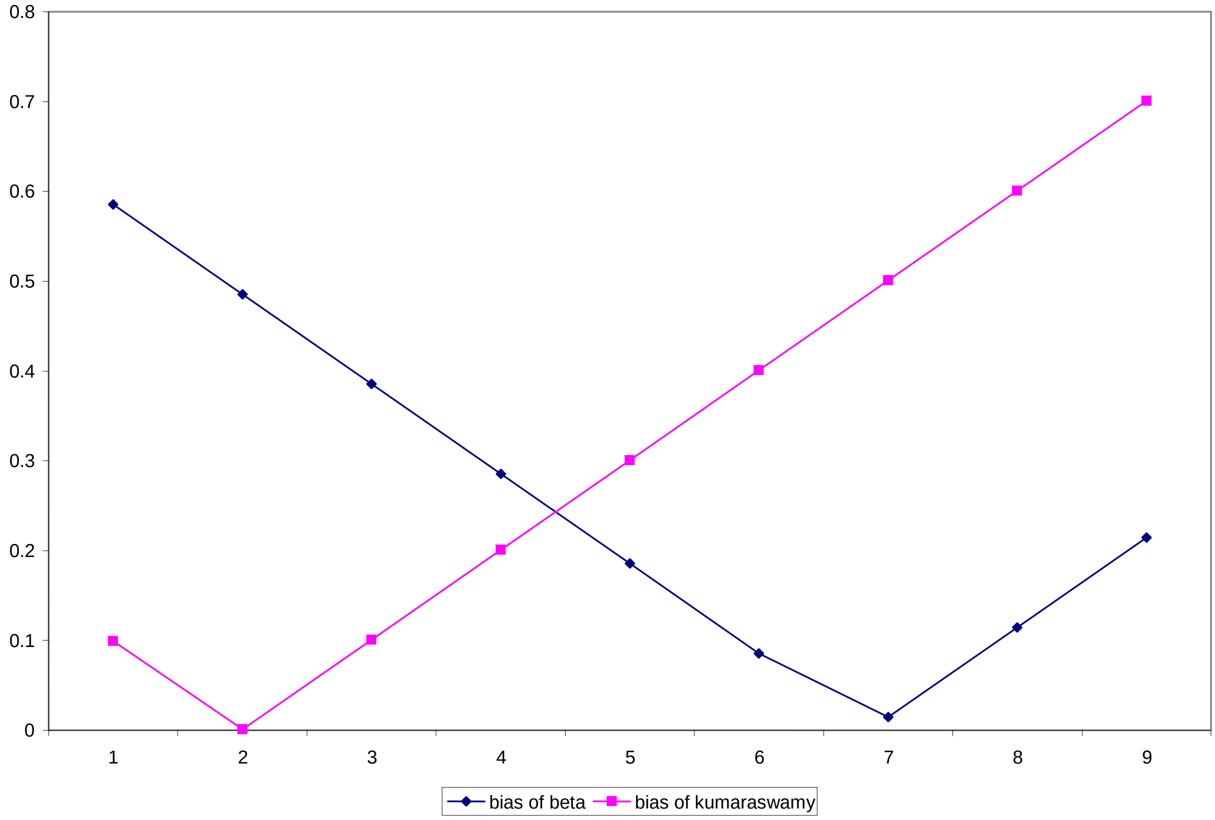

FIG 5b: THE BIAS PLOT OF PRIORS WHEN n=150

CONCLUSION

Having considered the necessary techniques as spelt out in the aim and objectives of the study, there is no doubt that the main intention of this work has been realized. It was observed that across the sample sizes considered over different values of x's, p1 and p2, the mean square error and bias of the Kumaraswamy prior are smaller than that of Beta prior when the proportion of stigmatized attribute increases from 0.1 to 0.4 irrespective of sample sizes, x, p1 and p2 as shown in the tables and graphs above.

REFERENCES

- Abid, M. Cheema, S.A. and Hussain, Z. (2011). An efficiency comparison of certain randomized response strategies. International Journal of business and social sciences, 2 (7), 188 – 198.
- Greenberg, B.G., Abul – Ela Abdel – Latif, A., Simmons, W.R. and Horvitz, D.G. (1969). The unrelated question RR Model: Theoretical framework, Journal of the American Statistical Association, 64, 52 – 539.
- Hussain, Z., Cheema, S.A. and Zafar, S. (2011): Extension of Mangat Randomized Response Model. International Journal of Bus. and Social Sciences: 2(8), 261 - 266
- Hussain, Z., Shabbir, J. (2009). An estimation of sensitive proportion utilizing Higher Order Moments of auxiliary variable. International Journal of Business and Social Sciences, 2(2), 121 – 125.
- Hussain, Z., Hamraz, M. and Shabbir, J. (2011a). An improved quantitative randomized response Model. International Journal of Business and Social Sciences, 2(4), 200 – 205.
- Hussain, Z., Anjum, S. and Shabbir, J. (2011b). improved logit estimation through mangat randomized response model. International journal of business and social sciences, 2(5), 179 – 188.
- Warner, S. L. (1965): Randomized response: a survey technique for eliminating evasive answer bias. Journal of American Statistical Association: 60. 63 – 69.