DIGITALCOMMONS
@WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 13 | Issue 2 Article 23

11-2014

Contrast of Bayesian and Classical Sample Size Determination

Farhana Sadia
University of Dhaka, Dhaka, Bangladesh, fsadia@isrt.ac.bd

Syed S. Hossain
University of Dhaka, Dhaka, Bangladesh, shahadat@isrt.ac.bd

Recommended Citation

Sadia, Farhana and Hossain, Syed S. (2014) "Contrast of Bayesian and Classical Sample Size Determination," *Journal of Modern Applied Statistical Methods*: Vol. 13 : Iss. 2 , Article 23.
DOI: 10.22237/jmasm/1414815720
Available at: http://digitalcommons.wayne.edu/jmasm/vol13/iss2/23

Journal of Modern Applied Statistical Methods
November 2014, Vol. 13, No. 2, 420-431.

ISSN 1538 − 9472

Contrast of Bayesian and Classical Sample Size Determination

Farhana Sadia
University of Dhaka
Dhaka, Bangladesh

Syed S. Hossain
University of Dhaka
Dhaka, Bangladesh

Sample size determination is a prerequisite for statistical surveys. A comprehensive overview of the Bayesian approach for computation of the sample size, and a comparison with classical approaches, is presented. Two surveys are taken as example to illustrate the accuracy and efficiency of each approach, and to make recommendations about which method is preferred. The Bayesian approach of sample size determination may require fewer subjects if proper prior information is available.

Keywords: Sample size determination, Bayesian methods, mean

Introduction

A good statistical study is one that is well designed and leads to a valid conclusion. The main aim of the sample size determination (SSD) is to find an adequate number of observations to be made to estimate the population prevalence with a good precision. That means an optimal sample size is required to give a desired level of validity of the results. Prior determination of a good sample size reduces expenses and time by allowing researchers to estimate information about a whole population without having to survey each member of the population (Cochran, 1977). A considerable number of criteria for SSD are available depending on the two types of inferential approaches-Frequentist and Bayesian.

Frequentist sample size determination methods depend directly on the unknown parameter of interest which in practice is often very hard to get whereas Bayesian way does not depend on the guessed value of the true parameter rather it depends on a prior distribution of the parameter (M'lan et al., 2008). Bayesian methods often results in providing a posterior distribution which combines the pre-experimental information of the parameter (prior distribution) with the

Pr. Sadia is a Lecturer in the Institute of Statistical Research and Training (ISRT). Email him at: fsadia@isrt.ac.bd. Dr. Hossain is Professor in the Institute of Statistical Research and Training (ISRT). Email him at: shahadat@isrt.ac.bd.

experimental data by utilizing the likelihood of the parameter (Pham-Gia, 1995). In case of Bayesian sample size determination, the marginal or prior-predictive distribution is used which is the mixture of the sampling distribution of the data and the prior distribution of the unknown parameters (M'lan et al., 2008). In this context, the minimal sample size determination using three different Bayesian approaches based on highest posterior density (HPD) intervals which are average coverage criterion (ACC), average length criterion (ALC) and worst outcome criterion (WOC) (Joseph et al., 1995) are examined herein. Sample sizes for two real life surveys were calculated using these criteria and were compared with the sample size determined by classical method as well as with the actual sample size utilized in these surveys.

Methodology

Bayesian Methods Used in Sample Size Determination

Let θ be an unknown parameter vector that is derived to be estimated and Θ be the parameter space for θ. Suppose it is desired to determine the sample size n where a random sample $X = (X_1, X_2, \ldots, X_n)$ is to be used for the estimation of θ. According to the Bayesian approach, if $f(\theta)$ is the prior distribution for the parameter and the likelihood function given the data $x = (x_1, x_2, \ldots, x_n)$ is $L(\theta; x) \propto f(x \mid \theta)$. The preposterior marginal distribution of x is thus given by

$$f(x) = \int_{\Theta} f(x \mid \theta) f(\theta) d\theta . \tag{1}$$

Now, the posterior distribution of θ given data x with sample size n is

$$f(\theta \mid x, n) = \frac{f(x \mid \theta) f(\theta)}{\int_{\Theta} f(x \mid \theta) f(\theta) d\theta} . \tag{2}$$

