1-1-2008

Comparison of confidence intervals for binomial proportions

Narain Armbya
University of Nevada, Las Vegas

COMPARISON OF CONFIDENCE INTERVALS FOR BINOMIAL PROPORTIONS

by

Narain Armbya

Master of Science
University of Nevada, Las Vegas
2004

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science Degree in Mathematical Sciences
Department of Mathematical Sciences
College of Sciences

Graduate College
University of Nevada, Las Vegas
December 2006

UMI Number: 1462874

INFORMATION TO USERS

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

Thesis Approval

The Graduate College
University of Nevada, Las Vegas

November 15, 2006

The Thesis prepared by

Narain Armbya

Entitled

Comparison of Confidence Intervals for Binomial Proportions

is approved in partial fulfillment of the requirements for the degree of

Master of Science in Mathematical Sciences

Examination Committee Chair

Dean of the Graduate College

Examination Committee Member

Examination Committee Member

Graduate College Faculty Representative

1017-53

ABSTRACT

Comparison of Confidence Intervals For Binomial Proportions

by

Narain Armbya

Dr. Rohan Dalpatadu, Examination Committee Chair
Associate Professor, Department of Mathematical Sciences
University of Nevada, Las Vegas

The main objective of this thesis is to compare the performance of confidence intervals for binomial proportions and also propose a Bayesian analysis for estimating the credible sets. In this thesis, a combination of analytical and numerical techniques is used to compare the Wald interval, exact interval and obtain the Bayesian credible sets for the binomial distribution when p is close to 0 or 1. Uniform and Beta priors were used and the credible sets were obtained. The statistical package R was used for this purpose.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

It is a great pleasure to thank many people whose help and suggestion were so valuable in completing this thesis. I am grateful to Dr. Ashok Singh and Dr. Dennis Murphy for their complete support, encouragement and guidance throughout my work, with out whose help, the completion of this thesis is not possible. It is because of them that my knowledge of statistics is stronger than it was before.

I would like to give my special thanks to Dr. Rohan Dalpatadu and Dr. Xin Li for their participation in my thesis committee. Their advice and patience is appreciated. I would also like to thank the College of Sciences at the University of Nevada, Las Vegas for offering me the financial support throughout my graduate studies.

This thesis is dedicated to my parents and my sister for their constant support. I am also extremely grateful to my friend Mira Capur for her support.

CHAPTER 1

INTRODUCTION

1.1. Confidence Intervals and Credible Sets

A confidence interval, for a population parameter, is an interval between two numbers with an associated probability *p* which is generated from a random sample of an underlying population, such that if the sampling was repeated numerous times and the confidence interval recalculated from each sample according to the same method, a proportion *p* of the confidence intervals would contain the population parameter in question. A $(1 - \alpha)100\%$ confidence interval for a parameter θ is the interval (l, u) such that

$$P(l \leq \theta \leq u) = 1 - \alpha.$$

A credible set is a Bayesian analogue of a confidence interval. A $(1 - \alpha)100\%$ credible set for a parameter θ is the set (l, u) such that

$$P(l \leq \theta \leq u \mid x) = \int_l^u g(\theta \mid x) d\theta \geq 1 - \alpha.$$

The coverage probability of the confidence interval is based on the sampling distribution of the parameter; in other words, how it varies over all possible samples. Hence the probabilities are determined pre-data. They do not depend on the particular sample that occurred. This is in contrast to the Bayesian credible set calculated from the posterior distribution that has a direct probability interpretation conditional on the observed sample data. The Bayesian credible set is more useful, i.e., it summarizes the belief about the parameter values that could credibly be believed given the observed data that occurred. In other words, it is post data. One need not be concerned about the data that could have occurred but did not (Bolstad, 2004).

1.2. Binomial Confidence Interval

The binomial distribution is a discrete probability distribution whose probability mass function f(x) determines the probability of obtaining exactly x successes out of n Bernoulli trials where each Bernoulli trial has success with a probability p and failure probability 1-p. The probability mass function of a binomial distribution is given by the formula

$$f(x) = {}^{n}C_{x} p^{x} (1-p)^{n-x}$$

The binomial distribution will be discussed in detail in Chapter 2.

The observed proportion of successes in a binomial sample is denoted by

$$\hat{p} = \frac{x}{n}.$$

One of the most basic and methodologically important problems in statistical practice is the interval estimation of the probability of success in a binomial distribution. For a binomial proportion, four types of confidence intervals can be distinguished: Wilson's score interval (Wilson, 1927), the Wald interval (Wald & Walfowitz, 1939), the adjusted Wald interval (Agresti & Coull, 1998), and the 'exact' Clopper-Pearson interval (Clopper& Pearson, 1934). The text-book confidence interval which is widely used in practice is the Wald interval given by the formula,

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution (Wald and Wolfowitz, 1939).

At first glance, one may think that the problem is too simple and has a clear and present solution when in fact, it does not. It is widely recognized that the actual coverage probability of the standard interval is poor for p near 0 or 1 unless n is very large; for example see Ghosh (1979) or Blythe and Still (1983). Even at the level of introductory statistics texts, the standard interval is often presented with the caveat that it should be used only when n·(min p, 1 − p) is at least 5 (or 10). Examination of the popular texts reveals that the qualifications with which the standard interval is presented are varied, but they all reflect the concern about poor coverage when p is near the boundaries. This is because the exact binomial distribution gets highly skewed as p tends to 0 or 1 and the normal approximation to the sampling distribution of $\hat{p}$ requires large samples before it actually takes hold.

1.3. Bayesian Statistics

Statistical analysis is the process of separating out systematic effects from the random noise inherent in all sets of observations. Jeff Gill in his book Bayesian Methods: A social and behavioral science approach has stated that there are three general steps in this process: collection, analysis and statistical inferences. Suppose there exists a statistical data analysis process with the following desirable characteristics:

- Overt and clear model assumptions

- A rigorous way to make probability statements about the real quantities of theoretical interest.
- An ability to update these statements as new information is received.
- Systematic incorporation of previous knowledge on the subject.
- Missing values handled seamlessly as a part of the estimation process.
- Recognition that population quantities are changing over time rather than forever fixed (i.e., the population of interest is dynamic rather than static).
- The ability to model a wide class of data types (with sufficient computing and programming expertise).
- Straightforward assessment of both model quality and sensitivity to assumptions.

Bayesian statistics possesses all of these qualities and the type of data researchers routinely encounter makes the Bayesian approach ideal in ways that frequentist data analysis cannot match. These advantages include avoiding the assumptions of infinite amounts of forthcoming data, recognition that fixed-point assumptions about human behavior are dubious, and a direct way to include existing expertise (or ignorance) in the analysis.

With Bayesian analysis, inferences about unknown model parameters are not expressed in the conventional way as point estimates with reliability assessed using the null hypothesis significance test. Bayesian analysts make no fundamental distinction between observations and the unknown parameters are treated as random variables themselves as a logical consequence of Bayesian conditional analysis. Bayesian statistical information about parameters is summarized in probability statements applied to samples or populations in the form of a *posterior distribution*:

the joint distribution of unknown parameters after observing the data and updating the model. These summary quantities include quantiles of the posterior distribution, the predictive quantities from the posterior and Bayesian forms of confidence intervals, the credible sets and the highest posterior density region.

The essentials of Bayesian thinking are contained in three general steps:

- Specify a probability model that includes some prior knowledge about the parameters if available for unknown parameter values.
- Update the knowledge about the unknown parameters by conditioning this probability model on observed data.
- Evaluate the fit of the model to the data and the sensitivity of the conclusions to the assumptions.

The value of a given Bayesian model is found in the description of the distribution of some parameter of interest in probabilistic terms (Gill,2002). The framework of the Bayesian analysis will be discussed in depth in Chapter 2.

1.4. Objective and Approach

As mentioned earlier, there many procedures for computing binomial confidence intervals, so the main objective of this thesis is to compare these procedures. More specifically, we will consider the Wald interval, the exact interval and a method for estimating the Bayesian credible sets for the binomial population proportion. A combination of analytical and numerical techniques is used to obtain the Bayesian credible sets for the binomial distribution when p is close to 0 or 1. Chapter 2 explains in detail the binomial model, the Wald interval and the Exact interval. Chapter 3 deals with the Bayesian framework. Uniform and Beta priors are used to calculate the posterior densities. The analytical approach used to get these posteriors is also discussed. Some examples and results are summarized in Chapter 4.

CHAPTER 2

CONFIDENCE INTERVALS FOR BINOMIAL PROPORTIONS

This chapter provides information on the working theory behind the Wald interval and the Exact interval for estimating a binomial proportion. The chapter is split into two sections. The first part provides a detailed description of the concept of binomial distribution, the second presents the details of Wald Interval and the Exact interval

2.1. The Binomial Distribution

Probability distributions are used to model randomness in populations; as such, statisticians usually deal with a family of distributions rather than a single distribution. There are two major types of probability distributions: discrete and continuous. A real valued random variable X is a function from a sample space into the real numbers, with the property that for every potential outcome X there is an associated probability P[X=x] which exists for all real values of x in the sample space. A random variable X is said to have a discrete distribution if the support of X, the sample space, is countable; in most situations, the random variable has integer-valued outcomes. The second major type of distribution has a continuous support region; in this situation, the sample space is some interval of the real line and the function used to model random behavior over the sample space is called a probability density function (pdf).

The purpose of this chapter is to introduce a particular type of discrete distribution, the binomial distribution, and its relation to other common discrete distributions. For each distribution, we will give its mean and variance and some other useful statistical descriptive measures and interrelationships that may aid understanding.

2.1.1. Specification of binomial distribution

The binomial distribution is based on the idea of a Bernoulli trial. A Bernoulli trial (named for James Bernoulli, one of the founding fathers of probability theory) is a random experiment with exactly two possible outcomes. A random variable X has a Bernoulli (p) distribution if

$$X = \begin{cases} 1 & \text{with probability p} \\ 0 & \text{with probability 1 - p} \end{cases}, \text{ where } 0 \le p \le 1.$$

The value X = 1 is often termed a "success" and p is referred to as the success probability. The value X = 0 is termed a "failure".

The binomial distribution gives the discrete probability distribution P(X = x) of exactly x successes out of n Bernoulli trials (Kotz 1969) . The binomial distribution is there fore given by

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x=1, 2, 3,.....n$$

The following figure is the plot of the binomial pdf for four values of p and n.

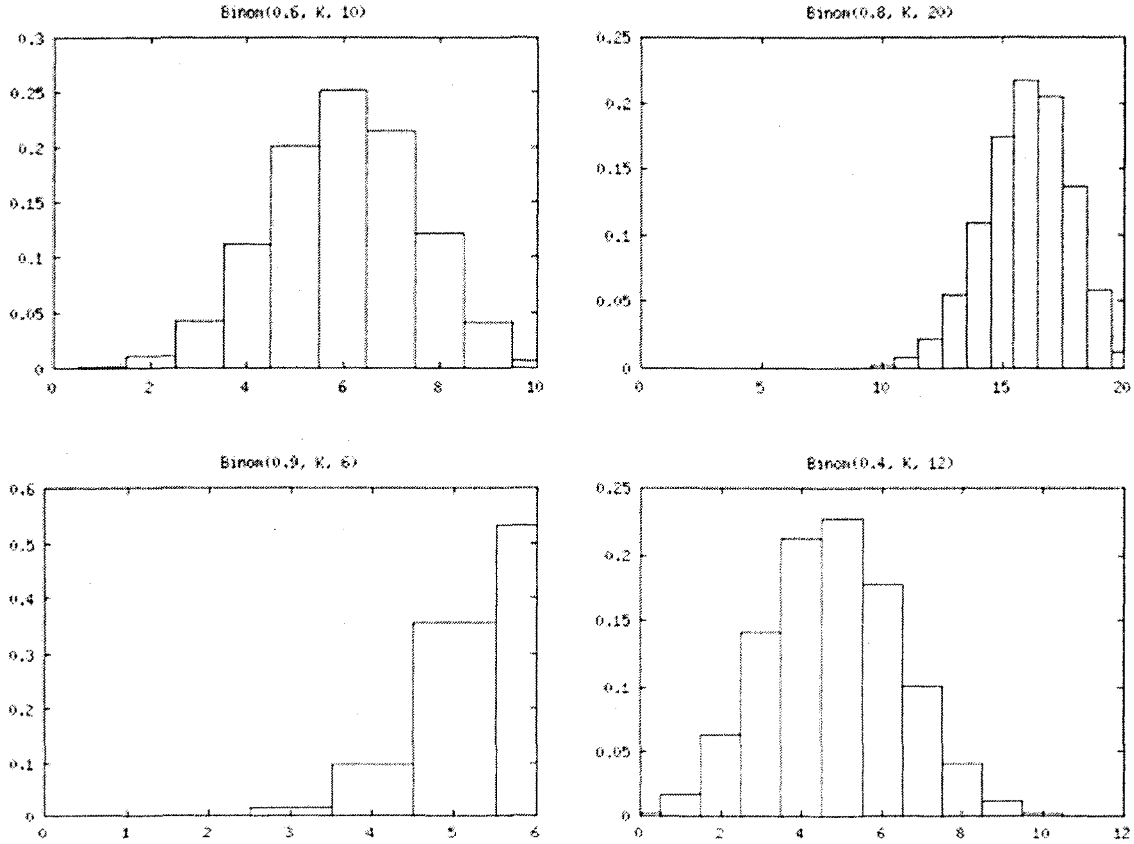


Figure 2.1. Plots of binomial probability distribution function.

