Advances and Applications in Statistics
Volume 33, Number 1, 2013, Pages 1-21
Published Online: March 2013
Available online at http://pphmj.com/journals/adas.htm
Published by Pushpa Publishing House, Allahabad, INDIA

EFFICIENT METHODS ON CONFIDENCE INTERVALS FOR PREDICTION INTERVALS

Mei Ling Huang*, Wai Kong Yuen and Miao Zhang

Department of Mathematics
Brock University
St. Catharines, Ontario
Canada L2S 3A1
e-mail: mhuang@brocku.ca

Abstract

Prediction intervals have many applications in economics, business, health studies, engineering, science and social science. In this article, we study confidence intervals for prediction intervals of the future value of a random variable based on quantile estimations. There are theoretical difficulties in this problem with few methods in the literature. We propose three methods based on weight functions combined with bootstrapping. The results of Monte Carlo simulations confirm that the proposed methods improve the efficiencies in mean square error and probability coverage relative to the existing methods. We also study a real world example which shows the improved efficiencies of the proposed methods.

2010 Mathematics Subject Classification: Primary 62G15; Secondary 62G05.

Keywords and phrases: bootstrapping, efficiency, HD estimator, kernel density estimation, *L*-statistics, probability coverage, quantile estimation, weighted empirical distribution function.

This research is supported by the Natural Sciences and Engineering Research Council of Canada.

*Corresponding author

Communicated by K. K. Azad

Received December 18, 2012

1. Introduction

A quantile estimator from a random sample is usually used to estimate the corresponding population quantile. It also plays a role in estimating the end points of an interval for predicting the value of another, independent, random variable drawn from the sampled population. Formally, we consider a random variable X with c.d.f. $F(x)$, $x \in \Re$, which is continuous and unknown. Thus we are interested in a future value of X being within an interval $[c, d]$ with probability $(1-\gamma)$, $0 < \gamma < 1$.

Definition 1.1. An interval $[c, d]$ (constants $c, d \in \Re$) is a *true* $(1-\gamma)100\%$ *prediction interval* of a real random variable X if

$$P(c \leq X \leq d) = 1 - \gamma, \quad 0 < \gamma < 1. \tag{1.1}$$

A simple prediction interval can be created by using the q_1th and $(1-q_2)$th population quantiles x_{q_1}, x_{1-q_2} for the end points c and d, respectively, so that

$$P(x_{q_1} \leq X \leq x_{1-q_2}) = 1 - \gamma, \quad 0 < \gamma < 1$$

where $c = x_{q_1}$, $d = x_{1-q_2}$, $q_1 + q_2 = \gamma$, $0 < q_i < \gamma$, $i = 1, 2$.

It is important to estimate the prediction interval given a random sample. By letting $q_1 = q_2 = \gamma/2$, we can estimate the true $(1-\gamma)100\%$ prediction interval by using point estimators $\hat{x}_{\gamma/2}$ and $\hat{x}_{1-\gamma/2}$ of the population quantiles $x_{\gamma/2}$ and $x_{1-\gamma/2}$ from a random sample $X_1, X_2, \ldots, X_n$ of size n from $F(x)$. Then $[\hat{x}_{\gamma/2}, \hat{x}_{1-\gamma/2}]$ is a point estimator for the true prediction interval.

There are a number of different approaches on the study of prediction intervals in the literatures. For example, exact nonparametric prediction intervals for the future value of a random variable were given in David and Nagaraja [2], Volterman and Balakrishnan [21]. Hall and Maiti [7] constructed bias-corrected estimators of mean squared error for computing prediction regions. Kim et al. [12] also used the bias-corrected bootstrap for interval forecasting on time series. Nolan and Ravishanker [14] proposed a method using simultaneous prediction intervals for time series. A number of authors studied the applications of prediction intervals. For example, Brodin and Rootzėn [1] used univariate and bivariate generalized Pareto distribution models to predict extreme wind storm losses. Vidoni [20] used a predictive likelihood method to predict future River Nidd annual maxima.

In this article, we are further interested in obtaining a confidence interval for the true prediction interval. For example, based on the current sample, we would like to be 95% $((1-\alpha)100\%)$ certain to predict that the quality of products by a manufacturer will be 90% $((1-\gamma)100\%)$ within standard limits. This notion can be applied to science, social science, engineering, economics, health studies and business. The example in Section 5 of this paper concludes statistically that we are 90% certain to predict that 80% of the IQ scores of children of age five whose mothers are in a non-depressed state in the UK are between 86.5 and 137.5. Formally, we define a confidence interval for a true prediction interval as follows.

