Confidence interval for the $100p$-th percentile for measurement error distributions

Clarena Arrieta $^1$, Reinaldo Castro Souza$^2$, Rodrigo Flora Calili$^1$

$^1$Postgraduate Program in Metrology, Pontifical Catholic University of Rio de Janeiro, Márques de São Vicente Street, 225, Gávea, Rio de Janeiro, Brazil

$^2$Department of Metrology, Pontifical Catholic University of Rio de Janeiro, Márques de São Vicente Street, 225, Gávea, Rio de Janeiro, Brazil

E-mail: clare.a18@gmail.com, reinaldo@ele.puc-rio.br, calili@puc-rio.br

Abstract. In metrology, one is usually comparing what indicates an instrument versus what it should be indicated by a reference pattern, resulting in the so called measurement errors. Most errors follow a normal distribution. If one wants to have some confidence about how small is such an error, the traditional confidence intervals for the unknown parameters are computed. This article illustrates, via simulation, the effectiveness of three intervals for estimating the $100p$-th percentile in a normal distribution, namely, the Traditional (T), the Normal Asymptotic (NA), and the Lawless (LA) intervals. Various samples of different sizes are drawn from a zero mean normal distribution, which resembles the measurement error distributions. For each one of them the 95% confidence interval is estimated by using the three approaches. From the simulation, it is possible to conclude that T and LA intervals are better than NA as they produce shorter length and wider coverage.

Keywords: Metrology, Percentile, Confidence interval, Normal distribution.

1. Introduction

We are often dealing with data survey in measuring procedures and in most of the cases the data come from a normal distribution. Also, we are usually interested in obtaining point or intervals estimates containing some parameter of interest with a given probability which is named confidence level. In practice, we usually choose this level as 95% [1].

A confidence interval is a random interval $(a,b)$ obtained from some observed data $X_1, X_2, \ldots, X_n$, in principle different from sample to sample, that frequently includes the value of an unobservable parameter of interest if the experiment is repeated. The confidence level $1 - \alpha$ of the interval measures the probability of containing the true parameter, that is, if we repeatedly take random samples using the same method with a 95% confidence level. For one given sample, we do not know whether the confidence interval covers or not the true parameter. The 95% probability only refers to the method that it is used, but not to the individual sample. The value $\alpha$ represents the error obtained when one claims that the probability that the true parameter lies in the interval $(a,b)$ is $1 - \alpha$. In other words, as the estimates computed are random variables, one can expect that for some exceptional cases the interval do not contain the true parameter. It is proposed in this article confidence intervals for the $100p$-th percentile $\kappa_p = \mu + Z_p \sigma$, where $Z_p$ denotes the $100p$-th percentile from the Standard Normal distribution.
i.e. \( N(0, 1) \), based on the work of Chakraborti and Li [2]. The idea is to contribute to the development of confidence intervals that are very much used in metrology studies. It is used 3 intervals coming from 3 different methods:

(i) Maximum-likelihood leading to the minimum variance unbiased estimator, which in turn provides a natural pivotal quantity. Then by approximating critical values (using the \( t_{n-1} \) approximation), it is obtained the so called \( T \) interval.

(ii) Use of a pivotal quantity based on a studentized version of the sample 100\(p\)-th percentile (an order statistic), and the asymptotic normality of order statistics, called the asymptotic normality interval (NA).

(iii) A method proposed by Lawless, based directly on the maximum likelihood estimator (MLE); a biased estimator of \( \kappa_p \), called the Lawless interval (LA).

2. Proposed confidence intervals

Let \( X_1, X_2, \ldots, X_n \) be iid \( N(\mu, \sigma^2) \), and \( \kappa_p = \mu + Z_p \sigma \) as before. Chakraborti S. and Li, J. propose the following intervals:

- **The Traditional Interval** (T).

