Published in final edited form as:

J Stat Plan Inference. 2009 December 1; 139(12): 4111–4122. doi:10.1016/j.jspi.2009.05.041.

Comparison of Bayesian Sample Size Criteria: ACC, ALC, and WOC

Jing Cao^{1,*}, J. Jack Lee², and Susan Alber³

¹Department of Statistical Science, Southern Methodist University, Dallas, Texas, 75275

²Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, 77030

³Department of Statistics, Oregon State University, Corvallis, Oregon, 97331

Abstract

A challenge for implementing performance based Bayesian sample size determination is selecting which of several methods to use. We compare three Bayesian sample size criteria: the average coverage criterion (ACC) which controls the coverage rate of fixed length credible intervals over the predictive distribution of the data, the average length criterion (ALC) which controls the length of credible intervals with a fixed coverage rate, and the worst outcome criterion (WOC) which ensures the desired coverage rate and interval length over all (or a subset of) possible datasets. For most models, the WOC produces the largest sample size among the three criteria, and sample sizes obtained by the ACC and the ALC are not the same. For Bayesian sample size determination for normal means and differences between normal means, we investigate, for the first time, the direction and magnitude of differences between the ACC and ALC sample sizes. For fixed hyperparameter values, we show that the difference of the ACC and ALC sample size depends on the nominal coverage, and not on the nominal interval length. There exists a threshold value of the nominal coverage level such that below the threshold the ALC sample size is larger than the ACC sample size, and above the threshold the ACC sample size is larger. Furthermore, the ACC sample size is more sensitive to changes in the nominal coverage. We also show that for fixed hyperparameter values, there exists an asymptotic constant ratio between the WOC sample size and the ALC (ACC) sample size. Simulation studies are conducted to show that similar relationships among the ACC, ALC, and WOC may hold for estimating binomial proportions. We provide a heuristic argument that the results can be generalized to a larger class of models.

Keywords

average coverage criterion; average length criterion; credible interval; coverage rate; interval length; worst outcome criterion

*Corresponding author: jcao@smu.edu.

1. Introduction

Sample size determination (SSD) is a crucial component in the design of experiments and clinical trials. There is a substantial literature on frequentist SSD approaches (Kraemer and Thiemann 1987, Cohen 1988, Desu and Raghavarao 1990). The two issues of *The Statistician* (44 (2) 1995, 46 (2) 1997) presented a summary of the work on Bayesian SSD. Frequentists and Bayesians have a fundamentally different approach for handling the unknown parameters that SSD methods often depend on. For example, suppose we want to find the sample size guaranteeing a certain precision for estimating the mean of a normal distribution with an unknown variance. A frequentist SSD method uses an estimate of the unknown variance from a previous study, or employs a sequential scheme (Liu 1997) where the estimate is obtained from a pilot study. In a Bayesian SSD method the model accounts for uncertainty about both the mean and variance with a prior distribution.

Bayesian methods for SSD can be classified into two approaches. One treats SSD as a decision problem and employs a utility function (Lindley 1997, Pham-Gia 1997) and the other uses a performance based approach to control inference for the parameter of interest with certain precision (Adcock 1988, Joseph and Belisle 1997, Joseph, du Berger, and Belisle 1997, Clarke and Yuan 2006, M'Lan, Joseph, and Wolfson 2008). This paper does not seek to contribute to the debate between Bayesians and frequentists SSD, nor does it take a Bayesian decision theoretical approach. Using the performance based approach, we compare three widely recognized Bayesian SSD criteria for a normal conjugate model and investigate their difference.

In a Bayesian analysis the posterior credible interval is frequently used as a summary of the inference about the parameter of interest. Several criteria for Bayesian SSD based on the length and coverage of these intervals have been proposed, among which the most commonly referenced are the average coverage criterion (ACC), the average length criterion (ALC), and the worst outcome criterion (WOC). The ACC, introduced by Adcock (1988), controls the coverage rate of fixed length credible intervals over the predictive distribution of the data. The ALC, proposed by Joseph and Belisle (1997), controls the length of credible intervals with a fixed coverage rate over the predictive distribution of the data. The WOC, also proposed by Joseph and Belisle (1997), guarantees that the desired coverage rate and interval length over all (or a subset of) possible datasets. Reviews of Bayesian SSD criteria, including these three criteria, are provided by Adcock (1997) and Pezeshk (2003).

The three sample size criteria have been applied to a variety of parameter estimation problems, including multinomial parameters (Adcock 1993), single normal means and differences between normal means (Joseph and Belisle 1997), binomial parameters (Joseph, Du Berger, and Belisle 1997, Rahmea and Joseph 1998, M'Lan, Joseph, and Wolfson 2008), and the exposure odds ratio in case-control studies (M'Lan, Joseph, and Wolfson 2006). In most cases, there is no closed form for either criteria. Wang and Gelfand (2002) proposed a generic simulation based approach to determine sample size based on these criteria.

The WOC usually produces a larger sample size than the ACC and the ALC. Sample sizes obtained by the ACC and the ALC are rarely the same, except for estimating the mean of a

normal distribution with known variance. This hinders the application of these SSD methods, because the literature contains little information about how much larger the WOC sample size is compared to the ACC and ALC sample size, and what determines the direction and magnitude of differences between the sample sizes given by the ACC and the ALC. Thus there is little guidance on which method to use. The goal in this paper is to compare the ACC, ALC and WOC on estimating the mean of a normal distribution with conjugate prior. For fixed hyperparameter values, we show that the difference of the ACC and ALC sample size depends on the nominal coverage, and not on the nominal interval length. There exists a threshold value of the nominal coverage level such that below the threshold the ALC is larger than the ACC, and above the threshold the ACC is larger. Furthermore, the ACC is more sensitive to changes in the nominal coverage. We also show that for fixed hyperparameter values, there exists an asymptotic constant ratio between the WOC sample size and the ALC (ACC) sample size. We then provide a heuristic argument that the results can be generalized to a larger class of models.