Definition 1.2. An interval estimator $[a, b]$ (random variables $a, b \in \Re$) is a $(1-\alpha)100\%$ confidence interval for the $(1-\gamma)100\%$ true prediction interval of a random variable *X* if

$$P(a \leq x_{\gamma/2} \ \& \ x_{1-\gamma/2} \leq b) = 1-\alpha, \quad 0 < \gamma < 1, \quad 0 < \alpha < 1. \qquad (1.2)$$

Our goal is to find a good estimator $[a, b]$ given a random sample. This problem is obviously related to computing confidence intervals for quantiles. Some existing methods for quantile estimations have been reviewed by Hutson and Ernst [11], Ernst and Hutson [6], and Wilcox [18]. While these methods perform reasonably well, they cannot be directly applied to relatively extreme quantiles, which is our main interest, due to some theoretical difficulties. In fact, the discussions about the formation of confidence intervals in the literature mostly use normal approximations for the sampling distributions of the estimates. This approximation, however, does not work well when the sample size is relatively small. On the other hand, if we make the additional assumption that *F* is known, there are various semi-parametric bias-reduction methods that provide estimations with good asymptotic properties. For some recent work in this direction, see e.g., Diebolt et al. [3], where the asymptotic normality of the extreme quantile of a Weibull tail-distribution was established and El Methni et al. [5], where similar results for Pareto-type and Weibull tail-distributions were derived.

There are few methods in the literature on confidence intervals for prediction intervals, which were first considered in Wilcox [19], where method L was introduced to derive such confidence intervals. The approach is based on taking a simple variation of (1.2) by choosing

$$a = \hat{x}_{\gamma/2} - z_{1-\alpha/2}\hat{\sigma},$$

$$b = \hat{x}_{1-\gamma/2} + z_{1-\alpha/2}\hat{\sigma}, \tag{1.3}$$

where $\hat{x}_{\gamma/2}$ and $\hat{x}_{1-\gamma/2}$ are point estimators of $x_{\gamma/2}$ and $x_{1-\gamma/2}$, respectively, $\hat{\sigma}$ is an estimator of the standard error of the quantile estimator $\hat{x}_{\gamma/2}$, and $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$th quantile of the standard normal distribution. Clearly, the success of this approach depends on the choices of the estimators for the quantiles and their standard errors, and how well the normality assumption holds.

The main objective of this article is to improve method L by considering other choices of the two types of estimators in (1.3). First, we attempt to improve the quantile estimation. Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ denote the order statistics of a random sample $X_1, X_2, ..., X_n$ from $F(x)$. It is well known that the simplest *p*th quantile estimator is the *p*th sample quantile.

Definition 1.3. The *p*th sample quantile is given by

$$SQ_p = X_{([np+0.5])}, \quad 0 < p < 1, \tag{1.4}$$

where $[np + 0.5]$ denotes the integer part of $(np + 0.5)$.

Examples of sample quantiles include the median and quartiles. Based on (1.4), the interquartile range (*IQR*) can be written as

$$IQR = SQ_{0.75} - SQ_{0.25}. \tag{1.5}$$

The IQR is a statistic that measures the spread of the distribution about the median. The variability of individual order statistics is a weakness of the sample quantiles in (1.4). To reduce this variability, we consider using L-statistics which are linear combinations of the order statistics, instead of the sample quantile SQ_p. We also consider weighted methods which can improve the estimations of extreme quantiles.

Second, we attempt to improve the estimation of the standard error of the quantile estimators. Method L in Wilcox [19] is based on an adaptive kernel density estimator for estimating the standard error of the quantile estimator. It is a useful method that provides satisfactory estimates, even when the sample size is small. The main weakness of this approach is that there are still theoretical difficulties for relatively extreme quantiles, in which case the probability coverages perform poorly. To address this problem, we consider the application of Huang and Brill [10]'s level crossing empirical distribution function (e.d.f.), and the corresponding weighted kernel density estimator. This estimator is shown to be more efficient on the tail of the distribution relative to the regular kernel density estimator and it is particularly efficient in the small sample case. We also use bootstrapping methods to estimate the standard errors for L-statistics quantile estimators. Based on these ideas, we propose weighted estimation methods combined with bootstrapping to improve the estimation of the standard error of quantile estimators.

In Section 2, we review method L for confidence intervals for prediction intervals. In Section 3, we list some quantile estimation methods, review the weighted level crossing kernel density estimator, and propose three new estimation methods for obtaining confidence intervals for prediction intervals. In Section 4, we describe the Monte Carlo simulation methods, and show the numerical results of the simulation efficiencies of these methods. The simulation results show that the efficiencies of our proposed methods relative to method L are greater than one in most situations. Finally, in Section 5, we apply these methods to a real world example, in which we analyze a data set from the IQ scores of children from non-depressed mothers. We found that the new methods give better predictions than method L.

2. Review of Method L

Typical approaches for computing confidence intervals of the form (1.2) break down when extreme quantiles are involved. For example, if we consider $(X_{(i)}, X_{(j)})$ as a confidence interval for the *p*th quantile, it is well known that the exact probability coverage of this interval is unlikely to reach a level close to $1 - \alpha$ even for $\alpha = 0.05$, unless the sample size is very large. To address this issue, Wilcox proposed method L, a simple approach that is based on (1.4) with the choice of sample quantiles as the estimators for the true quantiles, i.e., $\hat{x}_{\gamma/2} = SQ_{\gamma/2}$ and $\hat{x}_{1-\gamma/2} = SQ_{1-\gamma/2}$. More importantly, the confidence interval involves estimating the standard error of the sample quantiles and the use of a normal approximation. This extra flexibility allows one to form a confidence interval with high enough probability coverages.