Let \( \bar{X} \) denote the sample mean and \( S = \left( \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \right)^{\frac{1}{2}} \) the sample standard deviation. Then the MLE of \( \kappa_p \) is \( \hat{K}_p = \bar{X} + Z_p S \) which is biased since \( E(\bar{X}) = \mu \) but \( E(S) = \bar{C} \), with

\[
C = C(n) = \left[ \frac{n-1}{2} \right]^{1/2} \left[ \Gamma \left( \frac{n}{2} \right) \right]^{-1} \Gamma \left( \frac{n-1}{2} \right)
\]

Therefore \( \hat{K}_p = \bar{X} + CZ_p S \) is unbiased for \( \kappa_p \) and its variance is given by

\[
Var(\hat{K}_p) = \frac{\sigma^2}{n} \left[ 1 + nZ_p^2 (C^2 - 1) \right]
\]

Finally using the \( t_{n-1} \) approximation we obtain the interval

\[
\bar{X} + CZ_p S \pm t_{\frac{1}{2}, n-1} \frac{S}{\sqrt{n}} \sqrt{1 + nZ_p^2 (C^2 - 1)}
\]  

(i) (2)

- **The Asymptotic Normality Interval** (NA).

This interval comes from the asymptotic normality of the order statistic \( X_{(r)} \) with \( r = np \), which is a natural point estimator of \( \kappa_p \) [3]. This interval is symmetrically placed like the T interval, but is centered at \( X_{(r)} \). The proposed confidence interval is given by

\[
X_{(r)} \mp t_{\frac{1}{2}, n-1} \sqrt{2p(1-p)} \left[ \frac{2\pi p(1-p)}{n} \right]^{1/2}
\]

(iii) (3)

- **The Lawless Interval** (LA).

Lawless [4] considered a confidence interval for \( \kappa_p \) based on the biased MLE \( \hat{K}_p = \bar{X} + Z_p S \), by considering a pivotal variable following a noncentral \( t \)-distribution with \( (n-1) \) degrees of freedom and noncentrality parameter \( -\sqrt{n}Z_p \) [5]. The obtained interval is

\[
\bar{X} \mp \frac{t_{1-\frac{1}{2}[n-1,(-\sqrt{n}Z_p)]} S}{\sqrt{n}}
\]

(iii) (3)
3. Evaluating the confidence intervals
In the literature, there are two main tools used to evaluate confidence intervals \((a, b)\), they are the length \(b - a\) and the coverage probability \(\text{Prob}(a \leq \kappa_p \leq b)\). Heuristically, if one wants to increase the probability that the interval contains the true parameter one must increase the length of the interval, and when the confidence level changes, say from 0.9 to 0.95, the confidence interval also changes. In general, the length of a confidence interval increases as the confidence level increases or as the standard error increases, but decreases as the sample size \(n\) increases. The same can not be said for the coverage, that is, there exists a trade-off between length and coverage probability of the interval.

The strategy, as usual, is choosing the interval whose length is as small as possible and whose coverage probability is at least equal to the confidence level.

4. Simulation study
It is shown via simulation of different scenarios the comparison of the confidence intervals proposed by Chakraborti and Li [2]. It is used the statistical software R-studio.

4.1. Procedure
Six random samples of sizes \(n_1 = 25, n_2 = 50, n_3 = 100, n_4 = 200, n_5 = 500, n_6 = 1000\) were considered. For each sample size \(m = 1000, m = 2000\) and \(m = 5000\) random sample replications were generated according to the standard normal distribution \(N(0, 1)\) and the confidence level used was 0.95. We remark that larger sample sizes do not reduce the inherent variability, but lead to narrower confidence intervals that are a basis for estimating the random component of uncertainty. The considered percentiles are \(p_1 = 0.25, p_2 = 0.5, p_3 = 0.9\).

For the intervals T, NA and LA, denote LT, LNA and LLA as the average lengths, and CT, CNA and CLA as the coverages, respectively.

