Journal of Statistical Computation and Simulation
Vol. 78, No. 9, September 2008, 829–841

Confidence intervals for the mean of a normal distribution with restricted parameter space

HSIUYING WANG

Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan

(Received 30 July 2006; final version received 12 February 2007)

For a normal distribution with known variance, the standard confidence interval of the location parameter is derived from the classical Neyman procedure. When the parameter space is known to be restricted, the standard confidence interval is arguably unsatisfactory. Recent articles have addressed this problem and proposed confidence intervals for the mean of a normal distribution where the parameter space is not less than zero. In this article, we propose a new confidence interval, rp interval, and derive the Bayesian credible interval and likelihood ratio interval for general restricted parameter space. We compare these intervals with the standard interval and the minimax interval. Simulation studies are undertaken to assess the performances of these confidence intervals.

Keywords: likelihood ratio interval; Bayesian credible interval; rp interval; coverage probability

1. Introduction

Traditionally, statistical theory is established in a natural parameter space. However, in many real applications, such as physics experiments, the parameter space is known to be restricted. Mandelkern [1] pointed out the importance of statistical inference for the bounded parameter space in physics, and gave the example that the classical Neyman procedure is not satisfactory to many scientists where the parameter is known to be bounded. The main problem in Mandelkern [1], related to the normal distribution, is setting confidence intervals for the mean of a normal distribution $N(\theta, \sigma^2)$, where the mean is known to be not less than 0, $\theta \geq 0$, and the variance σ^2 is known. The other related literatures are Feldman and Cousins [2] and Roe and Woodroofe [3], which proposed several alternatives for setting confidence bounds in the bounded parameter space, $\theta \geq 0$. A crucial reason that the classical procedure is not satisfactory for the restricted parameter space is that it does not sufficiently utilize the information regarding the restriction.

Setting confidence bounds for the restricted parameter space, $\{\theta : \theta \geq 0\}$, is widely discussed in the literatures [1, 3, 4]. The problem also arises in many applications for the general restricted parameter space, $\{\theta : \theta \in (a, b)\}$, where *a, b* are some constants. In this

Journal of Statistical Computation and Simulation
ISSN 0094-9655 print/ISSN 1563-5163 online © 2008 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00949650701273902

article, setting confidence intervals of the mean of a normal distribution for general restricted parameter space is investigated.

Let X be a random sample from a $N(\theta, \sigma^2)$ population with unknown mean θ and known variance σ^2. Without loss of generality, σ^2 is assumed to be 1. Assume that the parameter space Ω of θ is known to be restricted and the bounds of the parameter space are known. Denote a and b as the lower bound and the upper bound of the parameter space. Here, it is feasible to assume that a and b are known in the real application, as in the case of physics experiments. Another example is to test if the mean of the lengths of products produced by a machine is less than or greater than a standard length. We can measure some products first, then obtain rough upper and lower bounds for the mean, and use them to be the bounds of the parameter space. It should be concluded that for most applications, it is not difficult to find rough bounds for the parameter of interest.

The goal here is to estimate θ. The standard $1-\alpha$ confidence interval for θ is

$$(x - z_{\alpha/2}, x + z_{\alpha/2}), \tag{1}$$

where $z_{\alpha/2}$ is the $\alpha/2$ upper cutoff point of the standard normal distribution. The interval is derived from the uniformly most powerful unbiased test for testing

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta \neq \theta_0, \tag{2}$$

which is equal to the likelihood ratio test when the parameter space is the natural parameter space. When Ω is known to be restricted to the interval (*a, b*); it is widely accepted to choose the interval

$$(\max(a, x - z_{\alpha/2}), \min(b, x + z_{\alpha/2})) \tag{3}$$

which is the intersection of (*a, b*) and interval (1), as a confidence interval for θ [5]. The standard interval (1) is derived from the classical Neynam procedure for the natural parameter space and interval (3) is restricted to the domain of the bounded parameter space. However, the construction of the standard interval does not sufficiently utilize the bounds of the parameter space compared with some other intervals that we describe below.

The second interval is the minimax interval. Evans *et al.* [6] show that for inference about a normal mean $\theta \in [-\tau, \tau]$, $\tau \leq 2z_{1-\alpha}$, the truncated Pratt interval:

$$I_{TP}(X) = I_P(X) \cap [-\tau, \tau],$$

has minimax expected length among random sets with at least $1-\alpha$ coverage probability, where $I_P(X)$ is the Pratt interval [7]

$$I_P(X) = \begin{cases} [(X-c), 0 \vee (X+c)], & X \leq 0 \\ [0 \wedge (X-c), X+c], & X > 0, \end{cases}$$

with $c = z_\alpha$

The third confidence interval is the interval derived from the likelihood ratio test. For testing hypothesis (2), a widely used test is the likelihood ratio test. It is called the unified approach in Feldman and Cousins [2] when the parameter space is $\Omega = [0, \infty)$.

The fourth approach is the Bayesian credible interval. The Bayesian credible interval for the general restricted parameter space when the parameter is given a uniform distribution over the restricted range will be given in section 2.3.

The fifth confidence interval is constructed from a modified *p*-value. For testing hypothesis (2), Wang [8] proposed a modified *p*-value,

$$r_{\theta_0}(x) = \frac{P(|Z| > |x - \theta_0|) - \min_{\theta \in (a,b)} P(|Z| > |x - \theta|)}{\max_{\theta \in (a,b)} P(|Z| > |x - \theta|) - \min_{\theta \in (a,b)} P(|Z| > |x - \theta|)}, \tag{4}$$

which is called an rp-value, as a modified measure of evidence against the null hypothesis instead of the usual *p*-value in the restricted parameter space (*a, b*). The confidence interval corresponding to the rp-value will be investigated in this article.