2. Performance based Bayesian SSD

Let $\theta \in \Theta$ be the parameter of interest and $\pi(\theta)$ the prior density of θ . Let n be the sample size, $x = (x_1, x_2, \dots, x_n)$ be the experimental data with sample space χ and density $f(x|\theta)$. The predictive marginal distribution of the data x is

$$f(x) = \int_{\Theta} f(x|\theta) \pi(\theta) d\theta,$$

and the posterior density of θ given x is

$$f(\theta|x) = f(x|\theta) \pi(\theta) / f(x).$$

A posterior credible interval is commonly used as a summary of the inference about θ . Both the ACC and the ALC are based on achieving a certain degree of precision provided by this interval estimation of θ .

2.1. Average coverage criterion

For a highest posterior density (HPD) credible interval with fixed length l , the corresponding coverage level varies with x across the data space χ . Adcock (1988) introduced the ACC sample size, which is the smallest integer such that for a fixed nominal interval length l the expected coverage level is at least $1 - \alpha$, where expectation is taken over the marginal distribution of x . The ACC sample size is the smallest integer n that satisfies

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

where $a(x, n)$ is the lower bound of the HPD interval of length l .

2.2. Average length criterion

The ALC, introduced by Joseph and Belisle (1997), reverses the roles of l and α , fixing a nominal coverage level $1 - \alpha$ which has a corresponding length that varies across x . The ALC sample size is the smallest integer such that for a fixed nominal coverage level $1 - \alpha$ the expected length is at most l , where expectation is taken over the marginal distribution of x . The ALC has a two-step formula. The first step is to find the HPD length $l'(x, n)$ that satisfies

$$\int_{a(x,n)}^{a(x,n)+l'(x,n)} f(\theta|x, n) d\theta = 1 - \alpha,$$

where $l'(x, n)$ is the length of the $1 - \alpha$ credible interval. Then choose the smallest integer n such that

$$\int_{\mathcal{X}} l'(x, n) f(x) dx \leq l. \quad (2)$$

Joseph and Belisle (1997) argued that because most researchers tend to report the length of the confidence interval of a fixed coverage, instead of the coverage of the interval of a fixed length, the ALC is more conventional than the ACC.

2.3. Worst outcome criterion

Joseph and Belisle (1997) also presented the WOC, which ensures the desired coverage rate and interval length over all (or a subset of) possible datasets. Such a conservative sample size can give more assurance than the “average” assurances provided by the ACC and ALC criteria. The WOC sample size is determined by the smallest integer n such that

$$\inf_{x \in \mathcal{S}} \left\{ \int_{a(x,n)}^{a(x,n)+l} f(\theta|x, n) d\theta \right\} \geq 1 - \alpha, \quad (3)$$

where $\mathcal{S}$ is a suitably chosen subset of the data space $\mathcal{X}$. For example, $\mathcal{S}$ can consist of the most likely 95% of the possible data $x \in \mathcal{X}$.

3. Comparison of sample sizes for the normal conjugate model

We investigate differences between sample sizes based on the ACC, ALC, and WOC for the normal model with conjugate prior. Three scenarios are considered: estimating a single normal mean (Scenario 1), estimating the difference between two normal means (Scenario 2), and estimating a single normal mean with sampling prior and fitting prior (Scenario 3) discussed by Wang and Gelfand (2002).

3.1. Scenario 1

We adopt the notation by Joseph and Belisle (1997) to facilitate comparison. Let n be the number of observations and $\mathbf{x}_n = (x_1, \dots, x_n)$ be the data. For $i = 1, \dots, n$, let x_i be normally

distributed with mean μ and variance σ^2 . Conditional on μ and σ^2 , $x_1, \dots, x_n$ are independent. Let μ have a normal prior with mean μ_0 and variance σ^2/n_0 .

When the variance σ^2 is known, all the three criteria give the same sample size, which is the smallest integer n that satisfies

$$n \geq \frac{4\sigma^2 z_{1-\alpha/2}^2}{l^2} - n_0 \quad (4)$$

where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution. The standard frequentist sample size, which assumes both σ^2 and μ are fixed, is the smallest integer n that satisfies

$$n \geq \frac{4\sigma^2 z_{1-\alpha/2}^2}{l^2}. \quad (5)$$

In the Bayesian model the hyperparameter n_0 controls the weight given to the prior mean μ_0 . The posterior expected value of μ is the average of the prior mean μ_0 , weighted by n_0 , and the sample mean $\bar{x}$, weighted by n . Therefore the quantity n_0 , which is call the prior sample size, can be interpreted as having n_0 observations centered at μ_0 before the study begins, and $n + n_0$ can be interpreted as the total effective sample size. The Bayesian sample size (4) gives n_0 fewer observations than the frequentist sample size (5). As n_0 approaches 0 the prior variance for μ approaches infinity and the Bayesian sample size (4) approaches the frequentist sample size (5).

Both sample size formulas (4) and (5) assume σ^2 is known. However, this is rarely the case, and a common practice for the frequentist is to replace σ^2 with an estimate from a previous study. If the estimate is smaller or larger than the true variance then the sample size will be too small or too large, respectively. Unlike the frequentist approach, a Bayesian only uses (4) when σ^2 is truly known.

When σ^2 is unknown the ACC, ALC, and WOC give different sample size. Let σ^2 have an inverse gamma prior with shape parameter ν and scale parameter β , and density proportional to $(\sigma^2)^{-(\nu+1)} \exp(-\beta/\sigma^2)$, $\nu > 0, \beta > 0$.

Adcock (1988) has shown that the ACC sample size is the smallest integer n that satisfies

$$n + n_0 \geq \frac{4t_{2\nu, 1-\alpha/2}^2 \beta}{l^2 \nu}. \quad (6)$$

As ν gets very large the uncertainty of the variance becomes very small. The limiting model, albeit with different hyperparameters, is the known variance model, for which the three sample sizes are the same.

Joseph and Belisle (1997) have shown that, for $\nu > 0.5$, the sample size based on the ALC is the smallest integer n that satisfies

$$2t_{n+2\nu,1-\alpha/2} \sqrt{\frac{2\beta}{(n+2\nu)(n+n_0)} \frac{\Gamma\left(\frac{n+2\nu}{2}\right) \Gamma\left(\frac{2\nu-1}{2}\right)}{\Gamma\left(\frac{n+2\nu-1}{2}\right) \Gamma(\nu)}} \leq l, \quad (7)$$

where $t_{d,1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the t distribution with d degrees of freedom. When $\nu < 0.5$ equation (7) is undefined because $(2\nu - 1)/2 < 0$ making the top right gamma function undefined. Therefore, in the remainder of the paper we assume that $\nu > 0.5$.