Using HPD interval approach, a sample size n most appropriate for estimating θ can be obtained by finding the n that gives the highest coverage of the equation (2) for a given fixed interval. The following three criteria are used in this paper. For details of these criteria, see (M'lan et al., 2008; Joseph et al., 1995; Joseph et al., 1997; Sahu et al., 2006). The average coverage criterion seeks the smallest n satisfying the following condition

$$\int_{\chi}\left\{\int_{a(x,n)}^{a(x,n)+l} f\left(\theta \mid x,n\right)d\theta\right\} f\left(x\right)dx \geq 1-\alpha\,, \tag{3}$$

where, $f(x)$ and $f(\theta \mid x, n)$ are given in equation (1) and (2) and $a(x, n)$ is the lower limit of the HPD credible set of length l for the posterior density $f(\theta \mid x, n)$. In general, $a(x, n)$ will depend both on the data x and the sample size n. ACC finds the minimum sample size n such that the expected coverage probability is at least $(1 - \alpha)$ for a given fixed HPD interval length l. The average length criterion seeks the smallest n satisfying the condition

$$\int_{\chi} l'\left(x,n\right) f\left(x\right)dx \leq l\,, \tag{4}$$

where l is the desired pre-specified average length. This average length criterion is used to find a sample size n that would fix the coverage probability $(1 - \alpha)$ of the HPD credible set for θ. The worst outcome criterion finds the smallest n satisfying the following condition

$$\inf_{x\in\chi}\left\{\int_{a(x,n)}^{a(x,n)+l} f\left(\theta \mid x,n\right)d\theta\right\} \geq 1-\alpha\,, \tag{5}$$

where both l and α are fixed in advance.

Bayesian Sample Sizes for Normal Mean

Let the data vector $x = (x_1, x_2, \ldots, x_n)$ consist of exchangeable components from a normal distribution with the unknown normal mean μ and variance σ^2. The precision of the data is then $\lambda = \sigma^2$. In this case, the prior distribution is a conjugate prior distribution. The prior distribution for μ and λ are $\lambda \sim gamma(v, \beta)$ and $\mu \mid \lambda \sim N(\mu_0, n_0\,\lambda)$. That means, the conjugate prior distribution for (μ, λ) is the normal-gamma conjugate prior distribution.

Sample Sizes for Single Normal Mean with Known Precision

If the precision λ is known, then the posterior distribution will be a normal distribution, i.e.

$$\mu \mid x \sim N\left(\mu_n, \lambda_n\right), \tag{6}$$

where

$$\lambda_n = (n + n_0)\lambda,$$
$$\mu_n = \frac{n_0\mu_0 + n\overline{x}}{n_0 + n},$$
$$\text{and } \overline{x} = \tfrac{1}{n}\sum_{i=1}^{n} x_i$$

In this case, the three Bayesian sample size criteria give the same solution because the posterior precision depends only on n and does not vary with the particular observed data vector x. This is also equivalent to that given by Adcock (1988) as

$$n \geq \frac{4Z^2_{1-\alpha/2}}{\lambda l^2} - n_0. \tag{7}$$

If a non-informative prior is used such that $n_0 = 0$, then inequality (7) reduces to the classical formulation.

Sample Sizes for Single Normal Mean with Unknown Precision

If the precision λ is unknown, then marginal posterior distribution of λ is given by

$$\mu \mid x \sim \sqrt{\left\{\frac{\beta_k}{(n+n_0)(\upsilon + \frac{n}{2})}\right\}} t_{2\upsilon+n}\mu_n \ , \tag{8}$$

where

$$\beta_n = \beta + \tfrac{1}{2}ns^2 + \tfrac{1}{2}\frac{nn_0}{n+n_0}(\overline{x} - \mu_0)^2,$$

and $t_{2\upsilon+n}$ represents a t-distribution with $2\upsilon + n$ degrees of freedom. In this case, different Bayesian sample size criteria will give different sample size if the posterior precision varies with the data.