4.2. Simulation results
The results of the study are summarized below in tables 1, 2 and 3 for \(p = 0.25, 0.5\) and 0.9 respectively. It is registered the average length and the coverage probabilities when the confidence level is 0.95. The highlighted lines in bold represent the best choice for length and coverage, and, of course, this is due to the fact that the biggest sample size was chosen, in fact when \(n = 500\) and \(n = 1000\) the coverage probabilities were very close to 95% and shorter lengths, the key point is that is does not depend on the number of replications. Now, for small sample sizes there is no evidence to say that the coverage is close to 95%.
Table 1. Lengths and coverages when the confidence level is 0.95% and $p = 0.25$.

| $n$ | Replicas | Length | Coverage |
|----|----------|--------|----------|
|    |          | LT     | LNA      | LLA      | CT       | CNA      | CLA      |
| 25 | 1000     | 0.3084 | 1.1115   | 0.9103   | 95.30    | 94.60    | 95.40    |
| 50 |          | 0.6288 | 0.7715   | 0.6295   | 95.70    | 95.50    | 95.90    |
| 100 |        | 0.3087 | 0.3794   | 0.3087   | 94.60    | 95.70    | 94.60    |
| 500 |        | 0.1945 | 0.2392   | 0.1945   | 94.40    | 95.30    | 95.30    |
| 1000 |       | 0.1375 | 0.1691   | 0.1375   | 95.30    | 95.45    | 95.10    |
| 25 | 2000     | 0.9138 | 1.1185   | 0.9160   | 94.30    | 94.40    | 94.35    |
| 50 |          | 0.6298 | 0.7728   | 0.6305   | 95.85    | 95.75    | 95.35    |
| 100 |         | 0.4114 | 0.5423   | 0.4117   | 94.70    | 94.60    | 94.90    |
| 200 |         | 0.3092 | 0.3600   | 0.3092   | 93.30    | 94.85    | 94.25    |
| 500 |         | 0.1946 | 0.2393   | 0.1946   | 94.55    | 95.45    | 94.90    |
| 1000 |        | 0.1375 | 0.1692   | 0.1376   | 94.90    | 94.75    | 94.78    |

Table 2. Lengths and coverages when the confidence level is 0.95% and $p = 0.5$.

| $n$ | Replicas | Length | Coverage |
|----|----------|--------|----------|
|    |          | LT     | LNA      | LLA      | CT       | CNA      | CLA      |
| 25 | 1000     | 0.8101 | 1.0153   | 0.8101   | 94.70    | 93.20    | 94.70    |
| 50 |          | 0.5665 | 0.7125   | 0.5685   | 95.90    | 95.40    | 95.90    |
| 100 |        | 0.3960 | 0.4963   | 0.3960   | 95.80    | 95.00    | 95.80    |
| 200 |         | 0.2781 | 0.3486   | 0.2781   | 95.20    | 95.60    | 95.20    |
| 500 |         | 0.1757 | 0.2202   | 0.1757   | 95.10    | 95.10    | 95.10    |
| 1000 |       | 0.1240 | 0.1554   | 0.1240   | 95.00    | 95.00    | 95.00    |
| 25 | 2000     | 0.8137 | 1.0198   | 0.8137   | 95.45    | 94.05    | 95.45    |
| 50 |          | 0.5653 | 0.7095   | 0.5653   | 95.05    | 95.05    | 95.05    |
| 100 |         | 0.3947 | 0.4947   | 0.3947   | 95.20    | 95.30    | 95.20    |
| 200 |         | 0.2783 | 0.3488   | 0.2783   | 95.45    | 95.45    | 95.45    |
| 500 |         | 0.1756 | 0.2201   | 0.1756   | 95.60    | 95.55    | 95.60    |
| 1000 |        | 0.1240 | 0.1555   | 0.1240   | 95.45    | 95.55    | 95.45    |

Table 3. Lengths and coverages when the confidence level is 0.95% and $p = 0.9$.