Joseph and Belisle (1997) have also derived the formula for the WOC sample size. Let $\int_{\mathcal{S}} f(x)dx = 1 - w$ and $f(x) \neq f(y)$ for all $x \in \mathcal{S}$ and $y \notin \mathcal{S}$. For example, if $w = 0.05$, then $\mathcal{S}$ is the 95% highest posterior density region according to the predictive distribution $f(x)$. The WOC sample size is the smallest integer n that satisfies

$$\frac{l^2(n+2\nu)(n+n_0)}{8\beta(1+(n/2\nu)F_{n,2\nu;1-w})} \geq t_{n+2\nu;1-\alpha/2}^2. \quad (8)$$

We begin our investigation of the relationship between the ACC and ALC sample size by developing an approximation to the ALC sample size given in equation (7) that will facilitate the comparison. Applying Stirling's approximation

$$\frac{\Gamma(z+a)}{\Gamma(z+b)} = z^{a-b} \left[1 + \frac{(a-b)(a+b-1)}{z} + O(z^{-2}) \right], \text{ as } z \rightarrow \infty,$$

to the ratio of gamma functions containing n in (7) gives

$$\frac{\Gamma\left(\frac{n+2\nu}{2}\right)}{\Gamma\left(\frac{n+2\nu-1}{2}\right)} = \sqrt{\frac{n+2\nu}{2}} + O((n+2\nu)^{-1/2}), \text{ as } n \rightarrow \infty. \quad (9)$$

For $n > 30$, the approximation using the first term in equation (9) is less than 3% below the true value, and we assume from here on that $n > 30$. Using the approximation in (9), the ALC inequality (7) can be written as

$$n+n_0 \geq \frac{4t_{n+2\nu,1-\alpha/2}^2 \beta}{l^2} \frac{1}{\nu C_\nu}, \quad (10)$$

where

$$C_\nu = \left(\frac{\Gamma(\nu)}{\Gamma\left(\frac{2\nu-1}{2}\right)} \right)^2 \frac{1}{\nu}.$$

We compare the ALC and ACC sample size in terms of the quantity $n + n_0$. Both n_0 and $n + n_0$ have the same interpretation as the known variance case: n_0 as a prior sample size and $n +$

n_0 as the total effective sample size. Using $z_{1-\alpha/2}$ to approximate $t_{n+2\nu, 1-\alpha/2}$, the ratio of $(n + n_0)$ based on the ACC sample size given by equation (6) and the ALC sample size given by equation (10) is

$$\frac{(n+n_0)_{ACC}}{(n+n_0)_{ALC}} = \frac{t_{2\nu, 1-\alpha/2}^2}{z_{1-\alpha/2}^2} C_\nu, \text{ as } n \rightarrow \infty, \quad (11)$$

where $(n + n_0)_{ACC}$ and $(n + n_0)_{ALC}$ are the total effective samples sizes based on the ACC and the ALC, respectively.

For $\nu > 0.5$, there is a threshold significance level α_ν^* such that

$$\frac{t_{2\nu, 1-\alpha/2}^2}{z_{1-\alpha/2}^2} C_\nu \begin{cases} > 1 & \text{if } \alpha < \alpha_\nu^*; \\ = 1 & \text{if } \alpha = \alpha_\nu^*; \\ < 1 & \text{if } \alpha > \alpha_\nu^*, \end{cases} \quad (12)$$

and consequently

$$(n+n_0)_{ACC} \begin{cases} > (n+n_0)_{ALC} & \text{if } \alpha < \alpha_\nu^*; \\ = (n+n_0)_{ALC} & \text{if } \alpha = \alpha_\nu^*; \\ < (n+n_0)_{ALC} & \text{if } \alpha > \alpha_\nu^*. \end{cases} \quad (13)$$

The threshold α_ν^* depends only on ν . As shown in figure 1, α_ν^* is an increasing function of ν . When $\nu = 1$ the threshold $\alpha_1^* \approx 0.101$, and as ν gets large α_ν^* approaches 0.156.

The relative differences between $(n + n_0)_{ACC}$ and $(n + n_0)_{ALC}$ depend only on ν and α . Figure 2 shows the relationship given by (13) for ν equal to 1, 2, 3, and 10. For any ν , as the distance between α and α_ν^* increases the relative difference between $(n + n_0)_{ACC}$ and $(n + n_0)_{ALC}$ increases. For any ν and any constant $d \in (0, \alpha_\nu^*)$, the relative difference between the two sample sizes is larger for $\alpha = \alpha_\nu^* - d$ than for $\alpha = \alpha_\nu^* + d$. Therefore the largest differences occur when α is small and $(n + n_0)_{ACC} > (n + n_0)_{ALC}$.

Under the assumed model, the WOC sample size is larger than the ACC and ALC sample size. Next we show that for fixed hyperparameter values, there exists an asymptotic constant ratio between the WOC sample size and the ALC (ACC) sample size. Using $z_{1-\alpha/2}$ to approximate $t_{n+2\nu, 1-\alpha/2}$ and $2\nu/\chi_{2\nu;w}^2$ to approximate $F_{n, 2\nu; 1-w}$, where $\chi_{2\nu;w}^2$ is the w quantile of the chi-squared distribution with 2ν degrees of freedom, the WOC inequality (8) can be written as

$$n+n_0 \geq \frac{8\beta z_{1-\alpha/2}^2}{l^2 \chi_{2\nu;w}^2}, \text{ as } n \rightarrow \infty. \quad (14)$$

Then we have

$$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ALC}} = \frac{2\nu}{\chi^2_{2\nu;w}} C_\nu, \text{ as } n \rightarrow \infty. \quad (15)$$

The ratio in (15) is greater than one for $\nu > 0.5$ and a small w (i.e., $w = 0.05$), which is in accordance with the fact that the WOC is a more conservative sample size criterion than the ALC. Note that the above ratio only depends on parameter ν given the subset of data space S , it does not depend on interval length l or the nominal coverage $1 - \alpha$. Based on (11) and (15), we have

$$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ACC}} = \frac{2\nu z_{1-\alpha/2}^2}{\chi^2_{2\nu;w} t_{2\nu,1-\alpha/2}^2}, \text{ as } n \rightarrow \infty. \quad (16)$$