2.1.2. Expectation and Variance

The population mean or expected value of X, when X ~ Bin(n, p), is given by

$$E(X) = np$$

and the variance is given by

$$Var(X) = np(1-p).$$

2.1.3. Properties of binomial distribution

The binomial distribution is unimodal, and belongs to the exponential family of distributions with respect to p/(1-p), since we can write

$$P(X = x) = \exp\left[x \log\left\{ \frac{p}{1-p} \right\} + \log\binom{n}{c} + n\log(1-p) \right].$$

It has also been shown that the binomial distribution belongs to the family of F-distributions when the duality between binomial and beta distribution is observed. Kemp (1968) has shown that binomial distribution is a generalized hypergeometric distribution.

The skewness of the distribution is positive if p < 1/2 and negative if p > 1/2. The distribution is symmetric if and only if p = 1/2.

2.1.4. Relation to other distributions

If n is large and p is small, so that np is moderate, then Poisson distribution is a good approximation of the binomial distribution. That is,

$$f(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} \qquad \text{where } \lambda = np\text{, the binomial mean.}$$

As n becomes large with p fixed, the pivotal quantity $\frac{X - np}{\sqrt{np(1-p)}}$ (where X ~ Bin(n, p)), approaches a normal(0, 1) distribution.

2.1.5. Applications of binomial distribution.

The binomial distribution arises whenever the underlying events are independent and identical Bernoulli trials; in particular, it is the distribution of the sum of n such trials. The importance of the distribution has evolved from its original application in gaming to many other areas.

Its use in genetics arises because the inheritance of biological characteristics depends on genes that occur in pairs. A more recent application is the study of number of nucleotides that are in same state in two DNA sequences. A number of scientists have provided applications of the binomial distribution in plant and animal ecology.

The number of defectives found in a random sample of size n from a stable production process is a binomial variable. Acceptance sampling is a very important application of the test for the mean of a binomial sample against a hypothetical value. This distribution is the sampling distribution of the test statistic in both the sign test and McNemar's test. Although appealing in their simplicity, the assumptions of independence and constant success probability for the binomial distribution are rarely precisely satisfied. Nevertheless, the model is often sufficiently accurate to enable useful inferences to be made.

2.2. Confidence Intervals

By definition, a confidence interval $\theta_l \leq \theta \leq \theta_u$ for an unknown parameter θ, with unreliability α, comprises all values θ_0 for which the null hypothesis H_0: $\theta = \theta_0$ would not have been rejected in the observed sample when a two-sided test with unreliability α (i.e., the Type I error) would have been applied. Any value θ_0 smaller than the lower bound θ_l in the sample at hand is 'improbably small', and any $\theta_0 > \theta_u$ is 'improbably large'. Given some best estimate of θ in a given sample, two numbers θ_l and θ_u have to be calculated that meet the required property. Moreover, the interpretation of a confidence interval has to be understood in a frequentist sense, i.e., in a framework of repeatedly taking samples of size n from the same population distribution, and calculating an interval with confidence coefficient $1 - \alpha$ for some unknown but fixed parameter of interest in each of these samples. Confidence intervals are then constructed so that in the long run, the proportion of intervals covering the fixed population parameter equals $1-\alpha$. Therefore, in this framework of repeated sampling under the same conditions, a probability statement can be made, saying that the probability that the stochastic interval will cover the unknown fixed population value equals $1 - \alpha$. This probability statement refers to the proportion of the time that a random interval will contain the true value of the parameter (a population constant).

The lower and upper bounds of a two-sided confidence interval are random in the sense that they may change from sample to sample. In a given sample, however, they are known numbers. On the other hand, the population parameter θ is a fixed but unknown number. It is this contraposition 'stochastic but unknown' versus 'fixed but unknown' that makes the interpretation of a confidence interval so difficult, because of the mind's tendency to think that the unknown quantity θ has a probability distribution. But as long as the concept of probability refers to the frequentist point of view — what happens if the sampling experiment is repeated — that is incorrect thinking. Only in Bayesian statistics, not in classical statistics, can a parameter have a probability distribution.

The idea of constructing confidence intervals in the framework of repeated sampling is illustrated in Figure 2.

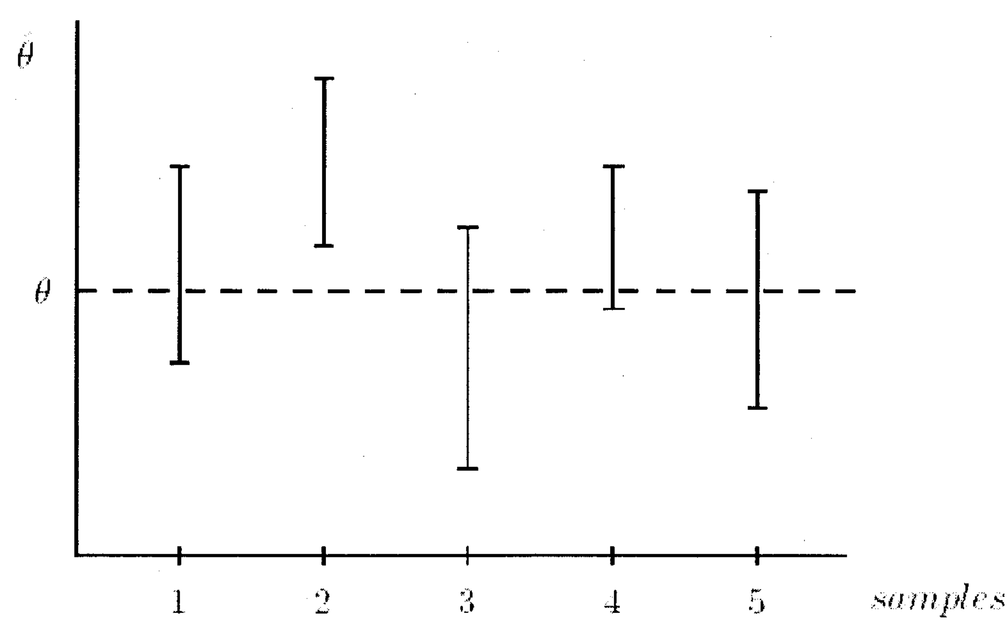


Figure 2.2. Confidence intervals for a fixed proportion parameter θ in 5 samples from the same population.

2.2.1. Wald Interval.

A common confidence interval procedure for estimating the binomial proportion p is the Wald interval given by the formula

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)^{th}$ percentile of the standard normal distribution (Wald and Wolfowitz, 1939). This procedure applies in large samples where the sampling distribution of $\hat{p}$ is reasonably approximated by a $N\left(p, \frac{p(1-p)}{n}\right)$ distribution.

This standard interval is easy to calculate and heuristically appealing because of its simplicity. In introductory statistics texts and courses, the confidence interval is usually presented along with some heuristic justification based on the central limit theorem. The larger the sample size n, the better the normal approximation, and thus the closer the actual coverage would be to the nominal level $1-\alpha$. The normal approximation used to justify the standard confidence interval for p has a significant error. The error is most evident when the true p is close to 0 or 1. In fact, it is easy to show that, for any fixed n, the binomial distributions converge to a degenerate distribution as $p \rightarrow 0$ or $p \rightarrow 1$. The length of the interval converges to 0.

Therefore, most major problems arise regarding coverage probability when p is near the boundaries. Poor coverage probabilities for p near 0 or 1 are widely remarked on, and generally, in the popular texts, a brief sentence is added qualifying when to use the Wald interval for p. For example, this confidence interval can be used if np and $np(1-p) \geq 5$ or 10; $n\hat{p}$ and $n\hat{p}(1-\hat{p}) \geq 5$ or 10, etc. Figures 3 and 4 show the coverage probabilities for n = 5 and 25 and p = 0 to 1, for the Wald interval at the nominal level of 95%.

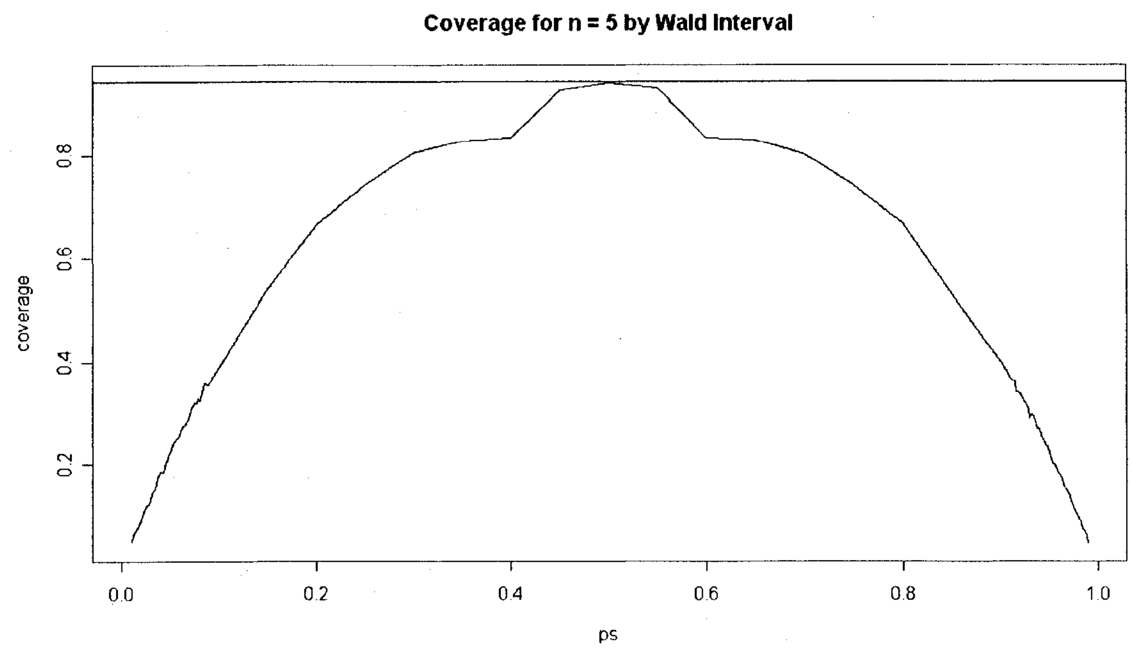


Figure 2.3. Coverage probability for n = 5 and p = 0 to 1 at 95% nominal level for the Wald interval.

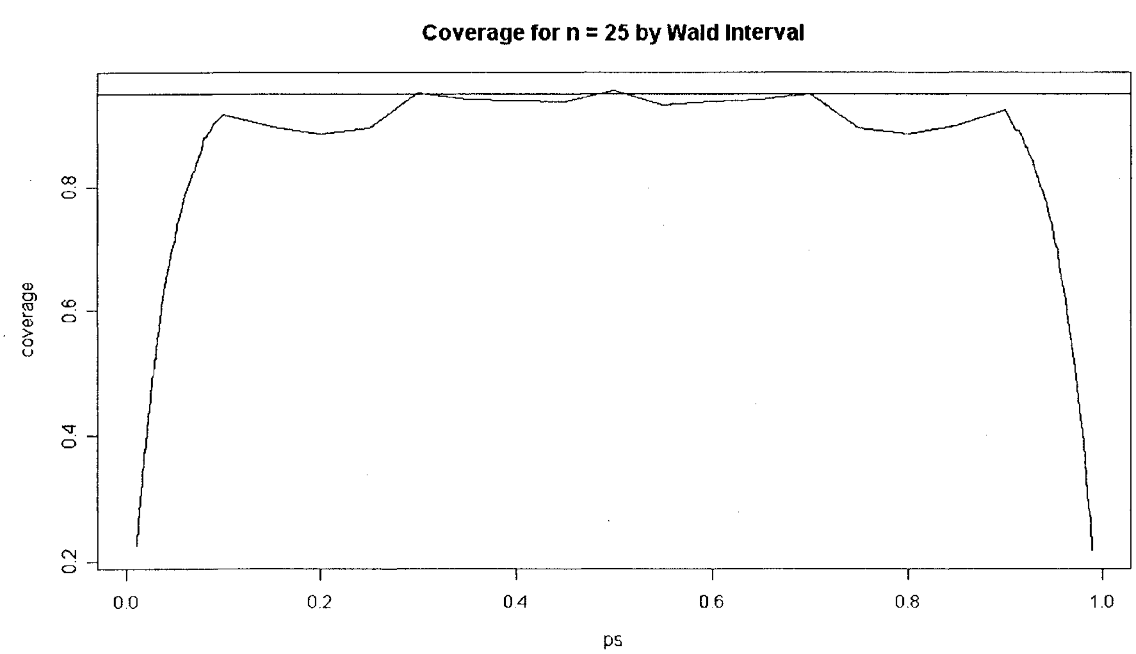


Figure 2.4. Coverage probability for n = 25 and p = 0 to 1 at 95% nominal level for the Wald interval.