To estimate the corresponding standard error $\hat{\sigma}$, we consider a classic expression (see e.g. Wilcox [18]) for the asymptotic squared standard error of the sample quantile SQ_p, given by

$$\sigma_p^2 \approx \frac{p(1-p)}{n[f^2(x_p)]}, \quad 0 < p < 1, \tag{2.1}$$

where $f(x_p)$ is the p.d.f. of *X* at the true quantile x_p. An estimator of $f(x_p)$ was derived based on an adaptive kernel density estimator (Silverman [17]) of the form

$$\hat{f}(t) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\lambda_i}K\{h^{-1}\lambda_i^{-1}(t - X_i)\}, \tag{2.2}$$

where *K* is the Epanechnikov kernel. The bandwidth is given by

$$h = \frac{1.06}{n^{1/5}}\min(s, IQR/1.34), \tag{2.3}$$

where IQR is given in (1.5), s is the sample standard deviation of

$X_1, \ldots, X_n$, and $\lambda_1, \ldots, \lambda_n$ are some adaptive constants (see Wilcox [19] for details). This leads to the following $(1-\alpha)100\%$ confidence interval of the $(1-\gamma)100\%$ prediction interval of a random variable X, $0 < \gamma < 1$, $0 < \alpha < 1$:

$$(SQ_{\gamma/2} - z_{1-\alpha/2}\, s_L,\ SQ_{1-\gamma/2} + z_{1-\alpha/2}\, s_L), \tag{2.4}$$

where SQ_p is defined in (1.4), and

$$s_L^2 = \widehat{\sigma_{\gamma/2}^2} = \frac{\frac{\gamma}{2}\left(1 - \frac{\gamma}{2}\right)}{n\widehat{f^2}(SQ_{\gamma/2})}. \tag{2.5}$$

Wilcox [19] compared method L with two other proposed methods and indicated that method L is the only satisfactory method when the sample size is 30. Clearly, the efficiency of the method depends heavily on the choice of the quantile estimators and methods of estimating their standard errors. We will investigate whether other estimation methods will improve the efficiency of this approach.

3. New Estimation Methods

In this section, we propose three new methods to improve method L, based on the applications of other quantile estimators and estimation methods of standard errors.

3.1. Method weighted L

Our first approach to improve method L is to consider another weighted kernel density estimator instead of the adaptive kernel density estimator in (2.2). Since the performance of method L is only satisfactory for relatively extreme quantiles, we consider the application of level crossing kernel density estimators (Huang and Brill [10]). Formally, we first define the level crossing e.d.f. as

$$F_{n(W)}(x) = \sum_{i=1}^{n} I_{(-\infty,\, x)}(X_{(i)})\, w_{i,n}, \tag{3.1}$$

where I_A is the indicator function of set A and

$$w_{i,n} = \begin{cases} \dfrac{1}{\sqrt{n(n-1)}}, & \text{when } i = 2, 3, \ldots, n-1, \\ \dfrac{1}{2}\left\{1 - \dfrac{n-2}{\sqrt{n(n-1)}}\right\}, & \text{when } i = 1, n. \end{cases} \tag{3.2}$$

Note that the level crossing e.d.f. puts less weight on the smallest and largest data points. Huang and Brill [10] showed that this estimator has better efficiency than the classical e.d.f., especially in the tail of the distribution when the sample size is small. This motivates the use of the level crossing kernel density estimator with a normal kernel, defined as

$$\widehat{f_W}(x) = \frac{1}{h}\sum_{i=1}^{n} w_{i,n}\phi\left(\frac{x - X_{(i)}}{h}\right), \tag{3.3}$$

where $w_{i,n}$ is given in (3.2), ϕ is the density of a standard normal distribution and we choose the bandwidth h as in (2.3) for simplicity. Combining with (2.1), we use the following estimator for the squared standard error:

$$s_{L(W)}^2 = \frac{\frac{\gamma}{2}\left(1 - \frac{\gamma}{2}\right)}{n\widehat{f_W^2}(SQ_{\gamma/2})}. \tag{3.4}$$

These results in the following $(1-\alpha)100\%$ confidence interval for the $(1-\gamma)100\%$ prediction interval of a random variable X, $0 < \gamma < 1$, $0 < \alpha < 1$:

$$(SQ_{\gamma/2} - z_{1-\alpha/2}\, s_{L(W)},\ SQ_{1-\gamma/2} + z_{1-\alpha/2}\, s_{L(W)}), \tag{3.5}$$

where SQ_p is given in (1.4).

3.2. Method HD-boot

Our second approach is based on the use of another quantile estimator instead of the sample quantile SQ_p in (1.4). Due to the variability of individual order statistics, it is not surprising that method L, which depends heavily on the sample quantiles, does not perform very well when the sample size is small. An alternative approach is to consider *L*-statistics of the form

$$T_n = \sum_{i=1}^{n} c_i X_{(i)}, \tag{3.6}$$

for some constants $c_1, ..., c_n$. In particular, we consider a non-kernel based *L*-quantile estimator introduced by Harrell and Davis [8].

Definition 3.1. The HD_p estimator for x_p is defined as

$$HD_p = \sum_{i=1}^{n} \left[\int_{(i-1)/n}^{i/n} \frac{1}{\beta((n+1)p, (n+1)q)} y^{(n+1)p-1} (1-y)^{(n+1)q-1} dy \right] X_{(i)}, \tag{3.7}$$

where $0 < p < 1$, $q = 1 - p$, and $\beta(s, t)$ is the beta function with parameters s, t.