We present numerical examples in Table 1 to compare the sample sizes based on the three criteria. Using the same hyperparameter values ($\nu = \beta = 2$ and $n_0 = 10$) as in Joseph and Belisle's paper (1997), we calculated the sample sizes using the R functions from Joseph's website (<http://www.medicine.mcgill.ca/epidemiology/Joseph>). We also included the approximate ALC and WOC ($w = 0.05$) sample sizes based on (10) and (14), denoted as $ALC_{(10)}$ and $WOC_{(14)}$. They are quite close to the exact sample sizes. For $\nu = 2$, the threshold α_2^* is approximately 0.132. We consider the sample sizes for interval lengths 0.3, 0.2, 0.1, and 0.05, with $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.20$. Across all of the α and l values, $(n + n_0)_{WOC}$ is about 3 times and a half larger than $(n + n_0)_{ALC}$. For each α the ratio $(n + n_0)_{ACC}/(n + n_0)_{ALC}$ is approximately constant across the four lengths. For $\alpha = 0.01$, $(n + n_0)_{ACC}$ is about twice as large as $(n + n_0)_{ALC}$; for $\alpha = 0.05$, $(n + n_0)_{ACC}$ is about 28% larger than $(n + n_0)_{ALC}$; while for $\alpha = 0.20$, $(n + n_0)_{ACC}$ is about 9% smaller than $(n + n_0)_{ALC}$. With a vague prior ($\nu = 2$), the ACC is more sensitive to the changes in the nominal coverage, especially when α gets close to 0. Note that $\alpha = 0.20$ and $\alpha = 0.05$ are almost equally distant from the threshold α_2^* . The relative differences are much smaller for $\alpha = 0.20$ because, as discussed above, the relative difference is smaller when α is greater the threshold. While the relative difference between $(n + n_0)_{ACC}$ and $(n + n_0)_{ALC}$ remains constant across the four lengths, as shown in Table 1, the absolute difference increases as length decreases.

Figure 2 also shows that for any α , as ν increases the relative difference between $(n + n_0)_{ACC}$ and $(n + n_0)_{ALC}$ decreases. In other words, as the inverse gamma prior on σ^2 becomes more informative, the ACC and the ALC tend to produce more similar sample size. The ACC and ALC sample size approaches sample size give in (4) as ν gets very large. Table 2 presents the numerical examples with $\nu = \beta = 10$ to demonstrate the comparison of the sample size under a stronger prior. The main difference of the results in Table 1 and 2 is that the ratios $(n + n_0)_{ACC}/(n + n_0)_{ALC}$, $(n + n_0)_{WOC}/(n + n_0)_{ALC}$, and $(n + n_0)_{WOC}/(n + n_0)_{ACC}$ become much closer to 1 in Table 2.

3.2. Scenario 2

In this section, we compare the three sample size criteria for estimating the difference between two normal means. This model was also discussed by Joseph and Belisle (1997).

Let $\mathbf{x}_{n1} = (x_{11}, \dots, x_{n11})$ and $\mathbf{x}_{n2} = (x_{12}, \dots, x_{n22})$ be two independent random vectors. For $i = 1, \dots, n_j$ and $j = 1, 2$, let x_{ij} be normally distributed with mean μ_j and variance σ^2 and μ_j have a normal prior with mean μ_{0j} and variance σ^2/n_{0j} . Sample sizes n_1 and n_2 are obtained to estimate $\theta = \mu_1 - \mu_2$ under the ACC, ALC and WOC criteria.

When the common variance σ^2 is known, all the three criteria give the same sample size. With the constraint $n_1 + n_{01} = n_2 + n_{02}$ which minimizes the posterior variance, the optimal sample size for n_1 satisfies

$$n_1 + n_{01} \geq \frac{8\sigma^2}{l^2} z_{1-\alpha/2}^2,$$

and $n_2 = n_1 + n_{01} - n_{02}$.

When the common variance σ^2 is unknown, σ^2 is assumed to have an inverse gamma prior with scale parameter ν and shape parameter β . Joseph and Belisle (1997) have derived the ACC, ALC, and WOC sample size solutions. The ACC is satisfied if

$$\frac{(n_1 + n_{01})(n_2 + n_{02})}{(n_1 + n_2 + n_{01} + n_{02})} \geq \frac{4\beta}{\nu l^2} t_{2\nu, 1-\alpha/2}^2.$$

The ALC is satisfied if

$$2t_{n_1+n_2+2\nu, 1-\alpha/2} \sqrt{\frac{2\beta(n_1+n_2+n_{01}+n_{02})}{(n_1+n_2+2\nu)(n_1+n_{01})(n_2+n_{02})} \frac{\Gamma(\frac{n_1+n_2+2\nu}{2})\Gamma(\frac{2\nu-1}{2})}{\Gamma(\frac{n_1+n_2+2\nu-1}{2})\Gamma(\frac{2\nu}{2})}} \leq l.$$

And the WOC is satisfied if

$$\frac{l^2(n_1+n_{01})(n_2+n_{02})(n_1+n_2+2\nu)}{8\beta(n_1+n_2+n_{01}+n_{02})[1+\{(n_1+n_2)/2\nu\}F_{n_1+n_2, 2\nu; 1-w}]} \geq t_{n_1+n_2+2\nu, 1-\alpha/2}^2.$$

Applying the same techniques in Section 3.1, we can approximate the ALC by

$$\frac{(n_1+n_{01})(n_2+n_{02})}{(n_1+n_2+n_{01}+n_{02})} \geq \frac{4z_{1-\alpha/2}^2}{l^2} \frac{\beta}{\nu} \frac{1}{C_\nu}, \text{ as } n_1+n_2 \rightarrow \infty,$$

and the WOC sample size by

$$\frac{(n_1+n_{01})(n_2+n_{02})}{(n_1+n_2+n_{01}+n_{02})} \geq \frac{8\beta z_{1-\alpha/2}^2}{l^2 \chi_{2\nu;w}^2}, \text{ as } n_1+n_2 \rightarrow \infty.$$