There are oscillations in the coverage in both the figures. These oscillations are due to the discreteness of the binomial distribution. Another reason is the presence of systematic bias in the coverage probability of the confidence interval. The bias is due mainly to the fact that the standard interval has the "wrong" center. The standard interval is centered at $\hat{p} = \frac{x}{n}$. Although $\hat{p}$ is the MLE and an unbiased estimate of p, as the center of a confidence interval it causes a systematic negative bias in the coverage. The standard interval is based on the fact that

$$W_n = \sqrt{\frac{n}{\hat{p}(1-\hat{p})}}(\hat{p} - p) \to N(0,1)$$

However, even for quite large values of n, the actual distribution of W_n may be significantly non-normal. Thus the very premise on which the standard interval is based is seriously flawed for moderate values of n.

A discussion on the coverages for some sample values of n and the confidence intervals is given in chapter 4.

2.2.2. Exact Interval

To avoid the approximation of the Wald interval, textbooks recommend the Clopper-Pearson "exact" confidence interval for p, based on inverting equal-tailed binomial tests of $H_0 : p = p_0$. Its endpoints that are the solutions p_u and p_l to the simultaneous equations

$$\sum_{k=x}^{n} \binom{n}{k} p^k (1-p)^{n-k} = \alpha/2$$

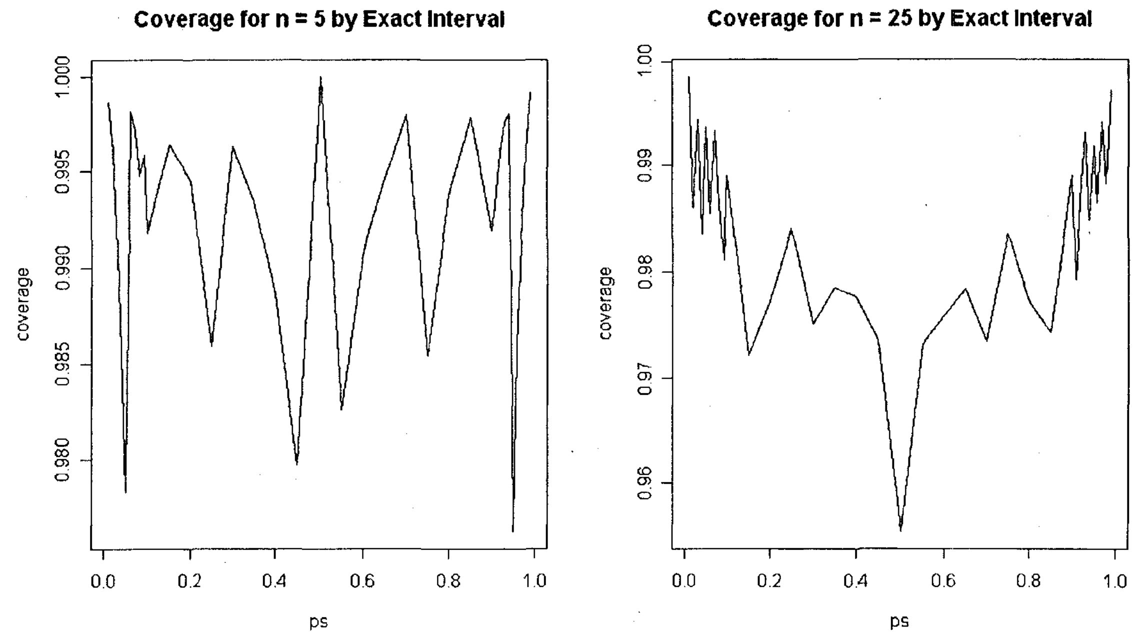
and

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = \alpha/2$$

except that the lower bound is 0 when x = 0 and upper bound is 1 when x = n. The interval estimator is guaranteed to have coverage probability of at least 1 – α for every possible value of p. When x = 1, 2, 3, …., n – 1, the confidence interval equals

$$\left[1+\frac{n-x+1}{xF_{2x,2(n-x+1),1-\alpha/2}}\right]^{-1} < p < \left[1+\frac{n-x}{(x+1)F_{2(x+1),2(n-x),\alpha/2}}\right]^{-1}$$

where $F_{a,b,c}$ denotes the 1-c quantile from the F-distribution with degrees of freedom a and b. Equivalently, the lower endpoint is the α/2 quantile of a beta distributionwith parameters x and n – x + 1, and the upper endpoint is the 1 – α/2 quantile of a beta distribution with parameters x + 1 and n – x. This exact interval is typically treated as the gold standard. However the procedure is necessarily conservative, because of the discreteness of the binomial distribution. This means that for any fixed parameter value, the actual coverage probability can be much larger than the nominal confidence level unless n is quite large. This is shown in both the plots of figure 2.5. The lower dotted line is the 95% coverage mark and the upper dotted line is the 100% coverage mark. As is seen here the coverage probability is higher than the 95% nominal level.



Figures 2.5. Coverage probability for n = 5 and 25 and p = 0 to 1 at 95% nominal level for the Exact interval.

CHAPTER 3

BAYESIAN ESTIMATION OF CREDIBLE SETS

The classical statistical approach considers a parameter as a fixed, but unknown constant to be estimated using data randomly sampled from the population of interest. A confidence interval for an unknown parameter is really a frequency statement about the likelihood that numbers calculated from a sample capture the true parameter value in repeated sampling. So the classical statistical approach cannot say there is a 95% probability that the true proportion is in any single interval, because it is either already in, or it is not. This is because under the classical approach, the true proportion is a fixed unknown constant, so it does not have a distribution; however, the sample proportion $\hat{p}$ does. Thus, we can say that there is a 95% chance the random interval contains p, in repeated samples of size m from the same population.

The Bayesian approach, on the other hand, treats the population model parameter as random instead of fixed. Actually, the data are treated as fixed realizations of a random process, accounted for by the likelihood function. Before looking at the current data, we use past information to construct a prior distribution model for the parameter. The prior distribution is chosen to reflect one's prior knowledge of p, which may vary from one person to the next. As a result, the mathematical form of a prior distribution is quite flexible. In particular, conjugate priors are a natural and popular choice of Bayesian prior distribution model, due to their mathematical convenience. The prior distribution of a parameter may be noninformative or informative. Noninformative priors are locally uniform in a certain range of parameter values. The range of possible values may be fixed or may be infinite. An informative prior distribution specifies a particular nonuniform shape for the distribution of the parameter. When new data are gathered, they are used to update the prior distribution. We then take the weighted average of the prior and data, expressed through the likelihood function, to derive what is called the posterior distribution model for the population model parameter. Point estimates, along with interval estimates (known as credibility intervals), are calculated directly from the posterior distribution. Credibility intervals are legitimate probability statements about the unknown parameter, since the parameter now is considered random. Under the Bayesian point of view, we can say that there is a 95% probability that the interval contains the population proportion.

The posterior distribution model is based on Bayes' theorem, which expresses the conditional probability of an event A, given that the event B has occurred, in terms of unconditional probabilities and the probability the event B has occurred, given that A has occurred. It is defined as

$$P(A \mid B) = \frac{P(A,B)}{P(B)} = \frac{P(A) \times P(B \mid A)}{P(B)}$$

In terms of probability density functions, the theorem takes the form

$$g(p \mid x) = \frac{f(x \mid p)g(p)}{\int_0^1 f(x \mid p)g(p)dp}$$

This is known as the posterior density of x, where f(x|p) is the likelihood function of the observed data x given the unknown parameter p, g(p) is the prior density of p and the denominator represents the marginal density of x. When $g(p \mid x)$ and g(p) both belong to the same family of distribution, g(p) and f(x|p) are called conjugate distributions and g(p) is called the conjugate prior for f(x|p). For example, the Beta distribution model is a conjugate prior for the proportion of successes p when samples have a binomial distribution.

With probability $1-\alpha$, a Bayesian credibility interval for p is given by (p_L, p_U), where p_L and p_U satisfy

$$\int_{p_L}^{p_U} g(p \mid x)dp = 1-\alpha.$$

This yields an interval estimate of p with probability $1-\alpha$.

In this chapter, a detailed overview of the estimation of the credible sets using uniform and beta priors is presented

3.1. Using Uniform priors

If there is no idea beforehand as to what the proportion p is, it is reasonable to choose a prior that does not favor any one value over the other. The idea is to be as objective as possible. In such a case a uniform prior that gives equal weight to all possible values of success probability p is to be used. Hence a reasonable prior is p ~ Unif(0,1).

The density function for this prior is given by

$$g(p) = \begin{cases} 1 & \text{where } 0 < p < 1, \\ 0 & \text{elsewhere.} \end{cases}$$

Let x ~ Bin(n, p) and p ~ Unif(0,1). Using Bayes' formula, the posterior density function is given by

$$g(p \mid x) = \frac{f(x \mid p)g(p)}{\int_0^1 f(x \mid p)g(p)dp}$$

that is

$$g(p|x)=\frac{p^x(1-p)^{n-x}\times 1}{\int_0^1\left(p^x(1-p)^{n-x}\times 1\right)dp}$$

Notice that the denominator is the normalizing constant and the product in the numerator is the kernel of the beta distribution.

The pdf of the beta distribution is:

$$\mathrm{f}(x|\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

where $0<\mathrm{x}<1$, $\alpha,\beta>0$ and $\Gamma(\mathrm{k})=(\mathrm{k}-1)!$ is the Gamma function when $\mathrm{k}\geq 1$ is integer-valued.

If we let x = p, α = x + 1 and β = n - x + 1, we get the following expression:

$$g(p|x)=\frac{\mathrm{p}^x(1-p)^{n-x}}{\left(\frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)}\right)\times\int_0^1\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\times\mathrm{p}^x(1-p)^{n-x}\,dp}$$

Since $\int_0^1\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\times\mathrm{p}^x(1-p)^{n-x}\,dp$ is the integral of the beta pdf over the parameter space for p. This expression is equal to one. Thus, after simplification we have

$$g(p|x)=\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\times p^x(1-p)^{n-x},$$

which corresponds to a Beta(x + 1, n - x + 1) density.

It worked out that the posterior distribution is a form of a beta distribution. The Bayes estimator of the proportion in the population p, under squared error loss, is just the posterior mean. If Y ~ Beta(α, β), then the mean of a beta distribution is

$$E(Y) = \frac{\alpha}{\alpha+\beta}.$$

Therefore, the Bayes estimator of p is

$$\hat{p}_{Bayes} = E(p \mid x) \approx \frac{x+1}{n+2}.$$

The credibility interval for the parameter p is then computed from that posterior beta distribution with parameters x + 1, n - x + 1.

This uniform prior is just one of an infinite number of possible prior distributions. Since the point of interest in this thesis, "p" is near 0 or 1, p is taken from Unif(0. 0.2) or Unif(0.8, 1). The simulations and the results are shown in the next chapter.

3.2. Using Beta Priors

When there is a prior knowledge as to the proportion p, a beta prior is a useful alternative. For a random variable p, where p ~ Beta(α, β), the pdf is

$$g(p \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

In fact, we can show that the Beta(1, 1) distribution is the Unif(0,1) distribution; that is

$$g(p \mid 1,1) = \frac{\Gamma(1+1)}{\Gamma(1)\Gamma(1)} p^{1-1}(1-p)^{1-1} \qquad 0 < p < 1,$$

which gives us

$$Beta(1,1) = \frac{\Gamma(1+1)}{\Gamma(1)\Gamma(1)} = 1 \qquad 0 < p < 1$$

This is the density function for the Uniform (0,1) distribution.

Figure 3.1 shows the plots of two beta pdfs. The first one is Beta(10, 100) and the second one is Beta(100, 10). This corresponds to the knowledge that the binomial probability comes from a known prior and has the value close to 0 (first plot) or 1 (second plot).