The derivations of the latter three intervals depend on the bounds of the parameter space. The coverage probabilities and the expected lengths of the five intervals are studied in this article.

The article is organized as follows. Section 2 describes the three confidence intervals where construction depends on the bounds of the parameter space, and the explicit forms of the intervals are given. Simulation studies for comparing the three intervals with the standard interval and the minimax interval are conducted in section 3. The performances of these intervals are based on their coverage probabilities and expected lengths. Section 4 provides a real data example. The coverage probabilities of the five intervals covering the true mean of this data set are given. A conclusion about the performances of the intervals is given in section 5.

2. Confidence intervals

In this section, three confidence intervals for the mean of a normal distribution with variance 1 are derived when the parameter space is (*a, b*).

2.1 *Likelihood ratio interval*

For testing hypothesis (2), the likelihood ratio statistic is

$$\Lambda(x) = \frac{(1/\sqrt{2\pi})\mathrm{e}^{-((x-\theta_0)^2/2)}}{\max_{\theta \in (a,b)} (1/\sqrt{2\pi})\mathrm{e}^{-((x-\theta)^2/2)}}. \tag{5}$$

Then $-2 \log \Lambda(x)$ has a chi-square distribution with degree of freedom 1. The $(1-\alpha)$ confidence interval derived from the likelihood ratio test is

$$\{\theta_0 : -2 \ \log \ \Lambda(x) < \chi^2_{1,\alpha}\}, \tag{6}$$

where $\chi^2_{1,\alpha}$ denotes the upper α cutoff point of a chi-square distribution with degree of freedom 1. By straightforward calculation (6) can be rewritten as

$$\left(x - \sqrt{\chi^2_{1,\alpha} - 2 \ \log \ H}, x + \sqrt{\chi^2_{1,\alpha} - 2 \ \log \ H}\right), \tag{7}$$

where $H = \max_{\theta \in (a,b)} \mathrm{e}^{-(x-\theta)^2/2}$. Since Ω is known to be restricted to (*a, b*), interval (7) is modified as

$$\left(\max(a, x - \sqrt{\chi^2_{1,\alpha} - 2 \ \log \ H}, \min(b, x + \sqrt{\chi^2_{1,\alpha} - 2 \ \log \ H})\right), \tag{8}$$

which is called the $(1-\alpha)$ likelihood ratio interval.

2.2 *Bayesian credible interval*

We derive Bayesian credible intervals when the parameter space is given a uniform prior distribution. Let the prior of the bounded parameter space be $< (\theta) = (1/b - a)I_{\theta\in(a,b)} \rightarrow \pi(\theta) = 1/(b-a)I_{\theta\in(a,b)}$, where $I_{\theta\in(a,b)}$ denotes the indicator function. The posterior distribution of θ given x is

$$f(\theta|x) = \frac{(1/\sqrt{2\pi})\mathrm{e}^{-(x-\theta)^2/2} I_{\theta\in(a,b)}}{c(x)},$$

where $c(x) = \int_a^b (1/\sqrt{2\pi})\mathrm{e}^{-(x-\theta)^2/2}\,\mathrm{d}\theta$. We wish to find an upper limit u and a lower limit l dependent on x, for which

$$P(l \le \theta \le u|x) = 1-\alpha$$

and the length $u - l$ is shortest. In Bayes theory, such an interval is called a credible interval. We have

$$(l, u) = \{\theta : f(\theta|x) \ge c'\}, \tag{9}$$

where c' satisfies $P(f(\theta|x) \ge c') = 1-\alpha$. Then equation (9) can be rewritten as $\{\theta : |x-\theta| \le d\}$, where d satisfies

$$1-\alpha = \int_{\max(x-d,a)}^{\min(b,x+d)} f(\theta|x)\,\mathrm{d}\theta. \tag{10}$$

By solving d in equation (10), the credible interval based on x with respect to $\pi(\theta)$ is

$$(x-d, x+d). \tag{11}$$

Since Ω is restricted to (a, b), equation (11) is modified as

$$(\max(a, x-d), \min(b, x+d)). \tag{12}$$

Since the solution d in equation (10) does not have a closed form, interval (12) should be derived by a numerical calculation.

2.3 *rp interval*

The fifth interval is derived from the rp-value (4). Wang [8] recommends using (4) as a measure of evidence against the null hypothesis in hypothesis (2) when the parameter space is restricted to (a, b). The motivation for proposing equation (4) is as follows.

The range of the usual p-value $p_{\theta_0}(x) = P(|Z| > |x-\theta_0|)$ is $(\min_{\theta\in(a,b)} P(|Z| > |x-\theta|), \max_{\theta\in(a,b)} P(|Z| > |x-\theta|))$ when θ_0 belongs to (a, b). Since no matter what value of $\theta_0 \in (a, b)$ is chosen to be the null hypothesis $\theta = \theta_0$, the p-value $p_{\theta_0}(x)$ is always greater than $\min_{\theta\in(a,b)} P(|Z| > |x-\theta|)$. Thus, the magnitude $\min_{\theta\in(a,b)} P(|Z| > |x-\theta|)$ should not be included in a measure of evidence against the null hypothesis. Hence, a reasonable measure of evidence should be the usual p-value minus this magnitude, which is

$$p_{\theta_0}(x) - \min_{\theta\in(a,b)} P(|Z| > |x-\theta|). \tag{13}$$

Moreover, usually we would decide to reject the null hypothesis or accept the null hypothesis by comparing the p-value with a value α, where $0 < \alpha < 1$. If we can transform (13) such that the range is (0, 1), then it is more reasonable to compare it with a value α between 0 and 1. Thus, for fixed x, we divide (13) by $\left(\max_{\theta\in(a,b)} P(|Z| > |x-\theta|) - \min_{\theta\in(a,b)} P(|Z| > |x-\theta|)\right)$ such that its range is (0, 1) which leads to (4).