Under the constraint $n_1 + n_{01} = n_2 + n_{02}$ it can be shown that

$$\frac{(n_1+n_{01})_{ACC}}{(n_1+n_{01})_{ALC}} = \frac{t_{2\nu,1-\alpha/2}^2}{z_{1-\alpha/2}^2} C_\nu, \text{ as } n_1 \rightarrow \infty,$$

and the same threshold significance level α_ν^* (12) exists regarding the ratio of the ACC and ALC sample size. Furthermore, there is an asymptotic constant ratio between $(n_1 + n_{01})_{WOC}$ and $(n_1 + n_{01})_{ALC}$ remains the same as in (15),

$$\frac{(n_1+n_{01})_{WOC}}{(n_1+n_{01})_{ALC}} = \frac{2\nu}{\chi_{2\nu;w}^2} C_\nu, \text{ as } n_1 \rightarrow \infty.$$

And we have,

$$\frac{(n_1+n_{01})_{WOC}}{(n_1+n_{01})_{ACC}} = \frac{2\nu z_{1-\alpha/2}^2}{\chi_{2\nu;w}^2 t_{2\nu,1-\alpha/2}^2}, \text{ as } n_1 \rightarrow \infty.$$

With the same hyperparameters values (ν , β , and $n_{01} = n_0$), the sample size n_1 to estimate the difference between two normal means is exactly twice as big as the sample size n to estimate a single normal mean under respective criteria. So the conclusions from the comparison based on Table 1 and 2 also apply in Scenario 2.

3.3. Scenario 3

Wang and Gelfand (2002) proposed to use sampling and fitting priors in the simulation-based approach for SSD. Sampling prior $\pi^{(s)}(\theta)$ is what is used to generate parameter θ which is likely to be some specified subset of the parameter space. Fitting prior $\pi^{(f)}(\theta)$ is the one actually used in the analysis when computing the posterior probabilities, and it is relatively noninformative to let the data drive the inference. Based on the simulation-based SSD approach, parameter θ is generated from sampling prior $\pi^{(s)}(\theta)$, and then data x are based on the assumed model $f(x|\theta)$. The posterior distribution of θ is calculated using the fitting prior $\pi^{(f)}(\theta)$. Applying the ACC, ALC, and WOC criteria, corresponding sample sizes can be obtained.

The previous scenarios are specific cases when the sampling and fitting priors are the same under the normal model. However, explicit sample size solutions for the three criteria under the two-prior SSD approach are not available. It prohibits the derivation of the ACC/ALC and WOC/ALC ratio in a closed form. As is explained in the next section, the difference

between the ACC and ALC sample sizes is caused by the shape of the posterior density curve of θ , which leads to a nonlinear relationship between the coverage and the interval length. Based on the argument, as long as $f(\theta|x)$ has a bell-shaped curve, we expect similar relationships among the ACC, ALC, and WOC sample size.

We conducted a simulation study to compare the sample size with two priors for a single normal mean. We used the normal model described in Section 3.1, where we assumed that the sampling prior for σ^2 is the inverse gamma prior $IG^{(s)}(\nu=4, \beta=4)$ and the fitting prior is $IG^{(f)}(\nu=2, \beta=2)$. The results are summarized in Table 3. Note that the sample sizes are smaller than their counterparts in Table 1 where the sampling prior and the fitting prior are the same ($IG(\nu=2, \beta=2)$). With a stronger sampling prior ($IG^{(s)}(\nu=4, \beta=4)$) in this simulation, the predictive distribution of data x is less spread out under the normal model, and a smaller sample size is needed to meet the criteria. Though an explicit form of the threshold significance level α_v^* is not available, the general conclusion still holds which states that for fixed hyperparameter values, the difference of the ACC and ALC sample size depends on the nominal coverage and not on the nominal interval length, and there exists an asymptotic constant ratio between the WOC sample size and the ALC sample size regardless of the nominal coverage and the nominal interval length.

4. Generalization to non-normal models

In this section we address the question of whether the results of sample size comparison based on the normal model can be generalized to other models. Because SSD for estimating a binomial parameter is a common situation in experiment design, we use the binomial model in this section. Let p be the binomial parameter to be estimated. Following Joseph, Wolfson, and Du Berger (1995), we assume that $p \sim \text{Beta}(a, b)$, $a, b > 0$ and $x_n|p \sim \text{Bin}(n, p)$, where $\text{Beta}(a, b)$ refers to a beta distribution with parameters a and b , and $\text{Bin}(n, p)$ represents a binomial distribution with parameters n and p . SSD under the binomial model does not have closed form solutions for the ACC, ALC, and WOC.

Under the normal model, we show that the ratio of the WOC and ALC sample size approaches a constant as the sample size increases regardless of the nominal coverage and the nominal interval length. M'Lan, Joseph, and Wolfson (2008) provided analytic approximations to the ALC sample size and the WOC sample size under the binomial model. For the ALC, the approximate sample size formula is

$$n+n_0=4\frac{z_{1-\alpha/2}^2}{l^2}\left[\frac{B(a+1/2, b+1/2)}{B(a, b)}\right]^2, a>1, b>1,$$

where $B(a, b)$ is the beta function with parameters a and b , and $n_0 = a + b$. For the WOC, the approximate sample size formula is

$$n+n_0=\frac{z_{1-\alpha/2}^2}{l^2}, a>1, b>1.$$

Then we have

$$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ALC}} = \frac{1}{4 \left[\frac{B(a+1/2, b+1/2)}{B(a, b)} \right]^2}, \text{ as } n \rightarrow \infty.$$

So a similar result between the WOC and ALC sample size also holds under the binomial model.