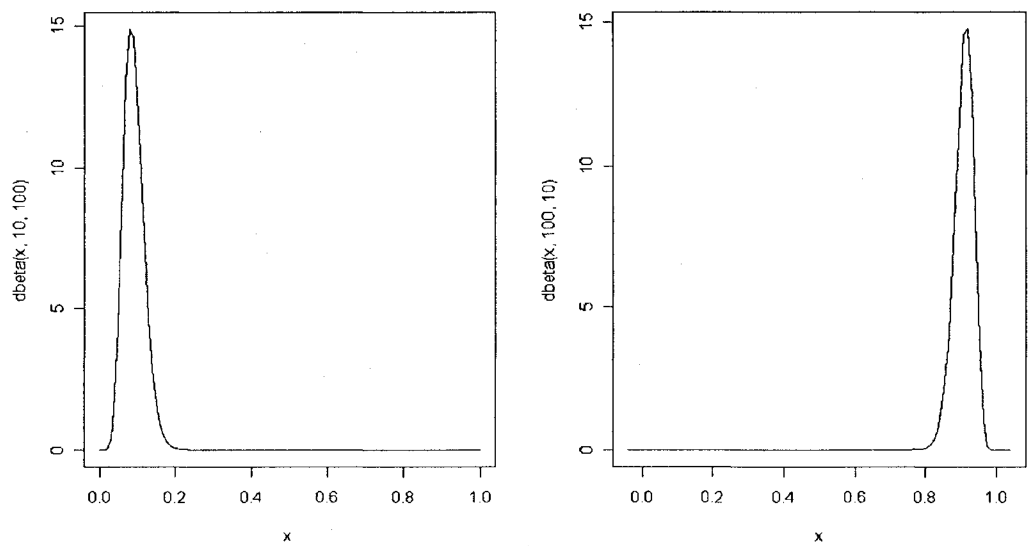


Figure 3.1. Prior distributions of Beta(10, 100) and Beta(100, 10)

Let x ~ Bin(n, p), as we defined in the previous section, and let $g(p) \sim \text{Beta}(\alpha,\beta)$; that is, the proportion p has beta prior distribution. Then,

$$g(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

where $0 < p < 1$, $\alpha > 0$, $\beta > 0$ and α,β are known arbitrary constants. The posterior density is given by

$$g(p \mid x) = \frac{f(x \mid p)g(p)}{\int_0^1 f(x \mid p)g(p)dp}$$

where $$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x}.$$

We then have

$$g(p|x)=\frac{\binom{n}{x}p^{x}(1-p)^{n-x}\times\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}}{\int_0^1\left(\binom{n}{x}p^{x}(1-p)^{n-x}\times\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}\right)dp}$$

$$=\frac{p^{\alpha+x-1}(1-p)^{n+\beta-x-1}}{\int_0^1\left(p^{\alpha+x-1}(1-p)^{n+\beta-x-1}\right)dp}$$

The denominator is the normalizing constant of a beta distribution. We then have the following expression,

$$g(p|x)=\frac{p^{x+\alpha-1}(1-p)^{n+\beta-x-1}}{\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(x+\alpha+n-x+\beta)}\int_0^1\frac{\Gamma(x+\alpha+n-x+\beta)}{\Gamma(x+\alpha)\Gamma(n-x+\beta)}\times\left(p^{x+\alpha-1}(1-p)^{n+\beta-x-1}\right)dp}$$

Since $\int_0^1\frac{\Gamma(x+\alpha+n-x+\beta)}{\Gamma(x+\alpha)\Gamma(n-x+\beta)}\times\left(p^{x+\alpha-1}(1-p)^{n+\beta-x-1}\right)dp$ is the integral of the beta pdf over the parameter space for p, this expression equals one. The posterior density is then

$$g(p|x)=\frac{\Gamma(x+\alpha+n-x+\beta)}{\Gamma(x+\alpha)\Gamma(n-x+\beta)}\times\left(p^{x+\alpha-1}(1-p)^{n+\beta-x-1}\right),$$

which is a Beta $(x+\alpha,\ n-x+\beta)$ density.

It worked out that the posterior distribution is a form of the prior distribution updated by the new data. In general, when this occurs we say the prior is conjugate.

The Bayes estimator, under squared error loss, is the mean of the posterior Beta $(x+\alpha,\ n-x+\beta)$ distribution.

$$\hat{p}_{Bayes}=E(p|X)=\frac{(x+\alpha)}{(n+\alpha+\beta)}\ .$$

The Bayes credible set for p is given by (p_l, p_u) where p_l, and p_u satisfy

$$\int_{p_l}^{p_u} g(p \mid X)dp = 1 - \alpha ,$$

where α is the significance level. The credible interval used in this thesis is computed from the posterior beta cumulative density function as

$$P\left[Beta_{\alpha/2}\left(x+\alpha, n-x+\beta\right) < p < Beta_{1-\alpha/2}\left(x+\alpha, n-x+\beta\right)\right] = 1-\alpha .$$

Then a 1 - α Bayesian credible interval for p is (p_l, p_u) where p_l, the α/2 quantile and p_u is the 1 - α/2 quantile of the beta distribution with parameters x + α, and n - x + β. Chapter 4 shows the simulations and results for this analysis.

CHAPTER 4

SIMULATIONS AND RESULTS OF EXPERIMENTS

In this chapter, results and simulations of the confidence intervals and the credible sets are discussed. The first section deals with the Wald and the Exact interval and the second section deals with the Bayesian credible sets.

4.1. The Wald Interval and the Exact Interval

95% confidence intervals were constructed for n = 5, 25, 50, 75, 100 and 1000 using the statistical software package R for the computations.

10,000 simulations were carried out to evaluate the coverage probabilities for both intervals at various values of n and p. The values of p used in this simulation are 0 to 0.1 (in increments of 0.01), 0.15, 0.2, 0.8, 0.85, and 0.9 to 1 (in increments of 0.01). Table 4.1 gives a sample of the output obtained for the Wald interval, with p = 0.05 and n = equal to 25. From left to right, the columns denote simulation number, the lower limit, the upper limit, the length of the confidence interval and whether or not p is captured by the confidence interval (1 if yes and 0 if it is not). Table 4.2 shows the same for the Exact Interval.

Table 4.1.Sample of the program output for Wald interval at n = 25 and p = 0.05

No.	Lower Limit	Upper Limit	Length	Hits
1	0	0	0	0
2	-0.036816	0.116816	0.153632	1
3	-0.0263469	0.1863469	0.212694	1
4	-0.0263469	0.1863469	0.212694	1
5	0	0	0	0
6	-0.036816	0.116816	0.153632	1
7	-0.036816	0.116816	0.153632	1
8	0	0	0	0
9	-0.036816	0.116816	0.153632	1
10	0	0	0	0
11	-0.0073849	0.2473849	0.25477	1
12	-0.036816	0.116816	0.153632	1
13	-0.036816	0.116816	0.153632	1
14	-0.036816	0.116816	0.153632	1
15	-0.036816	0.116816	0.153632	1
16	-0.036816	0.116816	0.153632	1
17	-0.036816	0.116816	0.153632	1
18	-0.036816	0.116816	0.153632	1
19	-0.0263469	0.1863469	0.212694	1
20	-0.0263469	0.1863469	0.212694	1
21	-0.036816	0.116816	0.153632	1
22	-0.0073849	0.2473849	0.25477	1
23	-0.036816	0.116816	0.153632	1
24	-0.036816	0.116816	0.153632	1
25	-0.0263469	0.1863469	0.212694	1
26	-0.036816	0.116816	0.153632	1
27	-0.036816	0.116816	0.153632	1
28	-0.036816	0.116816	0.153632	1
29	-0.0263469	0.1863469	0.212694	1
30	0	0	0	0
31	0	0	0	0
32	-0.036816	0.116816	0.153632	1
33	-0.0263469	0.1863469	0.212694	1
34	-0.036816	0.116816	0.153632	1
35	0	0	0	0
36	-0.036816	0.116816	0.153632	1
37	-0.036816	0.116816	0.153632	1
38	-0.036816	0.116816	0.153632	1
39	-0.036816	0.116816	0.153632	1
40	0.0162904	0.3037096	0.287419	1
41	-0.036816	0.116816	0.153632	1
42	-0.0263469	0.1863469	0.212694	1
43	0	0	0	0

Table 4.2.Sample of the program output for Exact interval at n = 25 and p = 0.05

No.	Lower Limit	Upper Limit	Length	Hits
1	0	0.1371852	0.137185	1
2	0.0098396	0.2603058	0.250466	1
3	0.0010122	0.2035169	0.202505	1
4	0.0098396	0.2603058	0.250466	1
5	0.0098396	0.2603058	0.250466	1
6	0.0010122	0.2035169	0.202505	1
7	0.0010122	0.2035169	0.202505	1
8	0	0.1371852	0.137185	1
9	0	0.1371852	0.137185	1
10	0.0010122	0.2035169	0.202505	1
11	0.0010122	0.2035169	0.202505	1
12	0.0254654	0.3121903	0.286725	1
13	0.0254654	0.3121903	0.286725	1
14	0.0098396	0.2603058	0.250466	1
15	0.0254654	0.3121903	0.286725	1
16	0	0.1371852	0.137185	1
17	0.0010122	0.2035169	0.202505	1
18	0	0.1371852	0.137185	1
19	0	0.1371852	0.137185	1
20	0.0010122	0.2035169	0.202505	1
21	0.0010122	0.2035169	0.202505	1
22	0.0010122	0.2035169	0.202505	1
23	0	0.1371852	0.137185	1
24	0	0.1371852	0.137185	1
25	0.0098396	0.2603058	0.250466	1
26	0.0254654	0.3121903	0.286725	1
27	0.0010122	0.2035169	0.202505	1
28	0.0254654	0.3121903	0.286725	1
29	0	0.1371852	0.137185	1
30	0.0098396	0.2603058	0.250466	1
31	0.0010122	0.2035169	0.202505	1
32	0.0010122	0.2035169	0.202505	1
33	0	0.1371852	0.137185	1
34	0.0098396	0.2603058	0.250466	1
35	0	0.1371852	0.137185	1
36	0	0.1371852	0.137185	1
37	0	0.1371852	0.137185	1
38	0.0098396	0.2603058	0.250466	1
39	0.0098396	0.2603058	0.250466	1
40	0.0010122	0.2035169	0.202505	1
41	0.0098396	0.2603058	0.250466	1
42	0.0254654	0.3121903	0.286725	1
43	0.0683115	0.4070374	0.338726	0

The lower limit of the Wald confidence interval when the value of p =0.5 (or when close to 0) has negative values. This is because of the normal approximation. Similarly for values of p near 1, the upper confidence limits tend to exceed 1. Tables 4.3 to 4.5 show the comparison of the coverage probabilities of the Exact interval and the Wald interval.

Table 4.3. Comparison of coverage probabilities of the two intervals, n = 5 and 25

p (n=5)	Exact	Wald
0.01	0.9991	0.0505
0.02	0.9961	0.0979
0.03	0.991	0.1397
0.04	0.9868	0.1848
0.05	0.9963	0.2212
0.06	0.9979	0.2659
0.07	0.9974	0.2965
0.08	0.9959	0.3342
0.09	0.9933	0.3716
0.1	0.9928	0.3998
0.15	0.998	0.5382
0.2	0.9938	0.6732
0.8	0.9924	0.6628
0.85	0.9977	0.5293
0.9	0.9889	0.4001
0.91	0.9945	0.375
0.92	0.996	0.3342
0.93	0.9975	0.2989
0.94	0.9979	0.2642
0.95	0.9765	0.2275
0.96	0.986	0.188
0.97	0.993	0.1438
0.98	0.9958	0.0937
0.99	0.9986	0.0509

p (n=25)	Exact	Wald
0.01	0.9978	0.2154
0.02	0.9853	0.3968
0.03	0.9941	0.5356
0.04	0.9824	0.6378
0.05	0.9931	0.7266
0.06	0.9842	0.7891
0.07	0.9931	0.8271
0.08	0.9876	0.8685
0.09	0.9781	0.9035
0.1	0.9903	0.9182
0.15	0.9738	0.8991
0.2	0.9807	0.8853
0.8	0.9801	0.8852
0.85	0.974	0.901
0.9	0.9905	0.9207
0.91	0.9788	0.8962
0.92	0.9878	0.8794
0.93	0.9937	0.8289
0.94	0.9846	0.7831
0.95	0.9934	0.7267
0.96	0.9922	0.6445
0.97	0.9944	0.5292
0.98	0.9873	0.3875
0.99	0.999	0.2233