Under this transformation, the range of (4) is (0, 1). Note that when the parameter space is the natural parameter space, (4) is exactly equal to the usual *p*-value.

The corresponding $(1-\alpha)$ interval to the rp-value is

$$\left\{\theta : \frac{P(|Z| > |x-\theta|) - \min_{\theta\in(a,b)} P(|Z| > |x-\theta|)}{\max_{\theta\in(a,b)} P(|Z| > |x-\theta|) - \min_{\theta\in(a,b)} P(|Z| > |x-\theta|)} > \alpha\right\} \tag{14}$$

Note that (14) can be rewritten as

$$\{\theta : P(|Z| > |x-\theta|) > \beta\}, \tag{15}$$

where

$$\beta = \alpha(\max_{\theta\in(a,b)} P(|Z| > |x-\theta|) - \min_{\theta\in(a,b)} P(|Z| > |x-\theta|)) + \min_{\theta\in(a,b)} P(|Z| > |x-\theta|)$$

Note that β depends on *x*,

$$\max_{\theta\in(a,b)} P(|Z| > |x-\theta|) = \begin{cases} P(|Z| > a-x) & \text{if } x < a, \\ 1 & \text{if } a \le x \le b, \\ P(|Z| > x-b) & \text{if } x > b \end{cases}$$

and

$$\min_{\theta\in(a,b)} P(|Z| > |x-\theta|) = \begin{cases} P(|Z| > b-x) & \text{if } x < a, \\ P(|Z| > \max(x-a, b-x)) & \text{if } a \le x \le b, \\ P(|Z| > x-a) & \text{if } b < x. \end{cases}$$

Interval (15) is equivalent to

$$(x - z_{\beta/2}, x + z_{\beta/2}). \tag{16}$$

Since Ω is known to be (*a, b*), (16) is modified as

$$(\max(a, x - z_{\beta/2}), \min(b, x + z_{\beta/2})). \tag{17}$$

The interval is called the $(1-\alpha)$ rp interval because it is derived from the rp-value.

3. Simulation study

The performances of the standard interval, minimax interval and those intervals derived in section 2 are investigated through simulation studies in this section. The coverage probabilities and the expected lengths of the five intervals are presented for some restricted parameter space. Figures 1 and 2 are the coverage probabilities of the five intervals with respect to parameter space $\Omega_1 = (-0.5, 0.5)$ and $\Omega_2 = (-1, 1)$, respectively. Figures 3 and 4 are the expected lengths of the five intervals with respect to Ω_1 and Ω_2, respectively. More simulation studies with respect to different restricted parameter spaces were conducted, but are not presented here. They show that the relationships of the performances of the five intervals for different restricted parameter spaces are similar to figures 1–4. Each simulation study is based on 1000 replicates. Note that the samples in the simulation study are randomly generated from $N(\theta, 1)$. It is possible to generate samples satisfying $\max(a, x - z_{\alpha/2}) > \min(b, x + z_{\alpha/2})$,