Joseph, Wolfson, and Du Berger (1995) studied Bayesian SSD in the binomial model. In Table 4, we reproduced a proportion of their results, where the prior for p is set to be $Beta(1, 1)$. As n increases, the ratio of the WOC and ALC sample size in Table 4 approaches 1.621, which is the nominal ratio of n_{WOC}/n_{ALC} when $a = b = 1$. Without closed form solution/approximation for the ACC sample size, we can not explicitly study the relationship between the ALC and ACC sample size. However, we can see that when $\alpha = 0.5$, the ACC sample size is smaller than the ALC sample size. As α decreases to 0.10 and less, the ACC sample size is larger. Also, as α decreases towards 0, the relative difference between the ACC and ALC sample size increases. The results hold regardless of the nominal length. In Table 5, we compared the sample sizes under the three criteria using the simulated-based approach, where we set the sampling prior for p to be $Beta(4, 1)$ and the fitting prior to be $Beta(1, 1)$. The results in Table 4 and 5 are consistent with the conclusions on the ACC, ALC, and WOC sample size discussed under the normal model.

Mathematical intractability in the ACC prohibits derivation of simple algebraic relationships between the ACC and the other two criteria for the binomial model. We provide a heuristic argument on why it is likely that there exists a threshold α for which equation (11) holds in a large class of models including the binomial model. We do not attempt to find a general approach for finding the threshold α , nor do we discuss the magnitude of differences between the ALC and ACC.

Instead of comparing the ALC sample size directly with the ACC sample size, our strategy is to investigate whether the interval coverage in ACC will be controlled at the nominal level with the ALC sample size. If the coverage is controlled, then the ACC sample size is no more than the ALC sample size. Otherwise, the ACC sample size is larger than the ALC sample size.

For a specific sample size, the predictive marginal distribution of the data $f(x)$ is fixed, and the possible posterior distributions are fully determined. We use Figure 3 to demonstrate the relationships between the coverage and interval length of credible intervals. Each column of Figure 3 shows three posterior density curves with small, moderate and large variances. Each posterior distribution is centered around its mean. This is because, for both ACC and ALC, it is the shape of the posterior distribution that matters, not the location. To present a numerical comparison, assume that the posterior distributions can be well approximated by normal distributions with small, moderate, and large standard deviations $s(x)$ proportional to 2, 5, and 8, respectively. Additionally, assume that the three posterior densities are

representative of the posterior densities with data x generated from $f(x)$, and they have respective weight 0.25, 0.5, and 0.25. For a fixed nominal coverage level, the interval length varies under different posterior distributions. The solid line in each plot shows the individual interval length, we call it the ALC-length. The length of the dotted line is exactly the nominal length l . The three plots in the each column shows the change in the coverage (d) with the ALC-length and the nominal length l . With the ALC-length, the nominal coverage $1 - \alpha$ is guaranteed, and the integral of the ALC-length is the nominal length l . Compare the three plots in the left column where $\alpha = 0.05$. From the ACC perspective, under each posterior density, the interval length is the fixed l . The coverage with $s(x)$ proportional to 2 is 0.999999, the coverage with $s(x)$ proportional to 5 is 0.95, and the coverage with $s(x)$ proportional to 8 is 0.78. The change of the coverage rate from the ALC-length to the nominal length is +0.05, 0, and -0.17, respectively. This nonproportional change in the coverage is because of the shape of the density curve which has convex flat tails, and thus, the relationship between the coverage and the interval length is not linear. When the nominal coverage is close to 1, in this case $1 - \alpha = 0.95$, the increment in the interval length only gains a very small increment in the coverage (see the first panel, $d = 0.05$). However, a decrease in the interval length by the same degree leads to a much larger decrease in the coverage (see the third panel, $d = -0.17$). When α gets bigger, because the density curve has a convex steep center, the relationship is reversed (see the plots on the right column with $\alpha = 0.30$).

With $\alpha = 0.05$, the ACC coverage with the ALC sample size is $0.25(0.999999) + 0.5(0.95) + 0.25(0.7794) = 0.9198$. To attain the nominal coverage $1 - \alpha = 0.95$, the ACC needs a larger sample size than the ALC. With $\alpha = 0.30$, the ACC coverage is $0.25(0.99) + 0.5(0.70) + 0.25(0.48) = 0.7175$. To attain the nominal coverage $1 - \alpha = 0.70$, the ACC needs a smaller sample size than the ALC.

Given the hyperparameters, it is the nominal coverage and the nominal length together that determine the ACC and the ALC sample size. However, we conjecture that it is the nominal coverage alone that determines the relative difference of the two. In general, this is because the posterior density curve has concave flat tails and convex steep center, and thus, the relationship between the coverage and the interval length is not linear.

When the sample size is large, the posterior density of θ , the parameter of interest, can be approximated by a normal distribution. Then the conclusion on the comparison of the ACC and the ALC holds. For small or moderate sample sizes, or when θ takes extreme values in the parameter space (such as a binomial parameter concentrates near 0 or 1), the posterior density of θ may be skewed or even have an irregular shape, which means the HPD interval will not be equal-tailed. However, the idea from the previous demonstration still applies. We suggest that when a posterior density has the common density shape of concave flat tails and convex steep center, the nominal coverage $1 - \alpha$ will determine the relative difference between the ALC sample size and the ACC sample size.

5. Example

As an alternative way of using medications prescribed by physicians, lifestyle modification including dietary habits can be used to treat high blood pressure. In this section, we compare the ALC, ACC, and WOC sample size in a study to assess the effect of sour milk on blood pressure (BP) in borderline hypertensive men. A previous study (Mizushima *et al.*, 2004) has provided an estimate of the change in systolic BP which was -4.3 mm Hg with 95% confidence interval of (-8.3, -0.4). Suppose a researcher is now interested to estimate the change with a total interval length $l=1, 2$, and 3 mm Hg, respectively.

Table 6 provides a summary of the sample sizes using the formula under *Scenario 1*, where an inverse gamma prior $IG(\nu=2, \beta=50)$ is assumed for σ^2 , the variance of the change in systolic BP. The ratios among the sample sizes are very close to the ratios in Table 1 with $\alpha=0.05$. It confirms our conclusion that under the normal model the relative difference between the ALC and ACC sample size is determined by α and ν and it is not affected by l and β . In addition, the relative difference between the WOC and ALC sample size only depends on ν . Based on Table 6, the researcher may choose the final sample size based on his/her preferred sample size criterion and practical (or budgetary) constraints.