The ACC sample size for unknown precision is similar to that for known precision because $\upsilon \mid \beta$ is the prior mean for precision λ, thus it is only necessary to substitute the prior mean precision for λ in inequality (7) and exchange the normal quantile Z with a quantile from a $t_{2\upsilon}$ distribution. If the degrees of freedom of t distribution do not increase with the sample size, equation (7) can give different sample size which is substantially different from those from inequality (8) and classical method of sample size.

The average length criterion seeks the minimum n satisfying the following condition

$$2t_{n+2\upsilon,1-\alpha/2}\sqrt{\left\{\frac{2\beta}{(n+2\upsilon)(n+n_0)}\right\}\frac{\Gamma\left(\frac{n+2\upsilon}{2}\right)\Gamma\left(\frac{2\upsilon-1}{2}\right)}{\Gamma\left(\frac{n+2\upsilon-1}{2}\right)\Gamma(\upsilon)}}\leq l. \tag{9}$$

When estimating a single normal mean with unknown precision λ with a gamma (v, β) prior distribution on λ, the ALC (4) is satisfied for large n. Although it does not appear feasible to solve inequality (9) explicitly for n, the left-hand side is straight forward to calculate given υ, β, α and n. Therefore, the exact smallest n can be found by a bi-sectional search algorithm (Chen et al., 1998).

For a single normal mean with unknown precision λ with a gamma (v, β) prior distribution on λ, the WOC is satisfied when n is sufficiently large so that

$$\frac{l^2(n+2v)(n+n_0)}{8\beta\left\{1+\left(\frac{n}{2v}\right)F_{n,2v,1-w}\right\}}\geq t^2{}_{n+2\upsilon,1-\alpha/2}, \tag{10}$$

where $F_{n,2v,1-w}$ denotes the 100(1 − w) percentile of an F distribution with n and 2v degrees of freedom and $\int_\chi f(x)\,dx = 1-w$. Therefore, the exact smallest n satisfying inequality (10) can be found by a bi-sectional search algorithm. If $X = \chi$, the sample size is not defined, because $F_{n,2v,1-w} \to \infty$ as $w \to 0$, hence inequality (10) cannot be satisfied for any n.

Results

Classical and Bayesian Sample Size for mean with Simple Random Sampling

For simple random sampling, computation of classical sample size for mean is made using the conventional formula (Cochran, 1977)

$$n = \frac{z^2_{\frac{\alpha}{2}} (CV)^2}{r^2}, \tag{11}$$

where, CV is the coefficient of variation and r the relative margin of error. Also note that with the population mean denoted by μ, $d = r\mu$ where d is absolute the margin of error and α is the level of significance. Pre-assigned values of α, r, CV can give an appropriate sample size n.

Bayesian sample sizes are computed using the three types of criteria given in earlier section and these criteria find the minimum sample size n satisfying the respective condition of the criteria. Table 1 gives the classical and Bayesian sample size for mean with $\alpha = 0.05$ considering simple random sampling assuming different prior distributions of mean. Also note that different length of the posterior credible interval for the mean are computed and given in Table 1. To make sense of Bayesian sample size in Table 1, gamma prior distributions for the precision are used with different types of parameters.

In Table 1 the coefficient of variation used in the classical method is CV=2. Table 1 shows that the three Bayesian criteria provides different sample sizes. It is also observed from Table 1 that Bayesian criteria ACC and WOC seem to lead similar sample sizes whereas ALC criteria provides the smallest sample sizes. For example, in case of l = 0.1 and a prior about mean, $u = v = 2$, ACC and WOC yield the sample size of n = 3074 and n = 3638 which are somewhat similar but ALC yields a sample size of n = 2405 which is smaller than that using ACC and WOC. However, from Table 1 the theoretical knowledge that $n_{ALC} \leq n_{ACC} \leq n_{WOC}$ is observed to be satisfied. It is important to note that, as long as non-informative prior approaches to informative prior, the sample size gradually reduces. For example, Bayesian sample sizes are larger than classical sample size for non-informative prior (1, 1) but they are smaller than the classical sample size when using the more informative prior. That means, if more informative prior information is in hand, Bayesian method could supply more parsimonious sample size than classical method would.