Sheather and Marron [16] indicated that HD_p performs as well as other *L*-quantile estimators when the sample size is large, with good asymptotic properties. Furthermore, it avoids the problem of selections of kernels or bandwidths in other kernel based estimators, e.g., Parzen [15]. In this approach, we use HD_p as the quantile estimator.

Due to the complexity of HD_p, it is not straightforward to obtain a close-form estimate for the standard error of HD_p. Instead, we will estimate the standard error using the bootstrap method (Efron and Tibshirani [4]).

Definition 3.2. Let $X_1^*, ..., X_n^*$ be a bootstrap sample obtained by resampling with replacement from *n* observations $X_1, ..., X_n$. Let $\widehat{x_p}^*$ be the estimate of the *p*th quantile based on this bootstrap sample. Repeat this procedure *B* times, yielding estimates $\widehat{x_{p1}}^*, ..., \widehat{x_{pB}}^*$. Then a bootstrap estimate of the squared standard error of $\widehat{x_p}$ is

$$s_{boot}^2 = \frac{1}{B-1}\sum_{i=1}^{B}(\widehat{x_{pi}}^* - \overline{x_p^*})^2, \quad 0 < p < 1, \tag{3.8}$$

where $\overline{x_p^*} = \frac{1}{B}\sum_{i=1}^{B}\widehat{x_{pi}}^*$.

The resulting $(1-\alpha)100\%$ confidence interval for the $(1-\gamma)100\%$ prediction interval of a random variable X, $0 < \alpha < 1$, $0 < \gamma < 1$, based on HD_p in (3.7) is given by

$$(HD_{\gamma/2} - z_{1-\alpha/2}\, s_{boot(HD)},\ HD_{1-\gamma/2} + z_{1-\alpha/2}\, s_{boot(HD)}), \tag{3.9}$$

where

$$s_{boot(HD)}^2 = \frac{1}{B-1}\sum_{i=1}^{B}(\widehat{x_{pi_{HD}}}^* - \overline{x_{p_{HD}}^*})^2, \quad p = \gamma/2, \tag{3.10}$$

is the estimate of the squared standard error of HD_p in (3.7).

3.3. Method weighted HD-boot

Our third approach is to combine the first two approaches by using the weighted e.d.f. in (3.1) and HD_p in (3.7). The key observation is that HD_p can be written as

$$HD_p = \frac{1}{\beta((n+1)p, (n+1)q)} \times \int_0^1 F_n^{-1}(y)\, y^{(n+1)p-1}(1-y)^{(n+1)q-1} dy, \tag{3.11}$$

where

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{(-\infty, x]}(X_i), \quad x \in \Re,$$

is the classical e.d.f.. While the classical e.d.f. puts equal weights on the ith order statistic $X_{(i)}$, the form of (3.11) allows us to easily generalize HD_p to a wider class of estimators by using weighted e.d.f. in place of F_n. We follow the approach in Huang [9] by replacing F_n with the level crossing e.d.f. $F_{n(W)}$ defined in (3.1), which leads to the following level crossing pth quantile estimator:

$$HD_{p(W)} = \frac{1}{\beta\{(n+1)p,\ (n+1)q\}} \int_0^1 F_{n(W)}^{-1}(y)\, y^{(n+1)p+1}(1-y)^{(n+1)q-1} dy$$

$$= \sum_{i=1}^{n} \left[\int_{p_{i-1,n}}^{p_{i,n}} \frac{y^{(n+1)p+1}(1-y)^{(n+1)q-1}}{\beta\{(n+1)p,\ (n+1)q\}} dy \right] X_{(i)}, \tag{3.12}$$

where $0 < p < 1$, $p_{i,n} = \sum_{j=1}^{i} w_{j,n}$, $i = 1, \ldots, n$, $w_{j,n}$ is given in (3.1), and $p_{0,n} \equiv 0$. In that work, it was also shown that from both the theoretical and computational point of view, $HD_{p(W)}$ is more efficient than HD_p in (3.7) in many cases. As in method HD-boot, we estimate the standard error of $HD_{p(W)}$ by the bootstrap method. The resulting $(1-\alpha)100\%$ confidence interval for the $(1-\gamma)100\%$ prediction interval of a random variable X, $0 < \alpha < 1$, $0 < \gamma < 1$, is given by

$$\left(HD_{\gamma/2(W)} - z_{1-\alpha/2}\, s_{boot(HD_W)},\ HD_{(1-\gamma/2)(W)} + z_{1-\alpha/2}\, s_{boot(HD_W)}\right), \tag{3.13}$$

where

$$s^2_{boot(HD_W)} = \frac{1}{B-1} \sum_{i=1}^{B} \left(\widehat{x_{pi_{HD(W)}}}^{*} - \overline{x^{*}_{p_{HD(W)}}} \right)^2, \quad p = \gamma/2, \tag{3.14}$$

is the squared standard error of $HD_{p(W)}$ in (3.12).