| $n$ | Replicas | Length | Coverage |
|----|----------|--------|----------|
|    |          | LT     | LNA      | LLA      | CT       | CNA      | CLA      |
| 25 | 1000     | 1.1093 | 1.3883   | 1.1179   | 95.20    | 92.20    | 95.20    |
| 50 |          | 0.7644 | 0.9627   | 0.7669   | 92.60    | 93.60    | 93.60    |
| 100 |        | 0.5348 | 0.6755   | 0.5355   | 94.10    | 93.90    | 94.10    |
| 200 |         | 0.3770 | 0.4768   | 0.3773   | 94.70    | 95.20    | 95.00    |
| 500 |         | 0.2373 | 0.3064   | 0.2373   | 94.50    | 94.40    | 94.80    |
| 1000 |        | 0.1675 | 0.2120   | 0.1675   | 95.70    | 94.60    | 96.10    |
| 25 | 2000     | 1.1174 | 1.3990   | 1.1225   | 95.30    | 91.55    | 95.35    |
| 50 |          | 0.7674 | 0.9664   | 0.7699   | 94.85    | 94.10    | 94.85    |
| 100 |        | 0.5369 | 0.6782   | 0.5378   | 95.30    | 94.45    | 95.45    |
| 200 |         | 0.3765 | 0.4763   | 0.3768   | 95.35    | 94.80    | 95.40    |
| 500 |         | 0.2370 | 0.3000   | 0.2371   | 95.60    | 94.15    | 95.65    |
| 1000 |        | 0.1674 | 0.2119   | 0.1674   | 94.50    | 94.85    | 94.75    |
| 25 | 5000     | 1.1149 | 1.3958   | 1.1227   | 94.70    | 91.40    | 94.46    |
| 50 |          | 0.7681 | 0.9673   | 0.7706   | 95.58    | 94.88    | 95.56    |
| 100 |        | 0.5357 | 0.6766   | 0.5365   | 95.08    | 94.32    | 94.94    |
| 200 |         | 0.3762 | 0.4759   | 0.3765   | 95.42    | 95.08    | 95.24    |
| 500 |         | 0.2372 | 0.3003   | 0.2373   | 95.36    | 95.08    | 95.24    |
| 1000 |        | 0.1675 | 0.2121   | 0.1676   | 95.12    | 94.92    | 95.08    |
In order to visualize better, the results are displayed in figure 1 below. From this figure 1, it is possible to see that in terms of length, the worst interval is NA as it has a wider length, but the other two have quite similar behavior. Moreover, figure 2 shows that the T and LA intervals have better coverage than the NA since they have less variability in coverage and are concentrated around the level of significance.

![Figure 1](image1.png)

**Figure 1.** Length vs Sample size.

![Figure 2](image2.png)

**Figure 2.** Coverage vs Sample size.
5. Conclusion

Based on the results obtained by simulation, in this study, it is possible to conclude that the least efficient interval among all three considered in this paper is the NA interval. Therefore, we recommend to use the T and LA intervals which have a better performance in the sense that they have shorter length and wider coverage.

However, for theoretical and practical aspects related to measurements it should be considered all possibilities contributing to the development of confident results in metrology applications. For example, the quality of measurement is quantified by means of evaluation of uncertainty of such a measurement [6], in particular when one is dealing with uncertainty problems where data follows a normal distribution. In this case, one can easily work with the intervals T and LA, as they seem to produce better and satisfactory results. For future works it is recommended the use the proposed methodology applied in real data rather than simulated data.

References

[1] ABNT 2010 Estatística-Vocabulário e símbolos Parte 1: Termos estatísticos gerais e termos usados em probabilidade NBR ISO, RJ 3534-1 69
[2] Chakraborti S and Li J 2007 Confidence Interval Estimation of a Normal Percentile Amer. Statist. 61 331-336
[3] Gibbons J D and Chakraborti S 2003 Nonparametric Statistical Inference New York: Marcel Dekker 168 xxiv+645
[4] Lawless, Jerald F 2003 Statistical Models and Methods for Lifetime Data Wiley-Interscience [John Wiley & Sons], Hoboken, NJ 2 xxii+630
[5] Owen D B 1968 A Survey of Properties and Applications of the Noncentral t-Distribution, Technometrics 10 445-478
[6] Pérez H M 2012 Estimación de incertidumbres Guía GUM Revista Española de Metrología 13 114-130.