Table 1. Classical and Bayesian Sample Size for SRS (α = 0.05)

Length (l)	Classical sample size	Different prior	Bayesian sample size		
			ACC	ALC	WOC
0.1	6147	Gamma (1,1)	7396	4819	7948
		Gamma (2,2)	3074	2405	3638
		Gamma (3,3)	2385	2028	2627
		Gamma (4,4)	2118	1877	2488
		Gamma (2,3)	4526	3612	4962
0.2	1537	Gamma (1,1)	1842	1198	2010
		Gamma (2,2)	761	595	823
		Gamma (3,3)	589	501	675
		Gamma (4,4)	522	463	614
		Gamma (2,3)	1147	463	1250
0.3	683	Gamma (1,1)	813	528	890
		Gamma (2,2)	333	260	392
		Gamma (3,3)	257	218	310
		Gamma (4,4)	227	201	255
		Gamma (2,3)	504	394	560
0.5	246	Gamma (1,1)	287	185	311
		Gamma (2,2)	114	88	136
		Gamma (3,3)	86	73	105
		Gamma (4,4)	76	67	89
		Gamma (2,3)	176	136	190

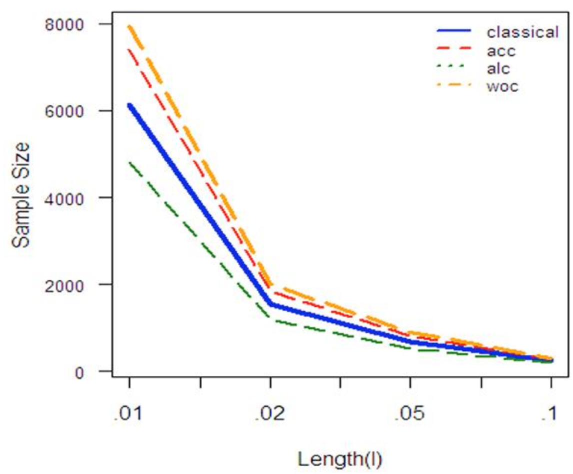

Figure 1. Classical and Bayesian sample size for different length with prior gamma (1,1)

Figure 1 gives the the comparison among classical sample sizes and Bayesian sample sizes for mean with respect to different length of the highest posterior density interval using the non-informative prior (1,1). Figure 1 shows that the Bayesian WOC criteria provides the largest sample size whereas Bayesian ALC criteria provides the smallest sample size and classical sample size and Bayesian ACC criteria give almost similar sample size. Figure 1 also elucidates that as length increases, sample size gradually decreases and this fact is true for both classical and Bayesian method of sample size determination.

Applicability of the Bayesian SSD in Real Life

The sample size determination (SSD) approaches from Bayesian perspective are grounded in theory and are eventually candidates for utilization in some real surveys. However, positive utilization of these methods in large-scale survey research in Bangladesh would depend on the computational features of the methods with respect to those usually used in such surveys. This study considered two recently conducted surveys as examples by comparing the sample sizes in these surveys with the hypothetical appropriate sample size computed using Bayesian criteria. The choice of the surveys was arbitrary; samples were selected mainly by considering availability in published format.

Household Income and Expenditure Survey 2010

Household Income and Expenditure Survey (HIES) is conducted by Bangladesh Bureau of Statistics (BBS), and is the main data source for estimation of poverty in Bangladesh. This survey provides valuable data on household income, expenditure, consumption, savings, housing condition, education, employment, health and sanitation, water supply and electricity, etc. (HIES, 2005). In the 2010 survey, a two-stage stratified random sampling technique was followed in drawing samples. The sample size of HIES 2010 was reported as 12,240 households, where 7,840 were from rural areas and 4,400 from urban areas. For making theoretically comparable, the required sample size was also calculated using the usual classical formula in equation (11) and multiplying it by design effect (deff) for adjustment of cluster sampling. Note that the choice of design effect = 1.6 is made on the basis of conventional practice in Bangladesh surveys where design effect is assumed to vary from 1.5 to 2.0.