4. Simulation Studies

4.1. Simulation efficiencies of the new estimators

The simulation of estimating the confidence intervals for prediction intervals is based on four cases of g-and-h distributions, which were considered in Wilcox [19]:

Definition 4.1. A random variable X has the g-and-h distribution if

$$X = \begin{cases} \dfrac{\exp(gZ) - 1}{g} \exp(hZ^2/2), & g > 0; \\ Z \exp(hZ^2/2), & g = 0, \end{cases} \tag{4.1}$$

where Z is a standard normal random variable and g, h are nonnegative parameters that determine the first four moments.

We choose g and h to have some characteristics of interest:

(i) The standard normal distribution $(g = h = 0)$;

(ii) A symmetric heavy-tailed distributions $(h = 0.2,\ g = 0)$;

(iii) An asymmetric distribution with relatively light tails $(h = 0,\ g = 0.2)$;

(iv) An asymmetric distribution with heavy tails $(g = h = 0.2)$.

For each of the above distributions, we simulate $m = 1000$ random samples of size $n = 30$, and use them to compute $(1 - \alpha)100\%$ confidence intervals for $(1 - \gamma)100\%$ prediction intervals using the four estimation methods: L, weighted L, HD-boot, weighted HD-boot, where we choose $\alpha = \gamma = 0.05$ (so that $p = 0.025$). For method HD-boot and method weighted HD-boot, we choose $B = 100$ as the number of resampling in calculating the bootstrap standard error. For each method, we calculate the simulation mean square error (*SMSE*), which is defined by

$$SMSE(s_{est}^2) = \frac{1}{m} \sum_{i=1}^{m} s_{est(i)}^2, \tag{4.2}$$

where $s_{est(i)}$ is the standard error of the estimator for the ith sample, $i = 1, \ldots, m$. The root simulation efficiency of *SMSE* for the proposed methods relative to method L is defined by

$$RSEFF(s_{proposed}) = \sqrt{\frac{SMSE(s_L^2)}{SMSE(s_{proposed}^2)}}, \tag{4.3}$$

where $SMSE(s_{proposed}^2)$ and $SMSE(s_L^2)$ are defined in (4.2).

We also study the performance of the probability coverages $(\hat{\alpha})$ of the confidence intervals based on the bias from the true level α $(= 0.05)$. The simulation value of the bias of an estimation method is defined by

$$SBias(\hat{\alpha}) = \frac{1}{m}\sum_{i=1}^{m}[\hat{\alpha}_{(i)} - \alpha]^2, \tag{4.4}$$

where $\hat{\alpha}_{(i)}$ is the value for the *i*th sample, $i = 1, \ldots, m$, and α $(= 0.05)$ is the true value. The simulation efficiency of probability coverages for $\hat{\alpha}_{proposed}$ relative to method L is defined by

$$SEFF(\hat{\alpha}_{proposed}) = \frac{SBias(\hat{\alpha}_L)}{SBias(\hat{\alpha}_{proposed})}. \tag{4.5}$$

4.2. Simulation results

The simulation results are shown in Figure A1, Figure A2 and Table A1 in Appendix A. Figure A1 shows the box-plots of the standard errors of the estimators. We can see that method weighted HD-boot has the smallest standard error in the four cases while method HD-boot and method weighted L also outperform method L. Figure A2 shows the box-plots of $\hat{\alpha}$. Again, all three proposed methods have smaller biases in probability coverage than the method L, with method weighted HD-boot achieving the highest simulation efficiency. Table A1 shows the results of the simulation efficiencies defined in formulas (4.3) and (4.5) in Section 4.1. All three proposed methods have 100% (24 out of 24 cases) efficiency greater than 1 relative to method L. Overall, the method weighted HD-boot is the most efficient method. All the results are computed using C++ programs.

5. An Example

Table 5.1. IQ scores of children of age five whose mothers are in non-depressed state

obs.	IQ score	obs.	IQ score	obs.	IQ score	obs.	IQ score
1	103	11	117	21	123	31	103
2	124	12	89	22	118	32	118
3	124	13	125	23	117	33	117
4	104	14	127	24	141	34	115
5	92	15	112	25	124	35	119
6	124	16	48	26	110	36	117
7	99	17	139	27	98	37	92
8	92	18	118	28	109	38	101
9	116	19	106	29	120	39	119
10	99	20	117	30	127	40	144

In the real world, there is a wide range of applications of the confidence intervals for prediction intervals in science, social science, economics and business. A good estimation method gives more accuracy in predicting the performance of random variables in the future. We consider a data set listed in Table 5.1 from health studies provided by Dr. Channi Kumar, Department of Psychiatry, Institute of Psychiatry, London, UK (Kumar [13]). It consists of the IQ scores of 40 randomly selected children of age five, whose mothers are in a non-depressed state.

We compare the performance of 90% confidence intervals $([a, b])$ for the 80% prediction interval for the IQ score, of the four methods. We study two different types of efficiencies of the estimation methods relative to method L. The first efficiency is based on the estimated standard errors:

$$EFF_{proposed}(s) = \frac{s_L}{s_{proposed}}. \tag{5.1}$$

The second efficiency is based on the ratio of the lengths of the confidence intervals:

$$EFF_{proposed}(CI) = \frac{|\, b_L - a_L \,|}{|\, b^*_{proposed} - a^*_{proposed} \,|}, \tag{5.2}$$

where a^*, b^* are the values based on the proposed new estimation methods. The results are summarized in Table 5.2.