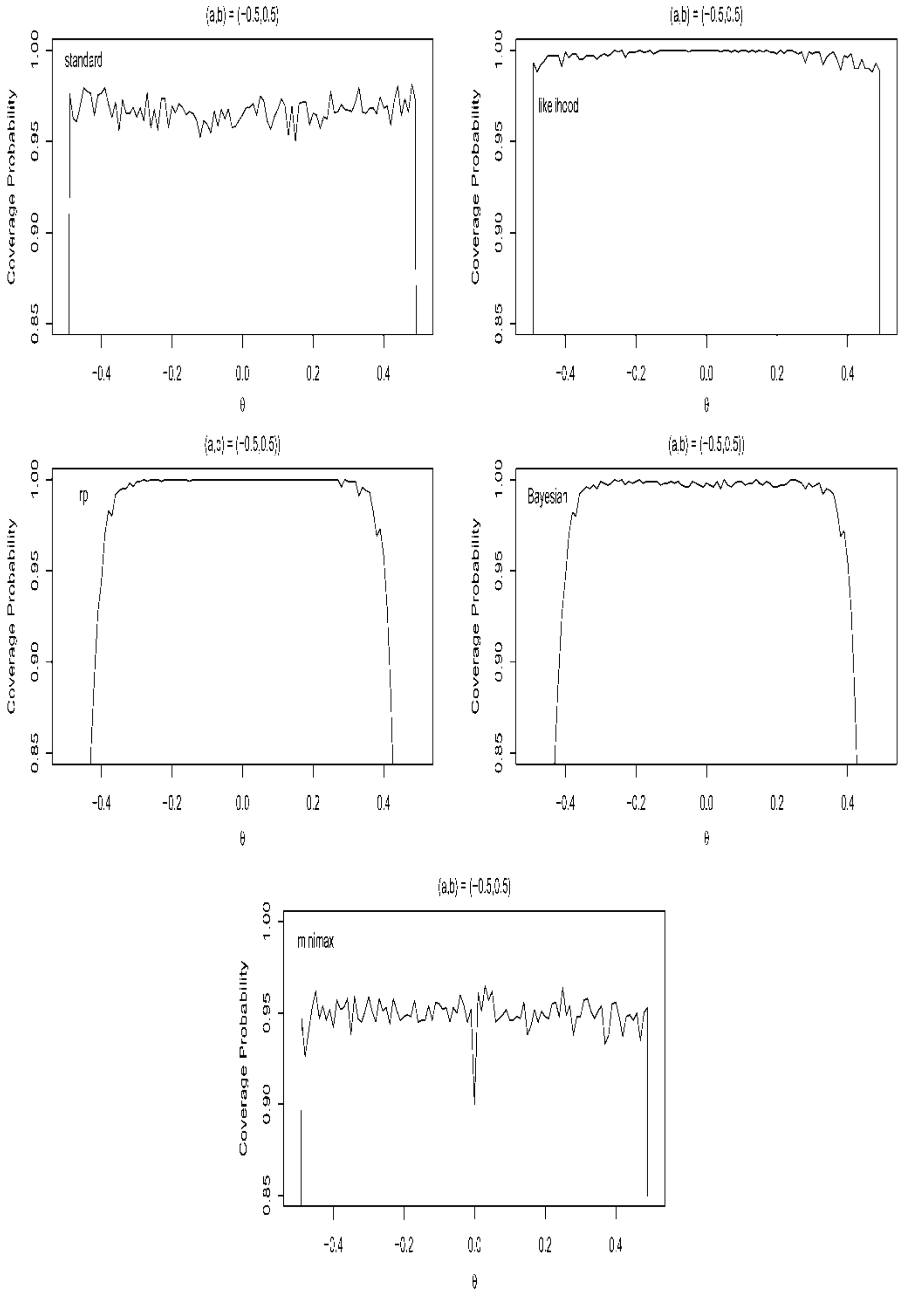

Figure 1. The coverage probabilities of the five level 0.95 intervals with respect to different parameter spaces $(-0.5, 0.5)$.

for example, $a < b < x - z_{\alpha/2} < x + z_{\alpha/2}$ or $x - z_{\alpha/2} < x + z_{\alpha/2} < a < b$. By the definition of the interval, if samples satisfy the condition, the confidence intervals based on them would be the empty set. Hence, we discard these kinds of samples in the simulation algorithm.

In the simulation study, the performances of the credible interval and rp interval are more similar than other intervals from both coverage probability and expected length aspects. They

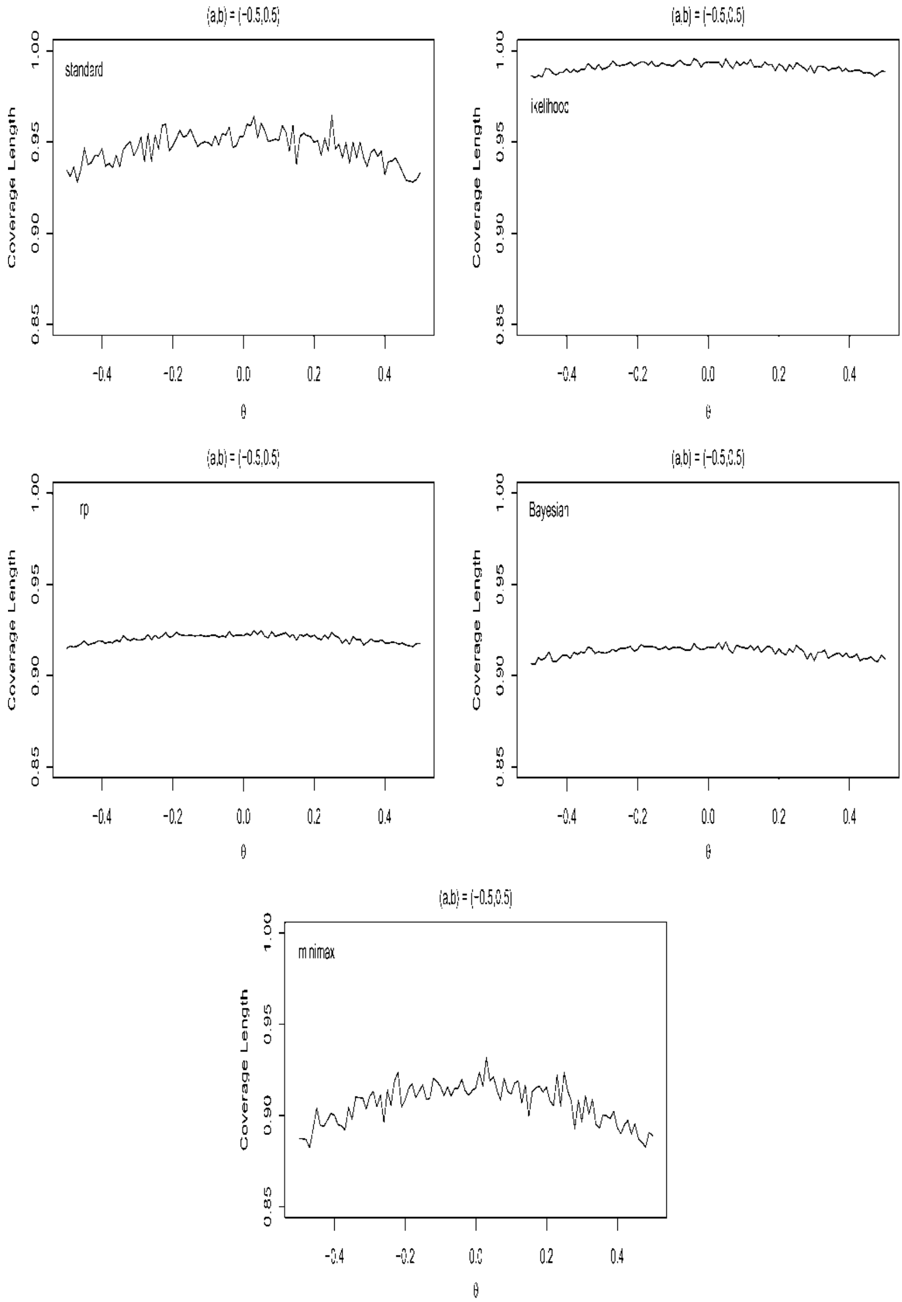

Figure 2. The expected lengths of the five level 0.95 intervals with respect to different parameter spaces $(-0.5, 0.5)$.