In this computation $CV(x)$, the pre-assumed value of the population coefficient of variation is computed from the HIES 2005 considering "Household Income" as the main interesting variable, $z^2_{\frac{\alpha}{2}} = 1.64$ and the maximum allowable relative margin of error r = 10%. The sample sizes actually used in HIES 2010 along with sample sizes computed using classical and Bayesian methods are given in Table 2. For the Bayesian approach the prior $(\alpha, \beta) = \left(\frac{1}{\sqrt{CV}}, std\right)$ is considered and the CV and standard deviation of "Household Income" computed from the HIES 2005 are used. The CV for rural and urban populations as 3.01 and 4.71 respectively are used in both approaches of sample size determination.

Table 2. Classical and Bayesian Sample Size for HIES 2010 (α = 0.1)

Region	Sample size actually used	Classical method	Hypothetically computed sample size		
			ACC	ALC	WOC
Rural	7840	3922	3621	8542	7096
Urban	4400	9604	9200	4908	5021
Total	12240	13526	12821	13450	12117

From Table 2, it can be observed that the total sample size used in the actual survey is almost same as that determined by the classical method as well as by the three Bayesian criteria. However, the urban-rural split of the sample sizes seems be of reverse order in the hypothetically determined methods. This could be due to the reason that the actual study allocated the size proportionally to the 70%-30% rural-urban population of Bangladesh whereas the classical and Bayesian SSD used in the hypothetical computation considered separate sample sizes for urban and rural domains, and because CV of household income in urban area is much higher than that in rural area, the urban sample sizes is obtained to be larger. It is obvious that the choice of higher sample size in urban area according to the computed sample size could have provided better precision than that possibly been attained in the actual survey.

The comparison between the classical and Bayesian SSD for the said survey reveals not much of difference except that the WOC criteria produced smallest sample size in comparison to the other methods. The ACC criteria and the classical method give almost a same sample size, which implies that with similar level of prior information even the most conservative Bayesian criteria gives as good sample size as the classical method. This result has been revealed in an extensive simulation with different level of significance and different level of precision. However, it can be expected that if higher level of prior information is in hand, Bayesian approach may possibly utilize them and reduce the required number of samples whereas the classical method do not have any option to utilize them. That means that, if a Bayesian approach is applied, as opposed to a classical approach for sample size determination, then it could have optimized the opportunity.

Bangladesh Demographic and Health Survey (BDHS) 2007

The Bangladesh Demographic and Health Survey (BDHS) is a periodic survey conducted in Bangladesh to serve as a source of population and health data for policymakers, program managers, and the research community. The sample size for BDHS 2007 was determined according to six divisions and two regions using BDHS 2004 with the help of the usual SSD formula (see the previous section) and the three Bayesian criteria (as described in the Methodology).

Table 3. Classical and Bayesian Sample size for six divisions and two regions of BDHS 2007 ($\alpha = 0.05$)

Region	Sample size actually used	Classical method	Hypothetically computed sample size		
			ACC	ALC	WOC
Dhaka	2726	376	394	213	398
Chittagong	2423	448	468	259	456
Khulna	1935	683	1124	677	1010
Barisal	1674	1071	1490	912	1256
Rajshahi	2403	707	1064	637	955
Sylhet	1949	267	166	69	187
Total (for division)	11485	3552	4706	2767	4262
Urban	5218	690	1111	669	998
Rural	7981	513	605	346	576
Total (for region)	11485	1203	1716	1015	1574

Considering the variable "children ever born" as the variable of interest, computations similar to those in the previous section were done. The actual sample sizes used in the survey along with the computed required sample sizes using Bayesian and classical methods are given in Table 3. The sample sizes are computed with two different perspectives about domains. Often only the rural/urban segregation of the estimates is needed from surveys; in such cases sample size may be calculated for only those two domains. BDHS 2007 makes separate estimates for the six administrative divisions of Bangladesh and hence these six domains were considered in the computation.