Table 4.4. Comparison of the coverage probabilities of the two intervals for n = 50 and 75

p(n=50)	Exact	Wald
0.01	0.9873	0.3879
0.02	0.9819	0.6343
0.03	0.9835	0.7884
0.04	0.9866	0.8666
0.05	0.9874	0.9133
0.06	0.9912	0.8103
0.07	0.9769	8.6156
0.08	0.9699	0.9109
0.09	0.9786	0.9423
0.1	0.9706	0.8762
0.15	0.9725	0.9371
0.2	0.9675	0.9362
0.8	0.964	0.9394
0.85	0.9704	0.9381
0.9	0.9711	0.878
0.91	0.9771	0.9436
0.92	0.9708	0.9093
0.93	0.9794	0.8651
0.94	0.9907	0.8028
0.95	0.9894	0.9178
0.96	0.9851	0.8651
0.97	0.9804	0.7788
0.98	0.9831	0.6267
0.99	0.984	0.3861

p(n=75)	Exact	Wald
0.01	0.9924	0.5286
0.02	0.9812	0.7794
0.03	0.9919	0.8935
0.04	0.989	0.8033
0.05	0.9649	0.8896
0.06	0.9746	0.9426
0.07	0.9814	0.8953
0.08	0.9723	0.9384
0.09	0.9767	0.9073
0.1	0.9689	0.9453
0.15	0.9642	0.9317
0.2	0.9578	0.9286
0.8	0.957	0.9319
0.85	0.9676	9.3223
0.9	0.9677	0.9443
0.91	0.9776	0.9089
0.92	0.9665	0.9401
0.93	0.9799	0.8979
0.94	0.9778	0.9411
0.95	0.9657	0.8951
0.96	0.9912	0.8031
0.97	0.9932	0.9003
0.98	0.9815	0.7785
0.99	0.9929	0.5319

Table 4.5. Comparison of the coverage probabilities of the two intervals for n = 100 and 1000

p(n=100)	Exact	Wald
0.01	0.9802	0.6316
0.02	0.9844	0.8671
0.03	0.9889	0.8041
0.04	0.9639	0.7073
0.05	0.9822	0.8767
0.06	0.9655	0.9346
0.07	0.9711	0.9137
0.08	0.9719	0.8965
0.09	0.9652	0.9457
0.1	0.9548	0.9341
0.15	0.9653	0.9333
0.2	0.9656	0.9321
0.8	0.9667	0.9343
0.85	0.9627	0.9324
0.9	0.9569	0.9299
0.91	0.9671	0.9431
0.92	0.9751	0.8988
0.93	0.973	0.9144
0.94	0.9681	0.9399
0.95	0.9822	0.8794
0.96	0.9654	0.9054
0.97	0.9899	0.7991
0.98	0.9844	0.8678
0.99	0.9806	0.6331

p(n=1000)	Exact	Wald
0.01	0.9773	0.9323
0.02	0.9574	0.9484
0.03	0.9582	0.9318
0.04	0.9582	0.9491
0.05	0.96	0.9425
0.06	0.9558	0.9511
0.07	0.9501	0.9474
0.08	0.9497	0.9484
0.09	0.9488	0.948
0.1	0.9554	0.9532
0.15	0.9495	0.9486
0.2	0.9537	0.948
0.8	0.9484	0.9463
0.85	0.9515	0.9449
0.9	0.9506	0.9531
0.91	0.9534	0.9491
0.92	0.9521	0.953
0.93	0.9508	0.9508
0.94	0.953	0.9488
0.95	0.9546	0.9407
0.96	0.9656	0.9471
0.97	0.957	0.9366
0.98	0.9582	0.9485
0.99	0.9789	0.9223

The above tables show that the coverage probability for the Exact interval is above the 95% nominal level. For small values of n, the coverage probability for this interval is very high at the extreme values of p. As p moves away from the extremes, the coverage probability approaches 0.95. As n increases, the coverage probability tends towards the nominal level and at very high values of n (n = 1000), the coverage probability almost equals the nominal level. But the coverage probability always remains higher than 0.95 because of the discreteness of the binomial distribution which results in this interval being conservative.

The coverage probability for the Wald interval, for low values of n, is very low at the extremes; moreover it is uniformly less than 0.95 for n = 5. As n increases, the coverage probability increases to 0.95 for each p and equals to 0.95 at very high values of n (n=1000). But the coverage probability of this interval is always less than the nominal level for the values of p near 0 or 1. For example, it is a mere 0.4 for n = 25 and p = 0.9, 0.88 at n = 50 and p = 0.9 and for very low value of p (0.01) and n (5), it is 0.05. Even at n = 1000 and p = 0.99 the coverage probability is less than 0.95 (0.92).

Graphs showing the coverage probability for both intervals at different values of n are shown below in Figures 4.1- 4.5.

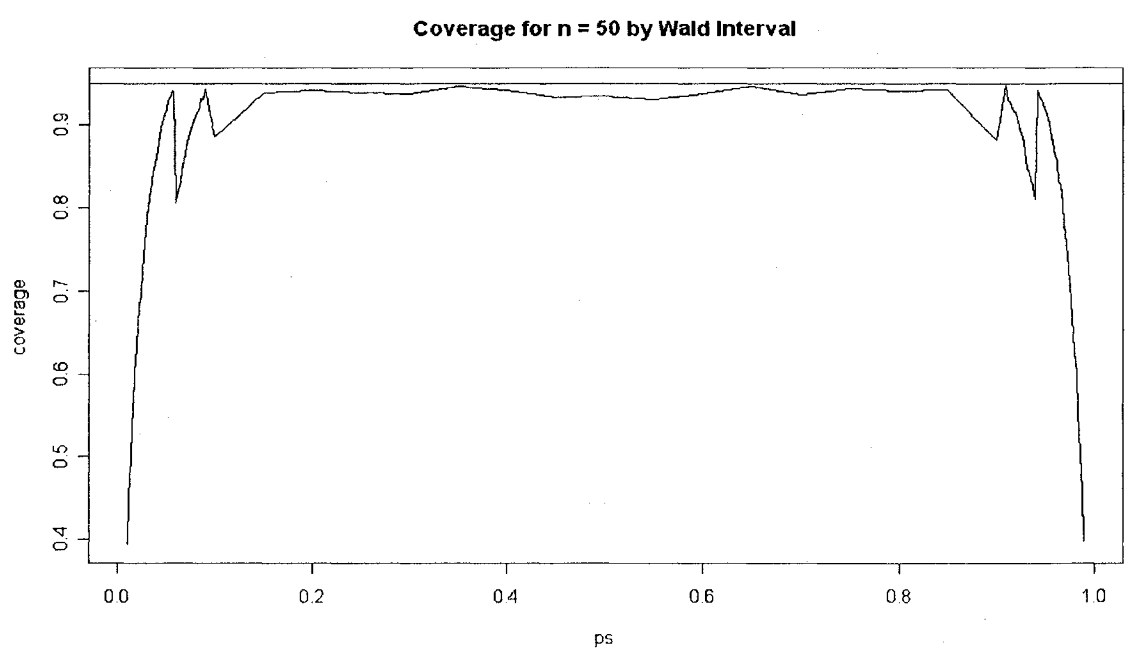


Figure 4.1. Coverage probability for n = 50 and p = 0 to 1 at 95% nominal level for the Wald interval.

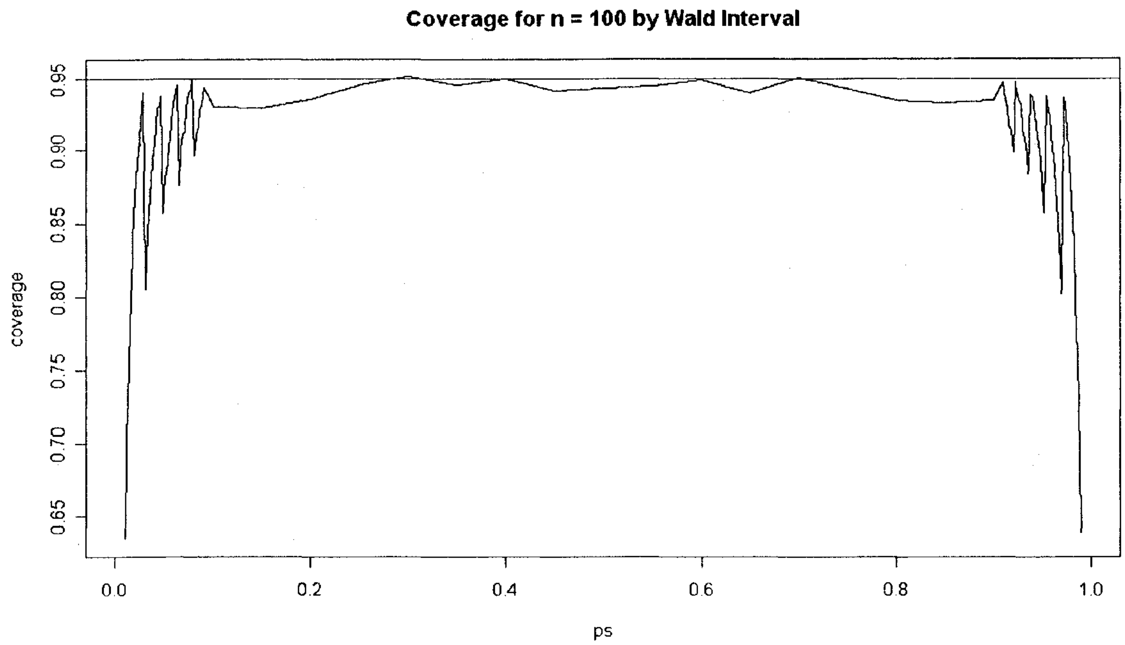


Figure 4.2. Coverage probability for n = 100 and p = 0 to 1 at 95% nominal level for the Wald interval.

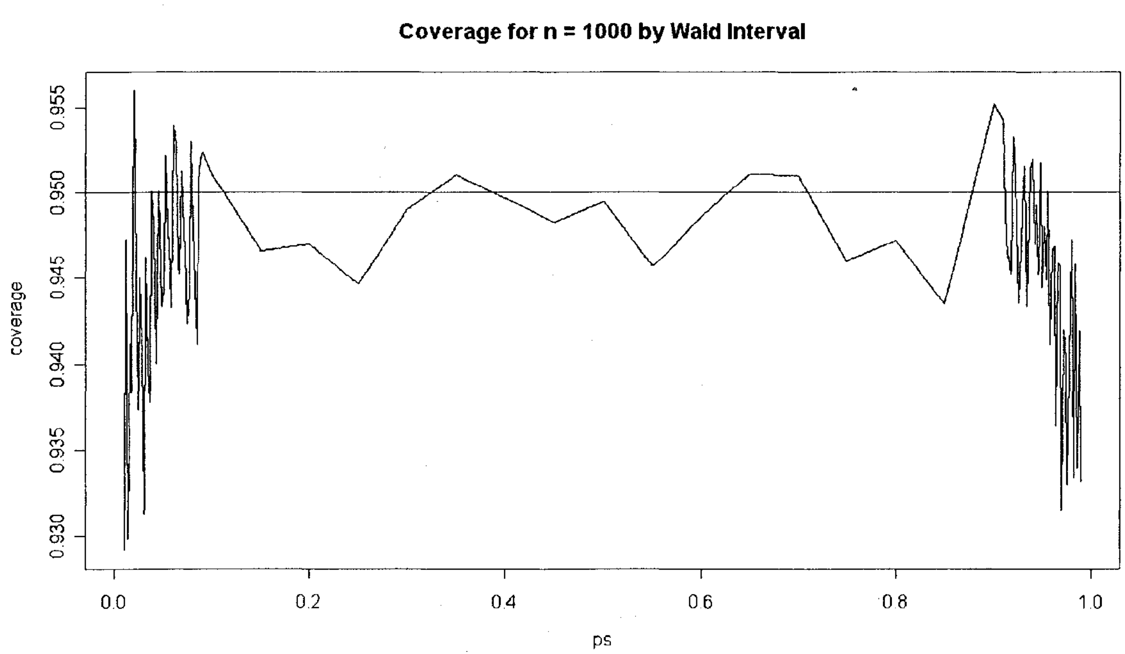


Figure 4.3. Coverage probability for n = 1000 and p = 0 to 1 at 95% nominal level for the Wald interval.

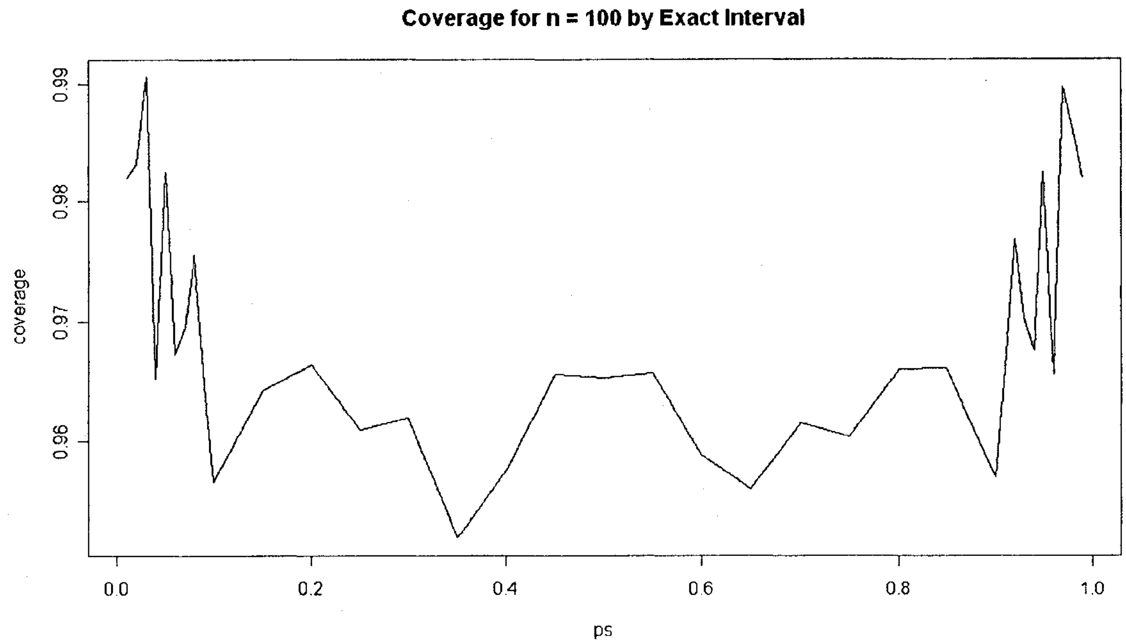


Figure 4.4. Coverage probability for n = 100 and p = 0 to 1 at 95% nominal level for the Exact interval.