Table 5.2. Efficiencies of the estimation methods for the IQ scores

Method	*s*	*EFF*(*s*)	Confidence Interval (C.I.)	*EFF*(*CI*)
L	5.2505	1	(83.3656, 135.6344)	1
Weighted L	4.5368	**1.1573**	(84.5393, 134.4607)	**1.0469**
HD-boot, *B* = 100	4.6924	**1.1189**	(85.0936, 138.7569)	0.9739
Weighted HD-boot, *B* = 100	4.3137	**1.2171**	(86.5262, 137.5550)	**1.0243**

In Table 5.2, all three proposed methods give smaller estimated standard errors than that of method L. In terms of the lengths of the confidence intervals, both method weighted L and method weighted HD-boot are more efficient than method L. In particular, method weighted HD-boot gives the smallest standard error and the second shortest confidence interval. The corresponding efficiencies of method weighted HD-boot relative to method L are $EFF_{HD-boot(W)}(s) = 1.2171$ and $EFF_{HD-boot(W)}(CI) = 1.0243$.

Finally, Figure 5.1 plots the values of the data set and the confidence intervals from both method L and method weighted HD-boot. The 90% confidence interval for 80% prediction interval from method HD-boot is (86.5262, 137.5550). Note that it is both shorter and fits the majority of the data better than that from method L.

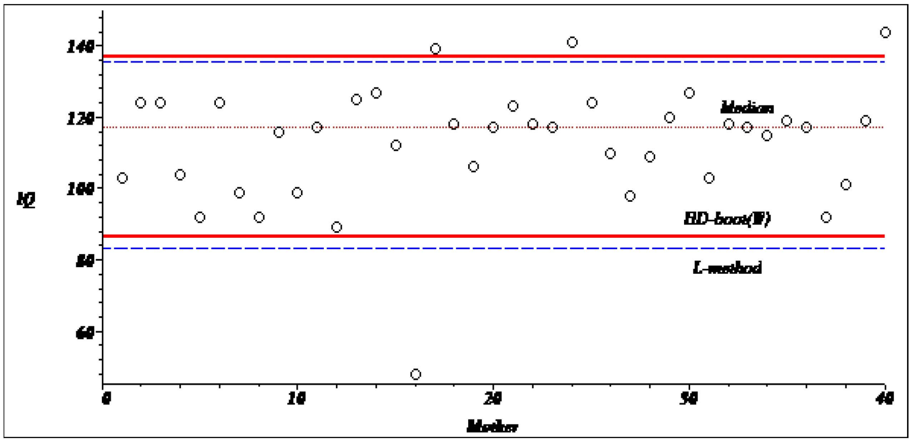


Figure 5.1. The 90% confidence interval for the 80% prediction interval for the IQ scores based on 40 children of age five. Blue dash: *method L*; Red solid: *method weighted HD-boot*; Brown dot: *median*.

6. Conclusions

We proposed three methods of deriving confidence intervals for prediction intervals. From the results of the simulation studies all three methods show improvements of efficiencies on MSE and probability coverages. In the example, we demonstrate that our methods provide shorter confidence intervals relative to method L. The idea of using HD_p, weighted e.d.f. and bootstrapping provides more efficient estimation methods to this field.

References

[1] E. Brodin and H. Rootzén, Univariate and bivariate GPD methods for prediction extreme wind storm losses, Insurance Math. Econom. 44 (2009), 345-356.

[2] H. A. David and H. N. Nagaraja, Order Statistics, 3rd ed., Wiley, New York, 2003.

[3] J. Diebolt, L. Gardes, S. Girard and A. Guillou, Bias-reduced extreme quantiles estimators of Weibull tail-distributions, J. Statist. Plann. Inference 138 (2008), 1389-1401.

[4] B. Efron and R. Tibshirani, An Introduction to the Bootstrap, Chapman and Hall, New York, 1993.

[5] J. El Methni, L. Gardes, S. Girard and A. Guillou, Estimation of extreme quantiles from heavy and light tailed distributions, J. Statist. Plann. Inference 142 (2012), 2735- 2747.

[6] M. D. Ernst and A. D. Hutson, Utilizing a quantile function approach to obtain exact bootstrap solutions, Statist. Sci. 18 (2003), 231-240.

[7] P. Hall and T. Maiti, On parametric bootstrap methods for small area prediction, J. R. Stat. Soc. Ser. B 68(2) (2006), 221-238.

[8] F. E. Harrell and C. E. Davis, A new distribution-free quantile estimator, Biometrika 69(3) (1982), 635-640.

[9] M. L. Huang, On a distribution-free quantile estimator, Comput. Statist. Data Anal. 37 (2001), 477-486.

[10] M. L. Huang and P. H. Brill, A distribution estimation method based on level crossings, J. Statist. Plann. Inference 124(1) (2004), 45-62.

[11] A. D. Hutson and M. D. Ernst, The exact bootstrap mean and variance of an *L*-estimator, J. Roy. Statist. Soc. B Part 1 62 (2000), 89-94.