6. Discussion

We have shown that for the normal model with conjugate prior the relative difference between the ALC and ACC sample size depends on the nominal coverage, and not on the interval length. There also exists an asymptotic constant ratio between the WOC sample size and the ALC sample size which does not depend on either the nominal coverage or the interval length. Given the prior, there is a threshold coverage such that the ACC sample size is larger than the ALC when the nominal coverage is above the threshold, and smaller than the ALC when nominal coverage is below the threshold. The relative difference between the ALC and ACC sample size increases as the difference between the nominal and threshold coverage increases. The ACC sample size is more sensitive to changes in the nominal coverage. For the normal model with a vague prior on the unknown variance (see $\nu=1$ in Figure 2), the ACC sample size can be more than four times of the ALC sample size with $\alpha=0.01$; it could reduce to less than 70% of the ALC sample size with $\alpha=0.2$.

We have also provided a heuristic argument that the existence of a threshold coverage extends to a larger class of models whose posterior distributions have the common density shape of concave flat tails and convex steep center. For further research, a formal argument needs to be developed regarding the general conclusion. Along the line, it is of practical interest to determine threshold coverage for the ACC and ALC methods in different models. In addition, more insight is needed to explain why the asymptotic ratio between the WOC sample size and the ALC sample size does not depend on either the nominal coverage or the interval length.

Lastly, we encourage the use of Bayesian methods for constructing credible intervals to estimate unknown parameters as they take the uncertainty of the parameters into considerations under proper distribution models. Frequentist's confidence intervals depend on the indirect fiducial argument and lack a formal way to incorporate the prior information

on the variation of the parameters. Bayesian ACC, ALC, and WOC criteria each has its merit and drawback. Our paper compares the three criteria and provides further information to help evaluating and choosing among these methods.

Acknowledgments

The authors thank Dr. Peter Mueller and Dr. Peter Thall for their constructive suggestions and Ms. LeeAnn Chastain for editorial assistance. This work was supported in part by grants from the National Cancer Institute CA16672, CA97007, and the Department of Defense W81XWH-06-1-0303.

References

- Adcock CJ. A Bayesian approach to calculating sample sizes. *The Statistician*. 1988; 37:433–439.
- Adcock CJ. An improved Bayesian procedure for calculating sample sizes in multinomial sampling. *The Statistician*. 1993; 42:91–95.
- Adcock CJ. Sample size determination: a review. *The Statistician*. 1997; 46:261–283.
- Clarke B, Yuan A. Closed form expressions for Bayesian sample size. *Annals of Statistics*. 2006; 34:1293–1330.
- Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Erlbaum; Hillsdale, NJ: 1988.
- Desu, MM.; Raghavarao, D. *Sample Size Methodology*. Academic Press; New York: 1990.
- Joseph L, Belisle P. Bayesian sample size determination for normal means and differences between normal means. *The Statistician*. 1997; 46:209–266.
- Joseph L, Du Berger R, Belisle P. Bayesian and mixed bayesian/likelihood criteria for sample size determination. *Statistics in Medicine*. 1997; 16:769–781. [PubMed: 9131764]
- Joseph L, Wolfson DB, Du Berger R. Sample size calculations for binomial proportions via highest posterior density intervals. *The Statistician*. 1995; 44:143–154.
- Kraemer, HC.; Thiemann, S. *How Many Subjects? Statistical Power Analysis in Research*. Sage; Newbury Park, CA: 1987.
- Lindley DV. The choice of sample size. *The Statistician*. 1997; 46:129–138.
- Liu W. On some sample size formulae for controlling both size and power in clinical trials. *The Statistician*. 1997; 46:239–251.
- Mizushima S, Ohshige K, Watanabe J, Kimura M, Kadowaki T, Nakamura Y, Tochikubo O, Ueshima H. Randomized controlled trial of sour milk on blood pressure in borderline hypertensive men. *American Journal of Hypertension*. 2004; 17:701–706. [PubMed: 15288885]
- M'Lan CE, Joseph L, Wolfson DB. Bayesian sample size determination for case-control studies. *Journal of the American Statistical Association*. 2006; 101:760–772.
- M'Lan CE, Joseph L, Wolfson DB. Bayesian sample size determination for binomial proportions. *Bayesian Analysis*. 2008; 3:269–296.
- Pezeshk H. Bayesian techniques for sample size determination in clinical trials: a short review. *Statistical Methods in Medical Research*. 2003; 12:489–504. [PubMed: 14653352]
- Pham-Gia T. On Bayesian analysis, Bayesian decision theory and the sample size problem. *The Statistician*. 1997; 46:139–144.
- Rahme E, Joseph L. Exact sample size determination for binomial experiments. *Journal of Statistical Planning and Inference*. 1998; 66:83–93.
- Wang F, Gelfand AE. A simulation-based approach to Bayesian sample size determination for performance under a given model and for separating models. *Statistical Science*. 2002; 17:193–208.

Figure 1.

Relationship between the threshold significance level α_ν^* and the hyperparameter ν .

Figure 2.

Dependence of $(n + n_0)_{ACC} / (n + n_0)_{ALC}$ on α and ν . The four vertical bars show the location of $\alpha_1^*, \alpha_2^*, \alpha_3^*, \alpha_{10}^*$.

Figure 3.

The change in the coverage rate of the ACC and the ALC (d = coverage with nominal length - coverage with ALC-length, x denotes the data generated from $f(x)$, and s denotes the standard deviation of x). The solid line represents the interval-specific length given the fixed nominal coverage and the dotted line represents the fixed nominal interval length.

Table 1

Comparison of sample sizes for estimating a single normal mean based on different criteria with varying nominal significance level α and length l of the credible interval. The hyperparameters are $\nu = \beta = 2$ and $n_0 = 10$.