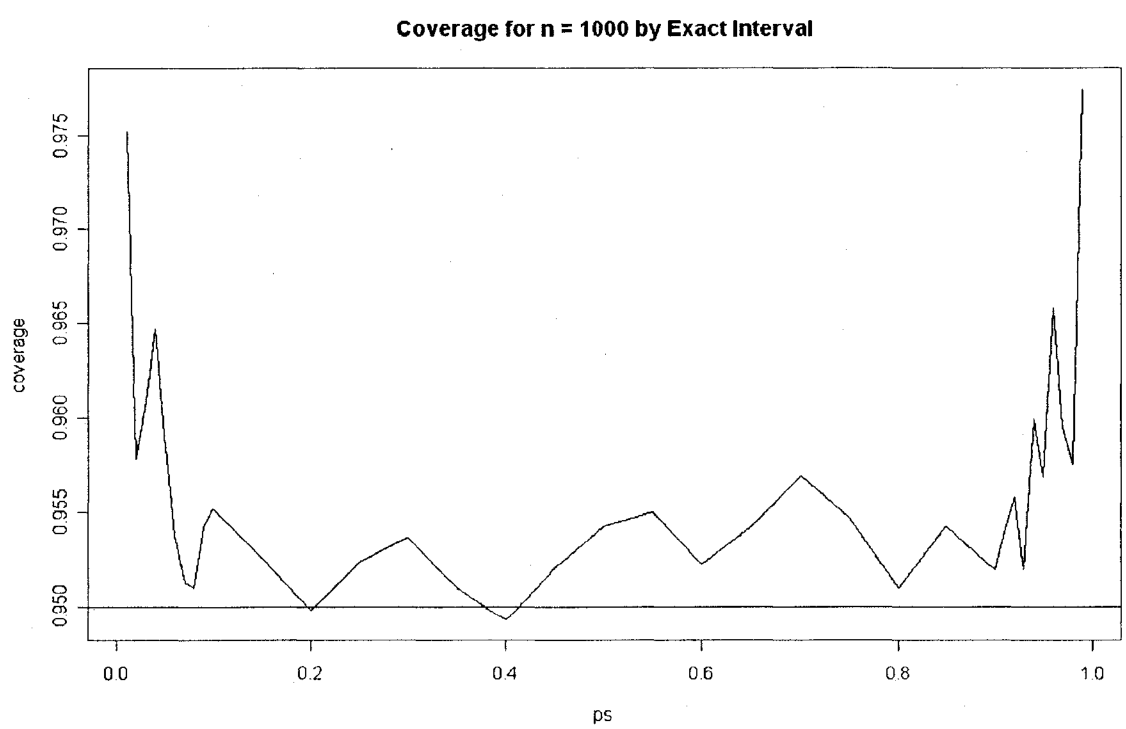


Figure 4.5. Coverage probability for n = 1000 and p = 0 to 1 at 95% nominal level for the Exact interval.

The graph of coverage probabilities for n = 5 and n = 25 for the Exact interval is given in page 16.

The Tables 4.6, 4.7 and 4.8 show the mean and the standard deviation of the lengths (upper limit - lower limit) of the confidence interval for both the Exact and the Wald Intervals.

Table 4.6. Mean and Standard deviations of the interval lengths for n = 5 and n = 25

n=5	Exact		Wald	
p	Mean	Std. dev.	Mean	Std. dev.
0.01	0.5316	0.0422	0.0355	0.15408
0.02	0.5406	0.0573	0.0696	0.21136
0.03	0.5508	0.0698	0.0994	0.24638
0.04	0.5572	0.076	0.1326	0.27843
0.05	0.5666	0.0841	0.1592	0.2988
0.06	0.5739	0.0889	0.1936	0.32076
0.07	0.5839	0.0947	0.2165	0.3328
0.08	0.5914	0.0982	0.2454	0.3448
0.09	0.5992	0.1018	0.2754	0.3551
0.1	0.6064	0.104	0.2993	0.3622
0.15	0.6411	0.1116	0.4217	0.3748
0.2	0.6712	0.1107	0.5173	0.3611
0.8	0.6735	0.1104	0.5088	0.3626
0.85	0.6404	0.1114	0.6416	0.3751
0.9	0.6067	0.0105	0.2991	0.3621
0.91	0.5993	0.1019	0.2778	0.35622
0.92	0.5902	0.098	0.246	0.34535
0.93	0.5836	0.0946	0.2181	0.33286
0.94	0.5742	0.0894	0.1917	0.3193
0.95	0.5672	0.0844	0.1636	0.3014
0.96	0.5574	0.0763	0.1347	0.2803
0.97	0.5491	0.0671	0.1026	0.2502
0.98	0.5406	0.0575	0.0662	0.20615
0.99	0.5311	0.0413	0.0358	0.15503

n=25	Exact		Wald	
p	Mean	Std. dev.	Mean	Std. dev.
0.01	0.1528	0.0304	0.0346	0.0668
0.02	0.1691	0.0412	0.0667	0.0841
0.03	0.1821	0.0464	0.0946	0.0916
0.04	0.1957	0.0506	0.118	0.0939
0.05	0.2074	0.0533	0.1395	0.0929
0.06	0.2189	0.0551	0.1591	0.0914
0.07	0.2291	0.0559	0.173	0.0892
0.08	0.2399	0.056	0.1292	0.087
0.09	0.2499	0.056	0.2034	0.0817
0.1	0.2592	0.0561	0.2164	0.0793
0.15	0.2984	0.0517	0.2664	0.0647
0.2	0.3298	0.0447	0.3026	0.0525
0.8	0.3306	0.0447	0.3024	0.0526
0.85	0.2932	0.0512	0.267	0.0638
0.9	0.2592	0.0557	0.2164	0.079
0.91	0.2502	0.0567	0.2032	0.0838
0.92	0.2414	0.056	0.1913	0.0849
0.93	0.2304	0.0556	0.174	0.0893
0.94	0.2175	0.0547	0.1574	0.0918
0.95	0.2078	0.0529	0.1392	0.0928
0.96	0.195	0.0508	0.1194	0.0937
0.97	0.1814	0.0463	0.0939	0.092
0.98	0.1685	0.0408	0.065	0.0836
0.99	0.153	0.0306	0.0358	0.0674

It is evident from the above tables that even though the Exact interval gives higher coverage probabilities, the confidence intervals have a wider length but a smaller standard deviation. The Wald interval lengths are smaller, i.e. the Wald interval gives a tighter confidence interval but a larger standard deviation when compared to the Exact interval.

Table 4.7. Mean and Standard deviations of the interval lengths for n = 50 and n = 75

n=50	Exact		Wald	
p	Mean	Std. dev.	Mean	Std. dev.
0.01	0.0876	0.0219	0.0333	0.0428
0.02	0.1027	0.0277	0.0597	0.0484
0.03	0.1153	0.0305	0.0812	0.0476
0.04	0.1267	0.0316	0.9764	0.0461
0.05	0.1378	0.0328	0.1107	0.0443
0.06	0.1402	0.0325	0.1238	0.0403
0.07	0.157	0.0313	0.1346	0.0395
0.08	0.1651	0.0315	0.1442	0.0374
0.09	0.1731	0.0315	0.1528	0.0356
0.1	0.1802	0.0305	0.161	0.0344
0.15	0.2104	0.0268	0.1944	0.0288
0.2	0.2336	0.023	0.218	0.0248
0.8	0.2338	0.023	0.2185	0.0245
0.85	0.2105	0.0272	0.1937	0.0291
0.9	0.1802	0.0305	0.161	0.0341
0.91	0.1732	0.0312	0.1529	0.0353
0.92	0.165	0.0317	0.1439	0.0373
0.93	0.1566	0.0319	0.1343	0.0385
0.94	0.148	0.0325	0.1235	0.0411
0.95	0.1379	0.0322	0.1108	0.0437
0.96	0.1275	0.0321	0.097	0.0462
0.97	0.1156	0.0309	0.0802	0.0482
0.98	0.102	0.0278	0.0591	0.0486
0.99	0.0879	0.0223	0.0331	0.0427

n=75	Exact		Wald	
p	Mean	Std. dev.	Mean	Std. dev.
0.01	0.0642	0.0174	0.0318	0.0314
0.02	0.0784	0.0212	0.0538	0.0324
0.03	0.0905	0.0225	0.0702	0.0308
0.04	0.1007	0.023	0.0838	0.028
0.05	0.1103	0.023	0.0942	0.0267
0.06	0.1187	0.0228	0.1039	0.0249
0.07	0.1262	0.022	0.1122	0.0242
0.08	0.1333	0.0214	0.1197	0.0234
0.09	0.1401	0.0201	0.1269	0.0226
0.1	0.1458	0.0206	0.1334	0.0219
0.15	0.1709	0.0179	0.1593	0.0187
0.2	0.1989	0.0153	0.1743	0.0162
0.8	0.1898	0.1542	0.1793	0.016
0.85	0.1715	0.0177	0.1592	0.0189
0.9	0.1458	0.0206	0.1327	0.0215
0.91	0.1397	0.0213	0.1267	0.0224
0.92	0.1333	0.0218	0.1196	0.0231
0.93	0.1237	0.0223	0.1121	0.0237
0.94	0.1185	0.0222	0.1039	0.0251
0.95	0.1101	0.0229	0.0946	0.0262
0.96	0.1007	0.0226	0.0831	0.0285
0.97	0.0903	0.0225	0.0707	0.0302
0.98	0.0784	0.0212	0.0539	0.0325
0.99	0.0641	0.0174	0.032	0.0314

4.2. Bayesian Credible Sets with Uniform and Beta Priors

In this subchapter the simulations and the results of constructing the Bayesian credible sets using uniform and beta priors are shown.

Table 4.8. Mean and Standard deviations of the interval lengths for n = 100 and n = 1000

n=100	Exact		Wald	
p	Mean	Std. dev.	Mean	Std. dev.
0.01	0.0526	0.0147	0.0298	0.0243
0.02	0.0655	0.017	0.0492	0.0237
0.03	0.0763	0.0174	0.0631	0.0215
0.04	0.0859	0.0176	0.0738	0.0202
0.05	0.0942	0.0176	0.0828	0.0191
0.06	0.1017	0.0172	0.0908	0.0184
0.07	0.1081	0.0167	0.0978	0.0177
0.08	0.1146	0.0163	0.1045	0.0172
0.09	0.1201	0.0159	0.1103	0.0165
0.1	0.1254	0.0155	0.1158	0.0162
0.15	0.1473	0.0137	0.1385	0.014
0.2	0.1641	0.0117	0.1556	0.0121
0.8	0.1641	0.0116	0.1556	0.012
0.85	0.1472	0.0137	0.1387	0.014
0.9	0.1256	0.0153	0.1156	0.016
0.91	0.1203	0.0158	0.1103	0.0166
0.92	0.1147	0.0162	0.1047	0.0168
0.93	0.1083	0.0165	0.0982	0.0176
0.94	0.1017	0.0171	0.0911	0.0181
0.95	0.0941	0.0174	0.0832	0.0191
0.96	0.0857	0.0177	0.0735	0.0201
0.97	0.0764	0.0177	0.0627	0.0218
0.98	0.0655	0.017	0.0493	0.0237
0.99	0.0525	0.0145	0.0301	0.0246

n=1000	Exact		Wald	
p	Mean	Std. dev.	Mean	Std. dev.
0.01	0.0133	0.0019	0.0121	0.002
0.02	0.0183	0.0019	0.0172	0.0019
0.03	0.0221	0.0018	0.021	0.0019
0.04	0.0253	0.0018	0.0242	0.0018
0.05	0.028	0.0018	0.0269	0.0018
0.06	0.0304	0.0017	0.0294	0.0017
0.07	0.0326	0.0017	0.0315	0.0017
0.08	0.0346	0.0017	0.0336	0.0016
0.09	0.0364	0.0016	0.0354	0.0016
0.1	0.0381	0.0016	0.0371	0.0016
0.15	0.0452	0.0014	0.0442	0.0014
0.2	0.0505	0.0012	0.0495	0.0012
0.8	0.0505	0.0012	0.0495	0.0017
0.85	0.0452	0.0014	0.0442	0.0014
0.9	0.0381	0.0016	0.0371	0.0016
0.91	0.0305	0.0016	0.0354	0.0016
0.92	0.0346	0.0017	0.0336	0.0016
0.93	0.0326	0.0017	0.0316	0.0017
0.94	0.0304	0.0017	0.0293	0.0017
0.95	0.028	0.0018	0.0269	0.0018
0.96	0.0253	0.0018	0.0242	0.0018
0.97	0.0022	0.0018	0.021	0.0018
0.98	0.0183	0.0019	0.0173	0.0019
0.99	0.0134	0.0019	0.2124	0.002

As discussed in section 3.1, a reasonable prior for p is a Uniform(0 1) distribution. But since the point of interest in this thesis is when p is close to 0 or 1, the condition p is from a Uniform(0.8, 1)distribution is imposed. 10,000 simulations were carried out to capture the coverage probabilities for credible sets with this prior for various n. Table 4.9 gives a sample of the output obtained for the credible sets. The first column shows the simulation number, the second column, the lower limit, the third column, the value of p obtained from Uniform(0.8, 1), upper limit in the fourth column, the length of the confidence interval in the fifth column and the final column, titled hits shows a value of 1 if p is captured by the confidence interval and 0 if it is not.