Table 3 shows that the hypothetically determined sample size (classical and Bayesian approach) is very much smaller than actually used sample size of BDHS 2007 for division and regions which indicates oversampling for reporting more reliable estimates of the rarer characteristics of that division or region. However, classical and Bayesian sample sizes are determined using the usual SSD formula due to unavailability of the used sample size formula of BDHS 2007. Because the coefficient of variation of "children ever born" is very low, so that a very much smaller sample size was obtained from classical and Bayesian approach than the used sample size of BDHS 2007. This may be explained because the Bayesian sample size using ACC and WOC criteria is larger than the classical sample size for all division and two regions. This table concludes that the Sylhet division has smallest sample size among all divisions for both approach and actually used sample size of BDHS 2007. Also note that Barisal division has the largest sample size among the other divisions of Bangladesh for classical and Bayesian approach but BDHS 2007 showed that the Dhaka division has the largest sample size among all divisions. This table also shows that the urban-rural sample size is present in reverse order in the hypothetically determined methods like the previous survey (HIES 2010). This statement indicates that the sample size allocation among the urban-rural strata and among the divisions could have possibly been done proportionally.

Conclusion

Results suggest that the classical sample size is larger than Bayesian sample size in the applications examined, although the estimated sample sizes in both methods (classical and Bayesian) are decreased when a larger margin of error is considered. Prior information can reasonably be utilized to improve Bayesian sample size estimation. In Bayesian approach of sample size determination, different prior are used in place of classical estimator. The estimated sample sizes decreased when moving towards informative prior from a non-informative prior. Results from this study show that the proper use of prior information may enhance the strength the of the Bayesian method of sample size determination. Thus, an optimized parsimony could be achieved by use of Bayesian sample size determination with substantially informative priors.

References

Adcock, C. J. (1988). A Bayesian approach to calculating sample sizes. *Journal of the Royal Statistical Society. Series D (The Statistician), 37,* 433-439.

BDHS (Bangladesh Demographic and Health Survey) 2004, National Institute of Population Research and Training (NIPORT), Mitra and Associates and ORC Macro, 2005, NIPORT, Dhaka.

BDHS (Bangladesh Demographic and Health Survey) 2007, National Institute of Population Research and Training (NIPORT), Mitra and Associates and ORC Macro, 2008, NIPORT, Dhaka.

Chen, M. H., & Shao, Q. M. (1998). Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics, 7*.

Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: John Wiley and Sons.

Joseph, L., & Belisle, P. (1997). Bayesian sample size determination for normal means and difference between normal means. *Royal Statistical Society, 46*(2), 209-226.

Joseph, L., Wolfson, D. B., & Berger, R. D. U. (1995). Sample size calculations for binomial proportions via highest posterior density intervals. *Royal Statistical Society, 44*(2), 143-154.

M'Lan, C. E., Joseph, L., & Wolfson, D. B.(2008). Bayesian sample size determination for binomial proportions. *International Society for Bayesian Analysis, 3*(2), 269-296.

Pham-Gia, T. (1995). Sample size determination in Bayesian statistics: A commentary. *Journal of the Royal Statistical Society, 44*(2), 163-166.

Report of the Household Income and Expenditure Survey (2005), Bangladesh Bureau of Statistics (BBS), Planning Division, Ministry of Planning, Government of the People's Republic of Bangldesh, May 2007.

Report of the Household Income and Expenditure Survey (2010), Bangladesh Bureau of Statistics (BBS), Planning Division, Ministry of Planning, Government of the People's Republic of Bangldesh, February 2010.

Sahu, S. K., & Smith, T. M. F. (2006). A Bayesian method of sample size determination with practical applications. *Royal Statistical Society, 169*(Part 2), 235-253.