have higher coverage probability in most of the parameter spaces compared with other intervals. However, the coverage probability drops when the true parameter is near the boundary. The coverage probability of the standard interval is higher than the minimax interval, which is around 0.95, and the expected length of the standard interval is longer than the minimax interval. The likelihood ratio interval also has higher coverage probability, but its expected length is longer than the other intervals. In the cases shown in figures 1–4, the credible and *rp*

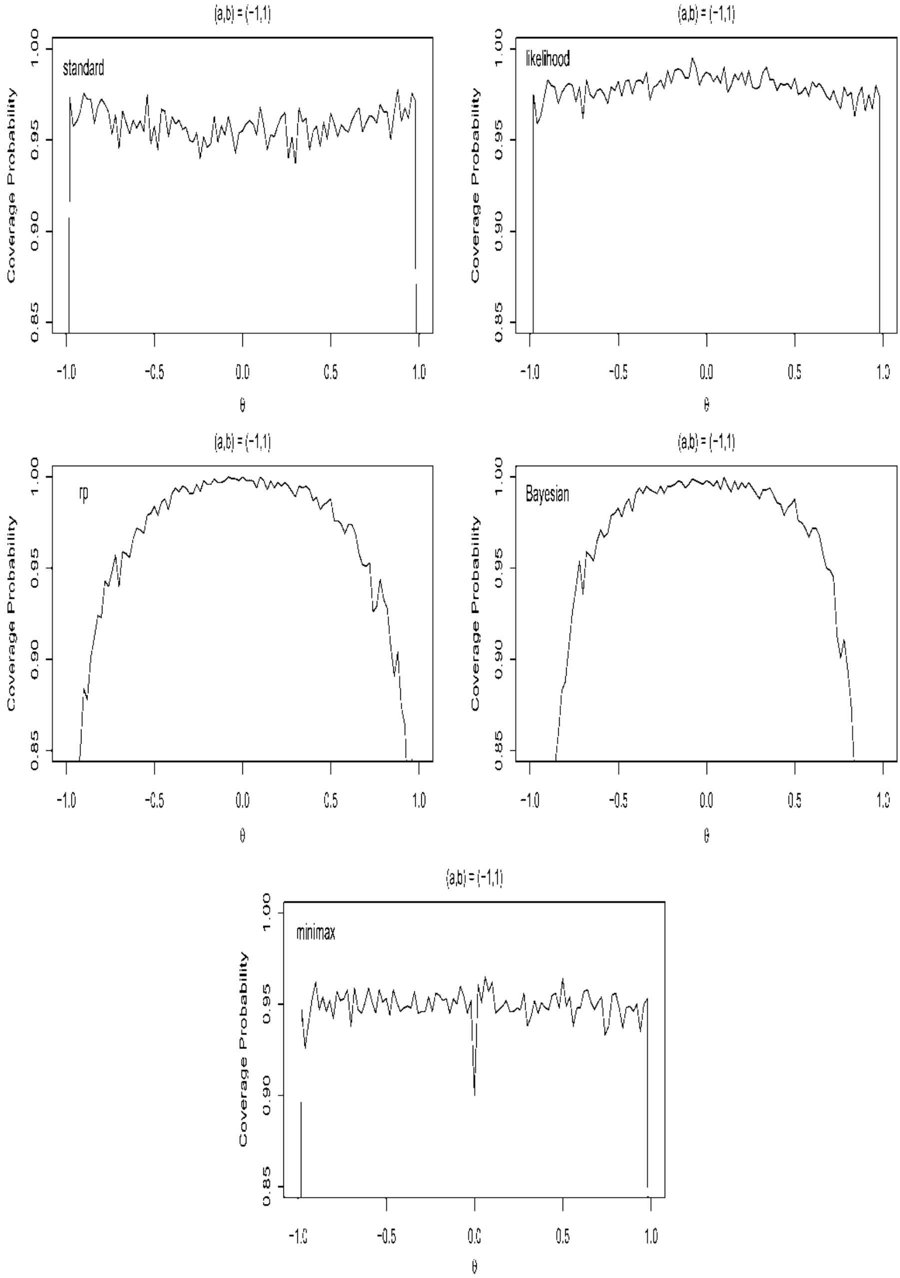

Figure 3. The coverage probabilities of the five level 0.95 intervals with respect to different parameter spaces $(-1, 1)$.

intervals perform well over the most range of parameters because they have higher coverage probabilities and shorter expected lengths than the standard interval. The disadvantage of the two intervals is their worse performance at the boundary of the parameter space. However, in real applications, the performance at the boundary may not be so important. When seeking the bounds of the parameter space, we may estimate some rough bounds of the parameter space such that the bounds can be larger than the true parameter. The true parameter probably does