[12] J. H. Kim, H. Song and K. K. F. Wong, Bias-corrected bootstrap prediction intervals for autogressive model: new alternative with applications to tourism forecasting, J. Forecast. 29 (2010), 655-672.

[13] R. Kumar, Anybody's child: severe disorders of mother-to-infant bonding, Br. J. Psychiatry 171 (1997), 175-181.

[14] J. P. Nolan and N. Ravishanker, Simultaneous prediction intervals for ARMA processes with stable innovations, J. Forecasting 28(3) (2009), 235-246.

[15] E. Parzen, Nonparametric statistical data modeling, J. Amer. Statist. Assoc. 74 (1979), 105-131.

[16] S. J. Sheather and J. S. Marron, Kernel quantile estimators, J. Amer. Statist. Assoc. 85 (1990), 410-416.

[17] B. W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, New York, 1986.

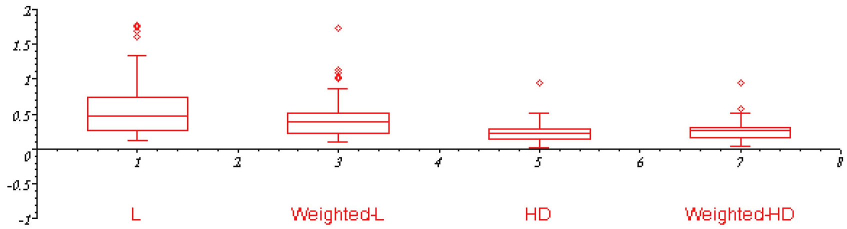
[18] R. R. Wilcox, Introduction to Robust Estimation and Hypothesis Testing, 2nd ed., Academic Press, San Diego, 2005.

[19] R. R. Wilcox, Confidence intervals for prediction intervals, J. Appl. Statist. 33(3) (2006), 317-326.

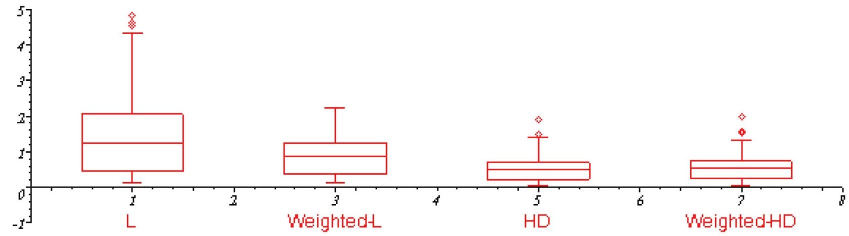
[20] P. Vidoni, Improved prediction intervals and distribution function, Scan. J. Statist. 36 (2009), 735-748.

[21] W. Volterman and N. Balakrishnan, Exact nonparametric confidence, prediction and tolerance intervals based on multi-sample Type-II right censored data, J. Statist. Plann. Inference 140 (2010), 3306-3316.