Table 4.9.Sample of the program output for the credible sets with Uniform(0.8, 1) prior at n = 25

No.	Lower Limit	p	Upper Limit	Length	Coverage
1	0.8663417	0.957935	0.9877045	0.121363	1
2	0.8417501	0.955421	0.978251	0.136501	1
3	0.8663417	0.911334	0.9877045	0.121363	1
4	0.8956527	0.985009	0.9952134	0.099561	1
5	0.8073366	0.860262	0.9441232	0.136787	1
6	0.8029478	0.822665	0.9206793	0.117732	1
7	0.8417501	0.947101	0.978251	0.136501	1
8	0.8015601	0.810856	0.8996962	0.098136	1
9	0.8237624	0.869454	0.9675322	0.14377	1
10	0.8029478	0.80504	0.9206793	0.117732	1
11	0.9302316	0.965354	0.999503	0.069271	1
12	0.9302316	0.996788	0.999503	0.069271	1
13	0.8015601	0.846656	0.8996962	0.098136	1
14	0.8956527	0.953157	0.9952134	0.099561	1
15	0.8015601	0.845044	0.8996962	0.098136	1
16	0.8020838	0.83094	0.9097693	0.107685	1
17	0.8029478	0.82205	0.9206793	0.117732	1
18	0.8073366	0.817307	0.9441232	0.136787	1
19	0.8663417	0.962234	0.9877045	0.121363	1
20	0.8044701	0.822392	0.932224	0.127754	1
21	0.8015601	0.848199	0.8996962	0.098136	1
22	0.8417501	0.948357	0.978251	0.136501	1
23	0.8012187	0.800909	0.8905614	0.089343	0
24	0.8129735	0.915483	0.9560198	0.143046	1
25	0.8237624	0.929205	0.9675322	0.14377	1
26	0.8009874	0.807842	0.8823877	0.0814	1
27	0.8015601	0.808707	0.8996962	0.098136	1
28	0.8015601	0.836618	0.8996962	0.098136	1
29	0.8073366	0.882564	0.9441232	0.136787	1
30	0.8129735	0.878934	0.9560198	0.143046	1
31	0.8029478	0.830809	0.9206793	0.117732	1
32	0.8663417	0.985265	0.9877045	0.121363	1
33	0.8029478	0.860651	0.9206793	0.117732	1
34	0.9302316	0.998572	0.999503	0.069271	1
35	0.8237624	0.954182	0.9675322	0.14377	1
36	0.8237624	0.949028	0.9675322	0.14377	1
37	0.8044701	0.915277	0.932224	0.127754	1
38	0.8663417	0.977294	0.9877045	0.121363	1

Simulations were carried out for various n and the summary of the results of the simulations is tabulated below.

Table 4.10. Summary of simulation results using a Uniform(0.8, 1) prior.

n	Mean Length	Std.dev	Coverage
5	0.1785	0.0143	0.9486
25	0.1434	0.0208	0.9497
50	0.1171	0.0242	0.9478
75	0.1006	0.0243	0.9526
100	0.08959	0.0234	0.9504
1000	0.03269	0.0103	0.9493

The mean length of the credible interval is the smallest of the three methods. This means that tighter confidence limits are obtained. The coverage probability is almost equal to 0.95, i.e. it is close to the nominal level of 95%.

The following figures show the plots of some uniform prior distributions and their posteriors.

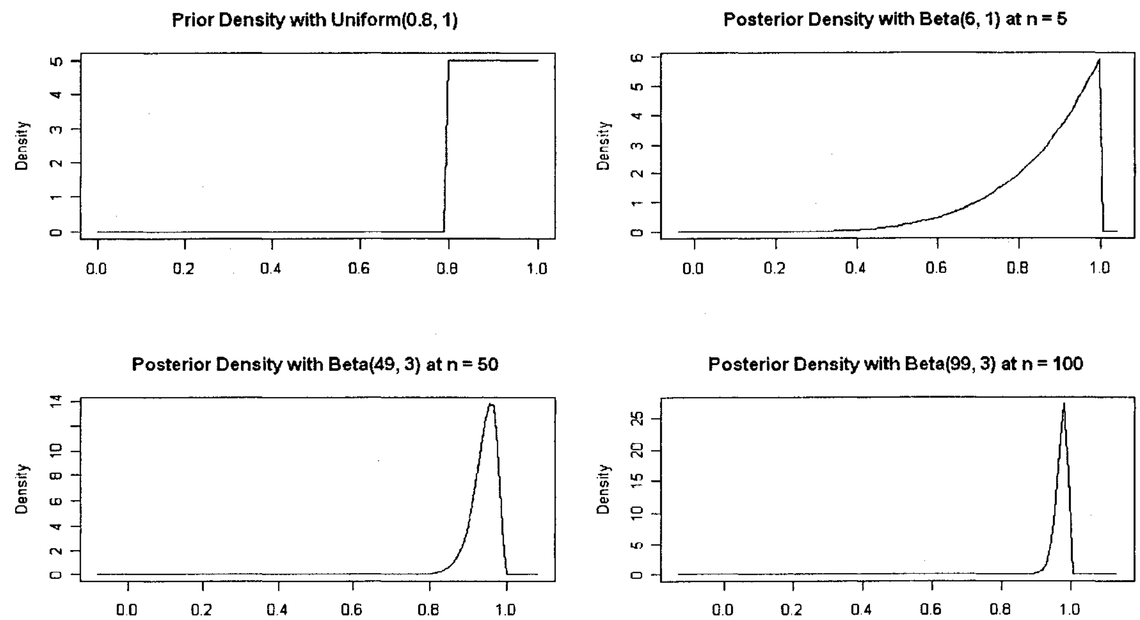


Figure 4.6. Uniform(0.8, 1) prior with its posterior distributions

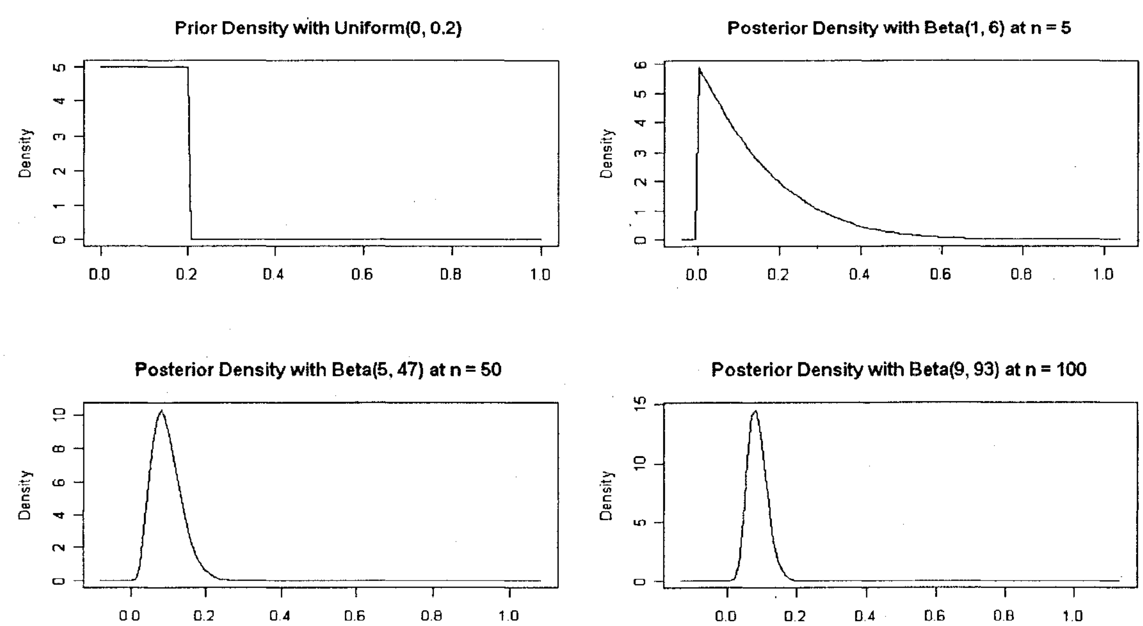


Figure 4.7. Uniform(0, 0.2) prior with its posterior distributions.

In this section, the results of the simulations using the beta prior are shown.

Table 4.11. Sample of the program output for the credible sets with Beta(10, 100) prior at n = 100

No.	p	x	n	alpha	beta	mean	lcl	ucl	sig	Cov.
1	0.180477	23	100	33	177	0.157143	0.111249	0.209217	0.05	1
2	0.207354	22	100	32	178	0.152381	0.107145	0.203882	0.05	0
3	0.168105	21	100	31	179	0.147619	0.103058	0.19853	0.05	1
4	0.158828	21	100	31	179	0.147619	0.103058	0.19853	0.05	1
5	0.150113	21	100	31	179	0.147619	0.103058	0.19853	0.05	1
6	0.136168	20	100	30	180	0.142857	0.098988	0.193162	0.05	1
7	0.171692	20	100	30	180	0.142857	0.098988	0.193162	0.05	1
8	0.18974	20	100	30	180	0.142857	0.098988	0.193162	0.05	1
9	0.18948	20	100	30	180	0.142857	0.098988	0.193162	0.05	1
10	0.177127	20	100	30	180	0.142857	0.098988	0.193162	0.05	1
11	0.128533	20	100	30	180	0.142857	0.098988	0.193162	0.05	1
12	0.115223	20	100	30	180	0.142857	0.098988	0.193162	0.05	1
13	0.131487	20	100	30	180	0.142857	0.098988	0.193162	0.05	1
14	0.129157	19	100	29	181	0.138095	0.094937	0.187775	0.05	1
15	0.114785	19	100	29	181	0.138095	0.094937	0.187775	0.05	1
16	0.129867	19	100	29	181	0.138095	0.094937	0.187775	0.05	1
17	0.120519	19	100	29	181	0.138095	0.094937	0.187775	0.05	1
18	0.138153	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
19	0.097407	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
20	0.135701	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
21	0.12967	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
22	0.132985	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
23	0.136284	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
24	0.134807	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
25	0.106409	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
26	0.140049	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
27	0.187603	18	100	28	182	0.133333	0.090904	0.182369	0.05	0
28	0.134971	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
29	0.16262	18	100	28	182	0.133333	0.090904	0.182369	0.05	1
30	0.110585	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
31	0.122284	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
32	0.135033	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
33	0.123648	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
34	0.115206	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
35	0.126473	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
36	0.147317	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
37	0.123193	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
38	0.120946	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
39	0.139787	17	100	27	183	0.128571	0.08689	0.176944	0.05	1
40	0.141925	17	100	27	183	0.128571	0.08689	0.176944	0.05	1

Table 4.12. Sample of the program output for the credible sets with Beta(100, 10) prior at n = 75