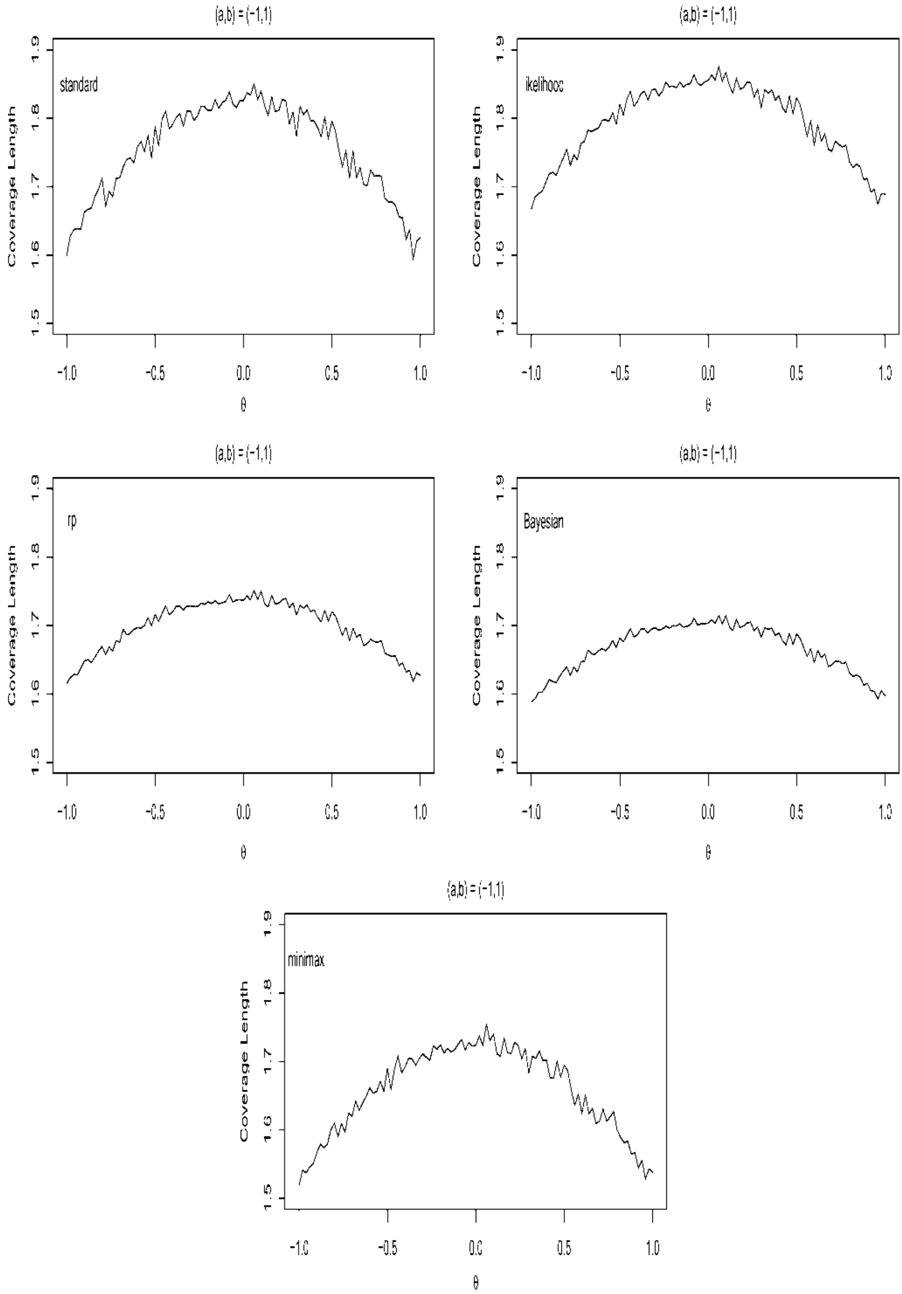

Figure 4. The expected lengths of the five level 0.95 intervals with respect to different parameter spaces $(-1, 1)$.

not occur at the boundary of the parameter space. Therefore, in this case, we do not need to be concerned about the performance of the intervals when the parameter is near the boundary of the restricted parameter space. Overall, the likelihood ratio, credible and rp intervals have higher coverage probabilities than the standard and minimax intervals; the likelihood ratio has longer expected length than the other intervals.

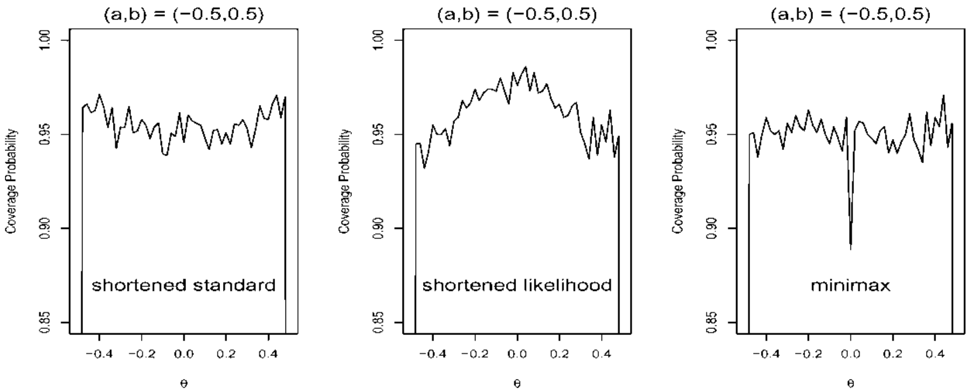

Figure 5. The coverage probabilities of the shortened standard, the shortened likelihood ratio and the minimax level 0.95 intervals with respect to different parameter spaces $(-0.5, 0.5)$.

From the simulation results, if we do not consider using a data-dependent approach, the wider confidence interval should have higher coverage probability. Only the credible and rp intervals are data-dependent and the other three intervals are not. This is the key point as to why these two intervals can have higher coverage probabilities and shorter average lengths.