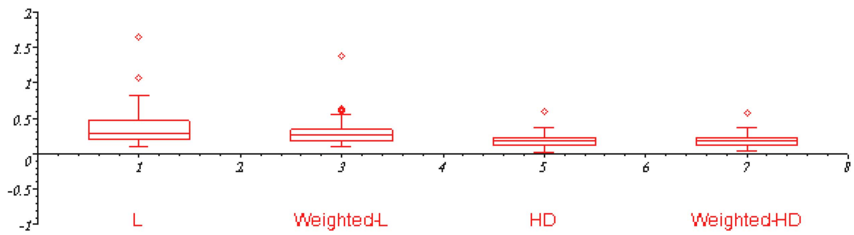
Appendix A. Simulation Efficiencies



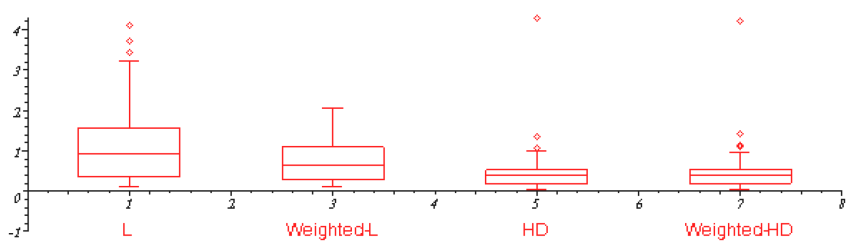
(a) $g = 0,\ h = 0$



(b) $g = 0,\ h = 0.2$

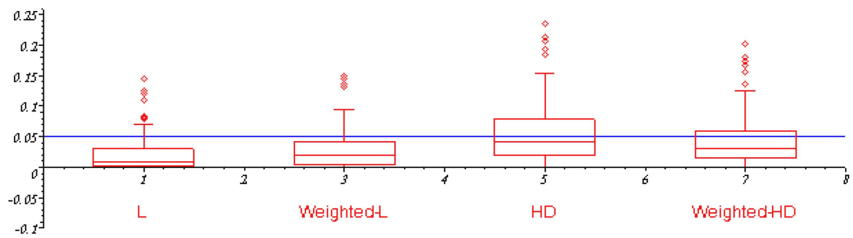


(c) $g = 0.2,\ h = 0$

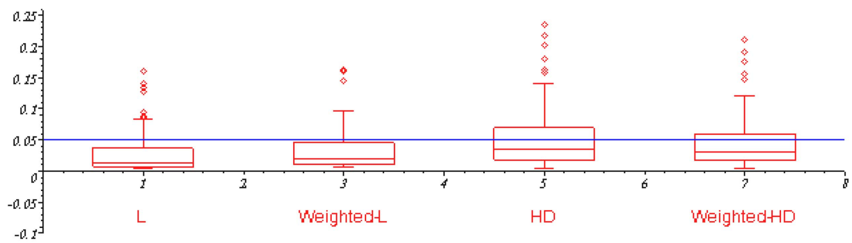


(d) $g = 0.2,\ h = 0.2$

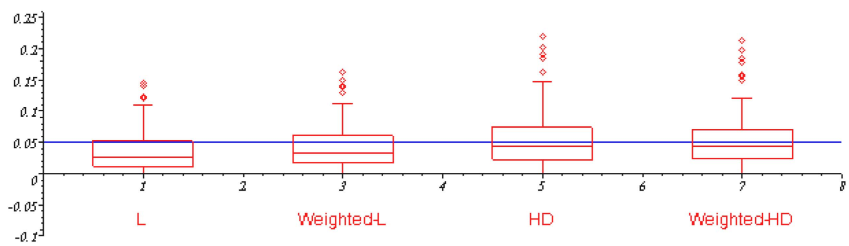
Figure A1. Box-plot of the standard error for estimating the quantiles $x_{\gamma/2}$, $x_{1-\gamma/2}$, $\alpha = \gamma = 0.05$, sample size $n = 30$; generated $m = 1{,}000$ times; bootstrapping $B = 100$.



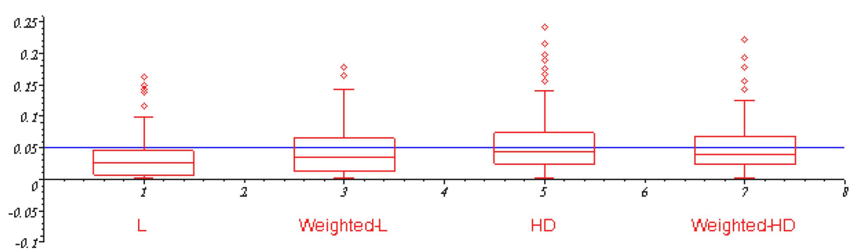
(a) $g = 0,\ h = 0$



(b) $g = 0,\ h = 0.2$



(c) $g = 0.2,\ h = 0$



(d) $g = 0.2,\ h = 0.2$

Figure A2. Box-plot of estimating α of confidence interval for predict interval, $\alpha = \gamma = 0.05$, sample size $n = 30$; generated $m = 1{,}000$ times; bootstrapping $B = 100$.

Table A1. Simulation efficiencies of estimating α and standard error of confidence interval for predict interval, $\alpha = \gamma = 0.05$, sample size $n = 30$; generated $m = 1000$ times; bootstrapping $B = 100$

True Distribution g	True Distribution h	Estimator and Efficiency		L	Weighted L	HD-bootstrap	Weighted HD-bootstrap
0	0	$\hat{\alpha}$	$\hat{\alpha}$	0.02186	0.02682	0.04351	0.04547
			SEFF $(\hat{\alpha})$	1	**1.1864**	**4.3359**	**6.2119***
		s_L	**SRMSE**	0.66064	0.45691	0.26549	0.26307
			SEFF	1	**1.4459**	**2.4884**	**2.5113***
0	0.2	$\hat{\alpha}$	$\hat{\alpha}$	0.02611	0.03343	0.04484	0.04653
			SEFF $(\hat{\alpha})$	1	**1.4418**	**4.6299**	**6.8847***
		s_L	**SRMSE**	1.72593	0.92465	0.64434	0.63007
			SEFF	1	**1.8666**	**2.6786**	**2.7393***
0.2	0	$\hat{\alpha}$	$\hat{\alpha}$	0.03466	0.04005	0.05132	0.05364
			SEFF $(\hat{\alpha})$	1	**1.5417**	**11.6212***	**4.2143**
		s_L	**SRMSE**	0.37738	0.29159	0.18569	0.18416
			SEFF	1	**1.2942**	**2.0323**	**2.0492***
0.2	0.2	$\hat{\alpha}$	$\hat{\alpha}$	0.03235	0.04006	0.04937	0.05108
			SEFF $(\hat{\alpha})$	1	**1.7757**	**28.0159***	**16.3426**
		s_L	**SRMSE**	1.42897	0.76046	0.49012	0.48093
			SEFF	1	**1.8791**	**2.9156**	**2.9713***

Note: (1) SEFF in **bold** is greater than 1: Proposed methods have 100% (24 out of 24 cases).

(2) $SEFF(\hat{\alpha}_{proposed}) = \left| \dfrac{SBias(\hat{\alpha}_L)}{SBias(\hat{\alpha}_{proposed})} \right|$, $SREFF(s_{proposed}) = \sqrt{\dfrac{SMSE(s_L^2)}{SMSE(s_{proposed}^2)}}$.

(3) Weighted HD-bootstrap has best efficiencies with "*": 75% (6 out of 8 cases).