No.	p	x	n	alpha	beta	mean	lcl	ucl	sig	Cov.
1	0.855084	65	75	165	20	0.891892	0.843447	0.932329	0.05	1
2	0.895905	70	75	170	15	0.918919	0.875634	0.953655	0.05	1
3	0.895612	64	75	164	21	0.886487	0.837128	0.927947	0.05	1
4	0.873737	66	75	166	19	0.897297	0.849802	0.936676	0.05	1
5	0.938065	70	75	170	15	0.918919	0.875634	0.953655	0.05	1
6	0.879484	64	75	164	21	0.886487	0.837128	0.927947	0.05	1
7	0.899392	70	75	170	15	0.918919	0.875634	0.953655	0.05	1
8	0.914322	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
9	0.882051	65	75	165	20	0.891892	0.843447	0.932329	0.05	1
10	0.903907	71	75	171	14	0.924324	0.882213	0.95778	0.05	1
11	0.930297	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
12	0.889599	74	75	174	11	0.940541	0.902325	0.969782	0.05	0
13	0.877516	70	75	170	15	0.918919	0.875634	0.953655	0.05	1
14	0.930285	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
15	0.897031	70	75	170	15	0.918919	0.875634	0.953655	0.05	1
16	0.941951	72	75	172	13	0.92973	0.88885	0.961847	0.05	1
17	0.863479	69	75	169	16	0.913514	0.869107	0.949479	0.05	0
18	0.91178	67	75	167	18	0.902703	0.856195	0.940986	0.05	1
19	0.898088	68	75	168	17	0.908108	0.862629	0.945254	0.05	1
20	0.899886	68	75	168	17	0.908108	0.862629	0.945254	0.05	1
21	0.841673	65	75	165	20	0.891892	0.843447	0.932329	0.05	0
22	0.894879	66	75	166	19	0.897297	0.849802	0.936676	0.05	1
23	0.878103	70	75	170	15	0.918919	0.875634	0.953655	0.05	1
24	0.909249	64	75	164	21	0.886487	0.837128	0.927947	0.05	1
25	0.888152	65	75	165	20	0.891892	0.843447	0.932329	0.05	1
26	0.942452	71	75	171	14	0.924324	0.882213	0.95778	0.05	1
27	0.86808	67	75	167	18	0.902703	0.856195	0.940986	0.05	1
28	0.931695	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
29	0.912767	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
30	0.844343	64	75	164	21	0.886487	0.837128	0.927947	0.05	1
31	0.878436	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
32	0.905176	68	75	168	17	0.908108	0.862629	0.945254	0.05	1
33	0.845634	63	75	163	22	0.881081	0.830842	0.923531	0.05	1
34	0.936504	74	75	174	11	0.940541	0.902325	0.969782	0.05	1
35	0.872118	67	75	167	18	0.902703	0.856195	0.940986	0.05	1
36	0.874051	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
37	0.95156	71	75	171	14	0.924324	0.882213	0.95778	0.05	1
38	0.852361	67	75	167	18	0.902703	0.856195	0.940986	0.05	0
39	0.904894	66	75	166	19	0.897297	0.849802	0.936676	0.05	1
40	0.941301	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
41	0.854342	56	75	156	29	0.843243	0.78763	0.891835	0.05	1
42	0.913493	66	75	166	19	0.897297	0.849802	0.936676	0.05	1
43	0.923036	69	75	169	16	0.913514	0.869107	0.949479	0.05	1
44	0.897717	65	75	165	20	0.891892	0.843447	0.932329	0.05	1

Tables 4.11 and 4.12 give a sample of the output obtained for the credible sets when the priors are Beta(10, 100) and Beta(100, 10) respectively. The first column shows the simulation number, the second column, the value of p obtained from the prior, the third column gives the value of x, n in the fourth column, the posterior values of alpha and beta in the 5th and 6th columns respectively, the mean of the posterior in the seventh column, the lower credible limit and the upper credible limit in eighth and ninth columns, the significance level and the last column, the shows a value of 1 if p is captured by the credible interval and 0 if it is not.

The summary of all the simulations with Beta(10, 100) prior is shown in table 4.13.

Table 4.13. Summary of simulation results using a Beta(10, 100) prior.

n	Mean length	Std.dev	Coverage
5	0.1245	0.0126	0.9511
25	0.0985	0.0069	0.9584
50	0.0885	0.0065	0.953
75	0.0818	0.0069	0.9684
100	0.0769	0.0071	0.9448
1000	0.0334	0.0043	0.9461

It is seen that the coverage probability for all values of n is almost equal to the nominal level of 95%. The mean credible interval length is also smaller than the Wald interval and the Exact interval. That is Bayesian credible intervals are tighter than the Wald and Exact intervals. Figure 4.8 shows the plots of some posterior distributions with Beta(10, 100) prior distribution.

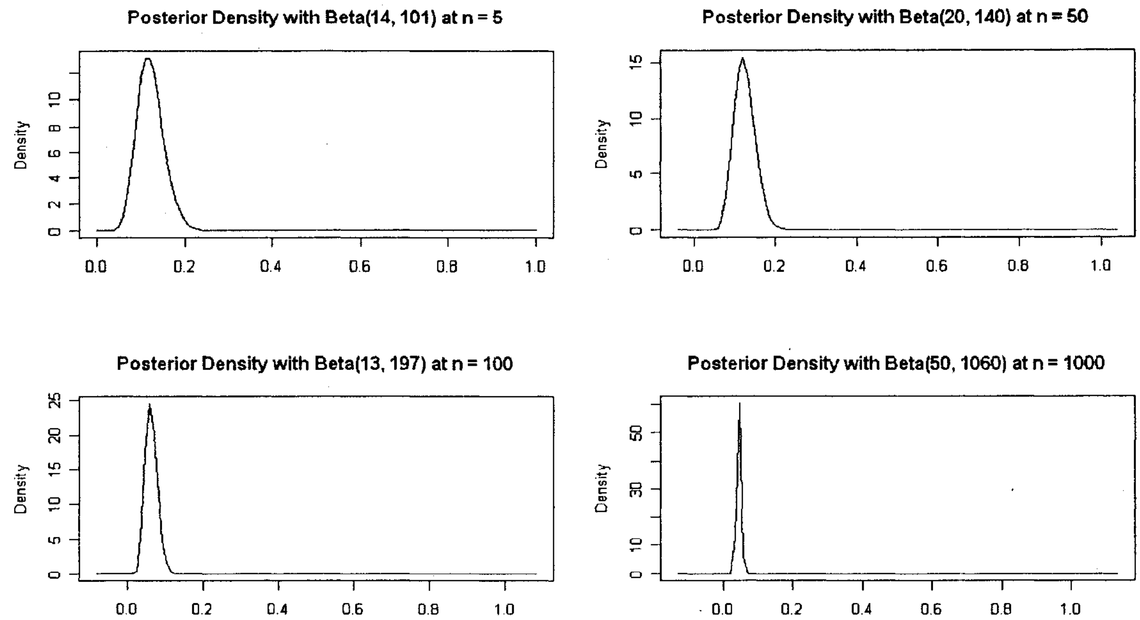


Figure 4.8. Posterior densities with Beta(10, 100) priors

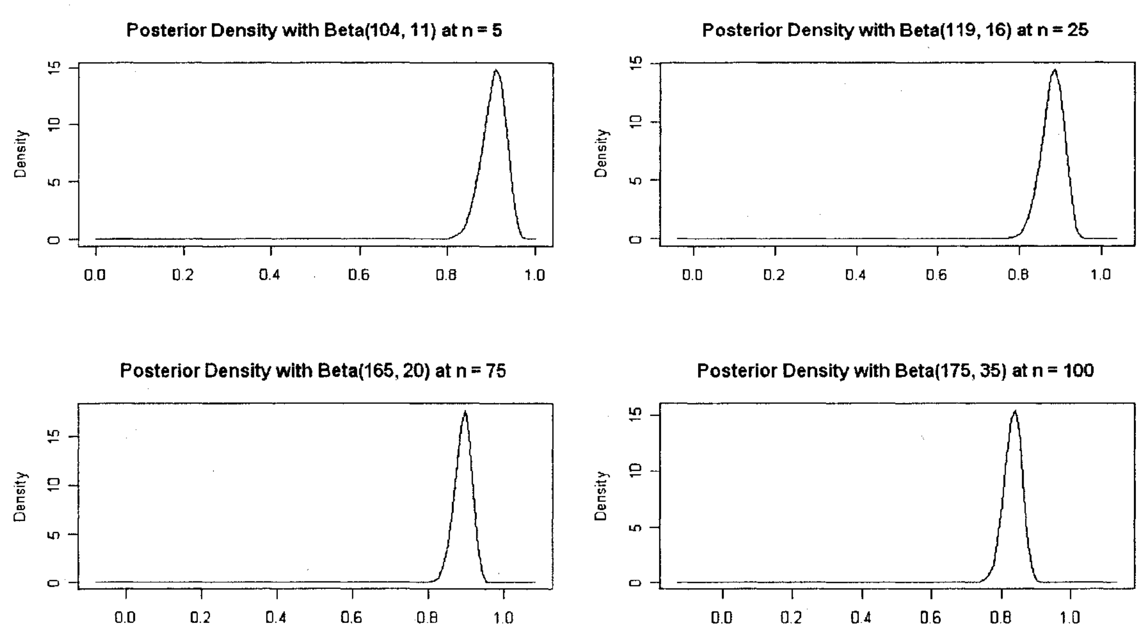


Figure 4.9. Posterior densities with Beta(100, 10) priors

Table 4.13. Summary of simulation results using a Beta(100, 10) prior.

n	Mean Length	Std.dev	Coverage
5	0.1234	0.013	0.9453
25	0.0991	0.0073	0.9515
50	0.0885	0.0062	0.9551
75	0.0818	0.0067	0.9496
100	0.0763	0.007	0.9478
1000	0.0333	0.0041	0.9456

Figure 4.9 shows the plots of some posterior distributions with Beta(100, 10) prior distribution. A summary of the simulation results using Beta(100, 10) is shown in table 4.13. It is seen that, for the values of p closer to 0 or 1, the credible intervals give very good coverage probabilities.

CHAPTER 5

CONCLUSIONS

This work is primarily a comparison of confidence intervals for a binomial proportion. The literature review showed that there are a lot of methods for analyzing and computing the confidence intervals of the parameter p for the binomial distribution. The methods compared here are the Wald interval, the Exact interval and the Bayesian credible sets.

The Wald interval gives a poor coverage probability when n is very low. For values of p closer to 0 or 1, even for large values of n. ($n > 50$), the coverage probability is less than the nominal value of 95% in most cases. The only advantage is its simplicity and the tighter confidence interval lengths.

The Exact interval consistently gives a higher coverage probability than the nominal 95% level. The coverage probability is very close to 0.95 only at high values of n. Even though the length of the Exact confidence intervals are longer than the Wald confidence intervals, the coverage probability is much better than the Wald interval.

The Bayesian credible sets consistently give the required coverage probabilities. The credible limit lengths are consistently smaller compared to the Wald and the Exact intervals. That is the Bayesian credible sets give have a tighter control over the lengths of the intervals. The priors are highly informative. This results in the posterior having smaller probability range and smaller confidence limits.

This thesis has shown that the Wald interval gives the lowest coverage probability of the three models. A lot of literature also points to the same. So it is recommended that the Wald Interval should be used only to demonstrate the generation of confidence intervals.

The computation of the confidence intervals for the binomial proportion should be done using the Exact Interval or the Bayesian method or the other methods like Score Interval, The Stern Interval etc.

REFERENCES

Bolstad, W. M., *Introduction to Bayesian Statistics,* Wiley-Interscience, (2004).

Wilson, E.B., *Probable inference, the law of succession, and statistical inference,* JASA, 22, 209-212 (1927)

Wald, A., and Wolfowitz, J., *Confidence limits for continuous distribution functions,* The Annals of Mathematical Statistics, 10, 105–118. (1939).

Agresti, A., and Coull, B.A. *Approximate is better than "exact" for interval estimation of binomial proportion,* The American Statistician, 52, 119–126, (1998).

Clopper, C.J., and Pearson, E.S., *The use of confidence or fiducial limits illustrated in the case of the binomial,* Biometrika, 26, 404–413, (1934).

Ghosh, B.K., *A comparison of some approximate confidence intervals for the binomial parameter,* Journal of the American Statistical Association, 74, 894–900, (1979).

Blyth, C.R., and Still, H.A. *Binomial confidence intervals*, Journal of the American Statistical Association, 78, 108–116, (1983).

Gill, J., *Bayesian Methods: Asocial and Behavioral Sciences Approach,* Chapmann and Hall/CRC, (2002).

Verzani J., *Using R for Introductory Statistic,* Chapman and Hall/CRC Press (2005).

R Reference Manuals.

Kotz, S. and Johnson, N.L., *Discrete Distributions*, Houghton Mifflin (1969).

VITA

Graduate College
University of Nevada, Las Vegas

Narain Armbya

Home Address:
1455 E Rochelle Ave, Apt#5
Las Vegas, NV 89119

Degrees:
Master of Science, Mechanical Engineering, 2004
University of Nevada, Las Vegas, Nevada

Thesis Title:
Comparison of Confidence Intervals for Binomial Proportions

Thesis Examination Committee:
Chairperson, Dr. Rohan Dalpatadu., Ph. D.
Committee Member, Dr. Dennis Murphy, Ph. D.
Committee Member, Dr. Xin Li, Ph. D.
Graduate Faculty Representative, Dr. Ashok Singh, Ph. D.