In addition, from figures 1 and 3, we may shorten the standard interval and the likelihood ratio interval such that the coverage probabilities of these two intervals are closer to 0.95 by choosing smaller cutoff points, and then compare their expected lengths with the minimax interval. Figures 5–8 are the comparisons of the coverage probabilities and the expected lengths of the shortened standard interval, the shortened likelihood ratio interval and the minimax interval. In figures 5 and 6, for the parameter space $(-0.5, 0.5)$ case, we shorten the standard interval to $(\max(-0.5, x - z_{4\times 0.05}), \min(0.5, x + z_{4\times 0.05}))$ and the likelihood ratio interval to $(\max(-0.5, x - \sqrt{\chi^2_{1,3\times 0.05} - 2\log H}), \min(0.5, x + \sqrt{\chi^2_{1,3\times 0.05} - 2\log H}))$. In figure 7 and 8, for the parameter space $(-1, 1)$, we shorten the standard interval and the likelihood ratio interval to $(\max(-1, x - z_{3\times 0.05}), \min(1, x + z_{3\times 0.05}))$ and $(\max(-1, x - \sqrt{\chi^2_{1,2\times 0.05} - 2\log H}), \min(1, x + \sqrt{\chi^2_{1,2\times 0.05} - 2\log H}))$ respectively. In these two cases, we shorten the standard and the likelihood ratio intervals such that their

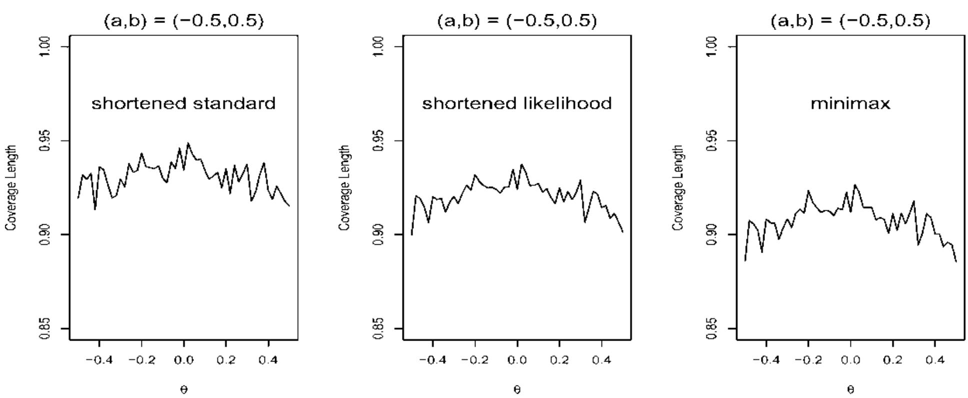

Figure 6. The expected lengths of the shortened standard, the shortened likelihood ratio and the minimax level 0.95 intervals with respect to different parameter spaces $(-0.5, 0.5)$.

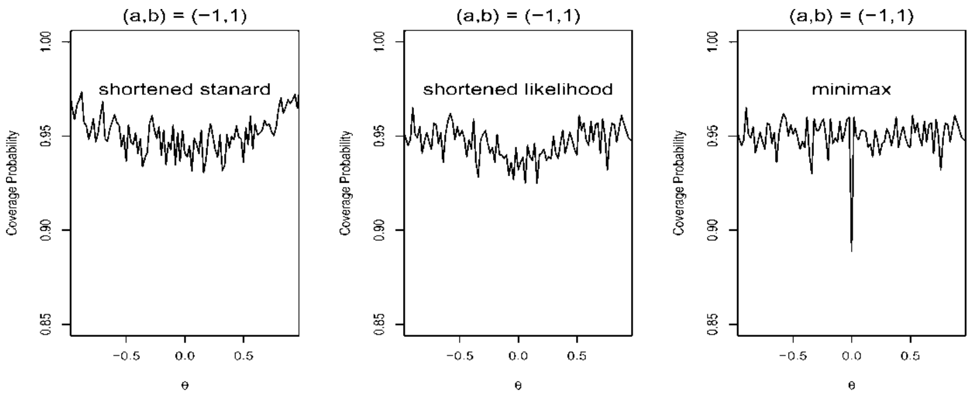

Figure 7. The coverage probabilities of the shortened standard, the shortened likelihood ratio and the minimax level 0.95 intervals with respect to different parameter spaces (−1, 1).

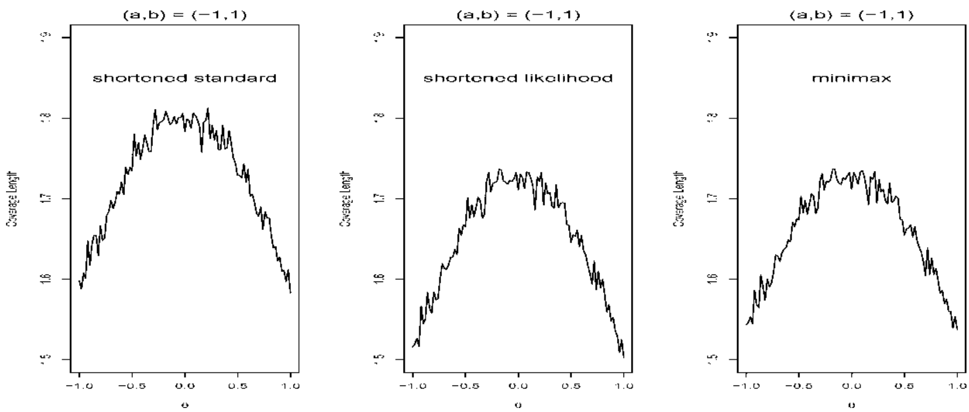

Figure 8. The expected lengths of the shortened standard, the shortened likelihood ratio and the minimax level 0.95 intervals with respect to different parameter spaces (−1, 1).

coverage probabilities are close to that of the minimax interval. The expected length of the standard interval is larger than those two intervals. From these plots, the performance of the shortened likelihood ratio interval is similar to the minimax interval in both coverage probability and expected length. Although the performance of the shortened likelihood interval is similar to the minimax interval, the modification depends on the parameter space, therefore, it is less useful in application. Compared to the shortened standard and the shortened likelihood intervals, the minimax interval is still a better choice among these three intervals.