α	length	ACC	ALC	ALC ₍₁₀₎	WOC	WOC ₍₁₄₎	$\frac{(n+n_0)_{ACC}}{(n+n_0)_{ALC}}$	ACC-ALC	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ALC}}$	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ACC}}$
0.01	0.3	933	456	453	1651	1650	2.023	477	3.564	1.761
	0.2	2110	1035	1032	3726	3724	2.029	1075	3.575	1.762
	0.1	8470	4162	4159	14928	14927	2.032	4308	3.581	1.762
	0.05	33907	16668	16665	59738	59737	2.034	17239	3.583	1.762
0.05	0.3	333	260	258	951	951	1.270	73	3.559	2.802
	0.2	761	595	593	2152	2152	1.274	166	3.574	2.804
	0.1	3074	2405	2404	8638	8638	1.277	669	3.581	2.804
	0.05	12324	9646	9645	34582	34582	1.277	2678	3.582	2.805
0.20	0.3	95	105	105	400	401	0.913	-10	3.565	3.905
	0.2	226	248	248	914	914	0.915	-22	3.581	3.915
	0.1	931	1022	1022	3687	3687	0.912	-91	3.582	3.929
	0.05	3752	4118	4118	14779	14779	0.911	-366	3.583	3.931

Table 2

Comparison of sample sizes for estimating a single normal mean based on different criteria with varying nominal significance level α and length l of the credible interval. The hyperparameters are $\nu = \beta = 10$ and $n_0 = 10$.

α	length	ACC	ALC	ALC ₍₁₀₎	WOC	WOC ₍₁₄₎	$\frac{(n+n_0)_{ALC}}{(n+n_0)_{ALC}}$	ACC-ALC	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ALC}}$	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ACC}}$
0.01	0.3	350	311	309	532	534	1.121	39	1.688	1.506
	0.2	800	710	707	1212	1213	1.125	90	1.697	1.509
	0.1	3229	2861	2858	4880	4882	1.128	368	1.703	1.510
	0.05	12944	11465	11462	19556	19557	1.129	1479	1.705	1.510
0.05	0.3	184	176	175	302	305	1.043	8	1.677	1.608
	0.2	426	407	405	695	698	1.046	19	1.691	1.617
	0.1	1731	1652	1651	2820	2822	1.048	126	1.703	1.626
	0.05	6952	6634	6632	11316	11319	1.048	318	1.705	1.627
0.20	0.3	69	69	69	121	125	1.000	0	1.658	1.658
	0.2	166	168	167	289	293	0.989	-2	1.680	1.699
	0.1	693	700	700	1197	1201	0.990	-7	1.700	1.717
	0.05	2801	2830	2830	4830	4834	0.990	-29	1.704	1.722

Table 3

Comparison of sample sizes for estimating a single normal mean based on different criteria with sampling prior $IG^{(\nu)}(\nu = 4, \beta = 4)$ and fitting prior $IG^{(\beta)}(\nu = 2, \beta = 2)$ and $n_0 = 10$.

α	length	ACC	ALC	WOC	$\frac{(n+n_0)_{ACC}}{(n+n_0)_{ALC}}$	ACC-ALC	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ALC}}$	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ACC}}$
0.01	0.3	537	392	926	1.361	145	2.328	1.711
	0.2	1217	886	2085	1.369	331	2.338	1.707
	0.1	4848	3551	8300	1.364	1297	2.334	1.711
	0.05	19383	14173	33133	1.367	5210	2.337	1.709
0.05	0.3	248	220	529	1.122	28	2.343	2.089
	0.2	567	504	1197	1.123	63	2.348	2.092
	0.1	2312	2053	4806	1.126	259	2.334	2.074
	0.05	9213	8217	19165	1.121	996	2.331	2.079
0.20	0.3	85	88	219	0.969	-3	2.336	2.411
	0.2	204	212	511	0.964	-8	2.348	2.436
	0.1	842	869	2046	0.969	-27	2.339	2.413
	0.05	3398	3512	8208	0.968	-114	2.333	2.411

Table 4

Comparison of sample sizes for estimating a binomial proportion based on different criteria with the uniform prior.

α	length	ACC	ALC	WOC	$\frac{(n+n_0)_{ACC}}{(n+n_0)_{ALC}}$	ACC-ALC	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ALC}}$	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ACC}}$
0.01	0.10	512	405	661	1.263	107	1.629	1.290
	0.05	2058	1633	2652	1.260	425	1.623	1.288
	0.01	51552	40923	66347	1.260	10629	1.621	1.287
0.05	0.10	274	234	382	1.169	40	1.627	1.391
	0.05	1105	945	1535	1.169	160	1.623	1.388
	0.01	27691	23693	38413	1.169	3998	1.621	1.387
0.10	0.10	183	164	269	1.114	19	1.633	1.465
	0.05	739	665	1080	1.111	74	1.622	1.460
	0.01	18533	16686	27053	1.111	1847	1.621	1.460
0.50	0.10	22	26	43	0.857	-4	1.607	1.875
	0.05	95	110	180	0.866	-15	1.625	1.876
	0.01	2444	2804	4547	0.872	-360	1.621	1.860

Comparison of sample sizes for estimating a binomial proportion based on different criteria with sampling prior $Beta^{(s)}(4, 1)$ and fitting prior $Beta^{(f)}(1, 1)$.

Table 5

α	length	ACC	ALC	woC	$\frac{(n+n_0)_{ACC}}{(n+n_0)_{ALC}}$	ACC-ALC	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ALC}}$	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ACC}}$
0.01	0.10	442	311	662	1.419	126	2.121	1.495
	0.05	1745	1230	2609	1.418	515	2.119	1.495
	0.01	43403	31004	65266	1.400	12399	2.105	1.504
0.05	0.10	229	185	385	1.235	44	2.070	1.675
	0.05	909	706	1511	1.287	203	2.137	1.661
	0.01	21991	17155	36622	1.282	4836	2.135	1.665
0.10	0.10	146	122	266	1.194	24	2.161	1.811
	0.05	583	499	1065	1.168	84	2.130	1.824
	0.01	14109	12032	25812	1.173	2077	2.145	1.829
0.50	0.10	18	22	44	0.833	-4	1.917	2.300
	0.05	66	84	178	0.791	-18	2.093	2.647
	0.01	1650	2055	4435	0.803	-405	2.157	2.686

Table 6

Comparison of sample sizes in the example based on different criteria with $\alpha = 0.05$, $\nu = 2$, $\beta = 50$, and $n_0 = 10$.

length	ACC	ALC	WOC	$\frac{(n+n_0)_{ACC}}{(n+n_0)_{ALC}}$	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ALC}}$	$\frac{(n+n_0)_{WOC}}{(n+n_0)_{ACG}}$
3	59	76	230	1.246	3.478	2.791
2	142	183	531	1.270	3.559	2.803
1	595	761	2152	1.274	3.574	2.804