4. Example

The comparison of the five intervals is illustrated by a real-data example, about testing the expected retirement age of civil servants. The data are the retirement ages of civil servants working for the Taiwan government in 2002. The retirement payment depends on the retirement age. There are two kinds of retirement payment systems for these civil servants: (1) a civil servant can continue receiving almost 85% salary every month after retirement if he retires after age 50; or (2) a civil servant can have a lump sum payment when he retires. Note that if a civil servant retires after 50, he can choose either the first or the second systems; however, if he retires before 50, he can only choose the second system. Most civil servants tend to choose the first system because the total amount of money paid out in the first system is higher than in the second one. However, some people will still choose to retire before 50 for personal reasons. According to the law, the retirement age can not exceed 65. From empirical knowledge, most civil servants' retirement ages are between 50 and 60. Therefore, we can choose 45 or 50 as a lower bound and 60 or 65 as an upper bound for the parameter space of the average age.

Example 1 We collected the data of civil servants' retirement ages from government records. There were a total of 9055 civil servants retiring in 2002. In this example, the population is the whole data set, *i.e.* 9055 civil servants' retirement ages. The mean of these ages is 53.31. The variance is 81.45. Suppose that the researchers do not have the whole data set, but the variance is known from empirical knowledge. We assume that retirement age follows a normal distribution $N(\theta, \sigma^2)$.

Since the mean of the whole data set is 53.31, we conduct simulation to investigate the five confidence intervals. We choose samples of five ages from the whole data set with 1000 replicates, and calculate the coverage probability when θ is the true mean 53.31. The expected length of the five intervals are also calculated. The coverage probabilities and the expected lengths of the five 0.95 intervals with respect to different bounded parameter spaces are shown in tables 1–4.

Note that before computing the intervals, the retirement data are normalized such that the variance of the data is 5 because the sample size is five. The expected lengths in tables 1–4 are the expected lengths of the confidence intervals of the normalized data. In this example,

Table 1. The coverage probabilities and the expected lengths of the five 0.95 intervals when the bounded parameter space is (45,65).

	Standard	Minimax	Likelihood	Credible	rp
Coverage probability	0.952	0.9382	0.952	0.946	0.959
Expected length	3.317	3.368	3.323	3.053	3.368

Table 2. The coverage probabilities and the expected lengths of the five 0.95 intervals when the bounded parameter space is (45,60).

	Standard	Minimax	Likelihood	Credible	rp
Coverage probability	0.955	0.937	0.957	0.969	0.969
Expected length	2.815	2.992	2.821	2.545	2.688

Table 3. The coverage probabilities and the expected lengths of the five 0.95 intervals when the bounded parameter space is (50,65).

	Standard	Minimax	Likelihood	Credible	rp
Coverage probability	0.954	0.941	0.987	0.957	0.985
Expected length	2.482	2.250	2.528	2.336	2.453

Table 4. The coverage probabilities and the expected lengths of the five intervals when the bounded parameter space is (50,60).

	Standard	Minimax	Likelihood	Credible	rp
Coverage probability	0.950	0.944	0.994	0.993	0.996
Expected length	1.986	1.881	2.029	1.861	1.912

the minimax interval has the smallest coverage probabilities for all four cases, at less than 0.95. However, the coverage probability of the minimax interval should be greater than 0.95 if the replicates are large enough and the data really follows a normal distribution. The small coverage probability may be due to the fluctuation of the simulation and small sample size.

In the these cases, the performance of rp interval is better than the other intervals because it has higher coverage probability and shorter expected length.

5. Conclusions

The purpose of this article is to propose improved confidence intervals for the mean of a normal distribution when the parameter space is restricted. Five intervals, the standard interval, minimax interval, likelihood ratio interval, Bayesian credible interval and rp interval, are discussed.

In simulation studies, the likelihood ratio interval has higher coverage probability than the standard interval, but it also has a longer expected length. The performances of the Bayesian interval and rp interval are more similar than the performances of the other two intervals. Although the credible interval has the advantage of shortest expected length in most cases, it does not have an explicit form, which needed to be derived from numerical calculations. The minimax interval performs better than the standard interval because it has shorter expected length.

Acknowledgements

The author thanks Professor T. Tony Cai at University of Pennsylvania for helpful discussions. The author thanks the associate editor and the referee for helpful comments.

References

[1] Mandelkern, M., 2002, Setting confidence intervals for bounded parameters. *Statistical Science*, **17**, 149–172.
[2] Feldman, G.J. and Cousins, R.D., 1998, Unified approach to the classical statistical analysis of small signals. *Physical Review* D, **57**, 3873–3889.
[3] Roe, B.P. and Woodroofe, M.B., 2001, Setting confidence belts. *Physical Review* D, **60**, 3009–3015.
[4] Zhang, T. and Woodroofe, M., 2003, Credible and confidence sets for restricted parameter spaces. *Journal of Statistical Planning and Inference*, **115**(2), 479–490.
[5] Cox, D.R. and Hinkley, D.V., 1974, *Theoretical Statistics* (London: Chapman and Hall).
[6] Evans, S.N., Hansen, B.B. and Stark, P.B., 2005, Minimax expected measure confidence sets for restricted location parameters. *Bernoulli*, **11**, 571–590.
[7] Pratt, J.W., 1961, Length of confidence intervals. *Journal of the American Statistical Association*, **56**, 549–567.
[8] Wang, H., 2006, Modified *P*-value of two-sided test for normal distribution with restricted parameter space. *Communication of Statistics – Theory and Method*, **35**, 1–14.